bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Noname manuscript No.
(will be inserted by the editor)

Impact of domain knowledge on blinded predictions of binding energies by alchemical free energy calculations

Antonia S J S Mey · Jordi Juárez Jiménez · Julien Michel

Received: date / Accepted: date

Abstract Keywords D3R · computer-aided drug design · protein-ligand interactions · alchemical free energy calculations

Within the context of the second D3R Grand Challenge several blinded binding free energies predictions were made for two congeneric series of FXR inhibitors with a semi-automated alchemical free energy calculations workflow featuring the FESetup and SOMD tools. Reasonable performance was observed in retrospective analyses of literature datasets. Nevertheless blinded predictions on the full D3R datasets were poor due to difficulties encountered with the ranking of compounds that vary in their net-charge. Performance increased for predictions that were restricted to subsets of compounds carrying the same net-charge. Disclosure of X-ray crystallography derived binding modes maintained or improved the correlation with experiment in subsequent rounds of predictions. The best performing protocols on D3R set1 and set2 were comparable or superior, to predictions made on the basis of analysis of literature SARs only, and comparable or slightly inferior, to the best submissions from other groups.

EaStCHEM School of Chemistry, University of Edinburgh, David Brewster Road, Edinburgh EH9 3FJ, United Kingdom
E-mail: mail@julienmichel.net Tel.: +44 (0)131 650 4797

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

1 Introduction

There is growing interest in the routine use of alchemical free energy (AFE) calculations for predictions of protein-ligand binding energies in structure-based drug discovery programs. [1–7] In particular building on pioneering work over three decades ago, [8,9] some modern alchemical relative free energy calculation protocols achieve in several diverse protein binding sites sufficiently accurate predictions of binding energies (RMSD under 1.5 kcal·mol^{-1}; R ca. 0.7 or better) to speed up hit-to-lead and lead optimisation efforts [10]. In favourable cases, AFE calculations can even reproduce subtle non-additivity of structure-activity relationships [11]. However, for a given set of protein-ligand complexes it remains difficult to anticipate the predictive power of AFE calculations. Uncertainties in binding modes [12–14] protonation/tautomeric states [15,16], binding site water content [17–19], and choice of potential energy functions [20,21], can profoundly influence the outcome of such calculations. Accordingly, there is much interest in defining as much as possible a domain of applicability for the technology [22].

Blinded predictions competitions, whereby participants submit physical properties computed by a model in the absence of knowledge of the actual experimental data, have been instrumental in driving methodological progress in a wide range of scientific fields [23–26]. Blinded predictions reduce the impact of unconscious biases on the design of protocols, and allow evaluation of molecular modelling methods in a context closer to their intended use in drug discovery. This is advantageous for academic groups that have expertise in computational methodologies, but lack resources to carry out prospective studies. It is also beneficial for the field to evaluate different methodologies applied to the same dataset with identical analysis protocols.

This report focuses on the predictions submitted by our group within the context of the second Drug Design Data Resource (D3R) Grand Challenge, that

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

ran between September 2016-February 2017.This complements previous reports from our group on blinded predictions of protein-ligand poses, rankings, binding free energies, [27], distribution coefficients [28], and host-guest binding free energies [29], as part of preceding D3R and SAMPL challenges [30,31]. The second D3R grand challenge focussed on a dataset of 102 inhibitors of the Faresnoid X receptor (FXR) provided by Roche. The competition featured poses predictions, dataset rankings, and relative binding free energy predictions for two subsets of 15 and 18 compounds referred to as set1 and set2 respectively. Our group only submitted predictions of relative binding free energies for the set1 and set2 subsets. Submissions were made before (stage1) and after (stage2) information about binding poses of representative set1 or set2 compounds was made available. This enabled an analysis of the impact of the available experimental data on the performance of the protocols. All input data download and submissions upload were conducted via the website of the D3R consortium [32].

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

2 Theory and Methods

2.1 Datasets

2.1.1 Blinded datasets

At the start of the challenge (stage1), the organisers released the pseudo apo-protein structure of ligand **10** as provided by Roche, as well as 36 ligand in SDF format to be used for the prediction of crystallographic poses, and an additional 66 ligands that should be used in affinity rankings. There were two subsets identified among these 102 ligands, set1 with 15 compounds and set2 with 18 compounds, for which relative binding free energies could be calculated. Ligand subsets set1 and set2 are depicted in Figure SI1. For the second stage of the challenge 36 X-ray structures were released, meaning that they could be used to prepare input files for alchemical free energy calculations. Once the competition was over a set of IC50 data for the entire dataset was released. The data stems from a scintillation proximity assay using only the FXR binding domain and a radioactive tracer. More information on the experimental binding assay as well as a study on other FXR inhibitors can be found in a series of publications [33–36]. Experimental relative binding free energies were estimated by eq. 1

$$\Delta\Delta G_{L1\rightarrow L2} = k_B T \ln \frac{IC_{50(\mathrm{L2})}}{IC_{50(\mathrm{L1})}}, \tag{1}$$

where L1 and L2 represent two ligands for which a relative energy difference is computed and k_B and T are the Boltzmann constant and temperature respectively.

2.1.2 Literature datasets

In order to test the computational protocols before submission of blinded predictions, retrospective studies were carried out using available literature data. A set of inhibition and structural data for 3-aryl isoxazole analogs of the non-steroid agonist GW4064 has been previously published [34,36]. The data consists of two

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

different ligand series, where the first series contains 8 compounds (LitSet1) and the second series 17 (LitSet2). The same experimental IC50 assay as described for the blinded dataset was used. Relative binding free energies were computed from the reported IC50s with equation 1. A summary of the compounds present in LitSet1 and LitSet2 can be found in the SI.

2.2 Methods

The methodology used for the calculations of relative binding free energies of FXR ligands was single topology molecular dynamics alchemical free energy. Several operations are necessary to produce from an input set of protein atom coordinates and 2D descriptions of ligands, a set of output relative free energies of binding. Currently this is implemented by a semi-automated workflow as depicted in figure 1.

2.2.1 Docking input preparation

For the two sets of literature data crystal structure with PDB ID 3DTC was used for the ligands taken from Feng et al [34], and crystal structure with PDB ID 3XVF for data taken from Richter et al [36].

For stage 1 of the competition protein and ligands were prepared as such: The FXR structure provided by the organizers was used as an initial template. The structure was converted to a mol2 file using fconv for the docking calculations, after initial cleaning with Maestro 11 (beta) [37]. For the alchemical free energy calculations it was necessary to model a missing region consisting in residues A459-K464. To this end the protein residues between N448 and Q476 were replaced by the loop observed in X-ray structure with PDB ID:3OKH. Subsequently, ACE capping groups were added to residues M247 of the main chain and D743 of the co-activator fragment. Similarly an NME capping group was attached to D755 of the co-activator fragment. Ligand 3D structures were generated from 2D sdf

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

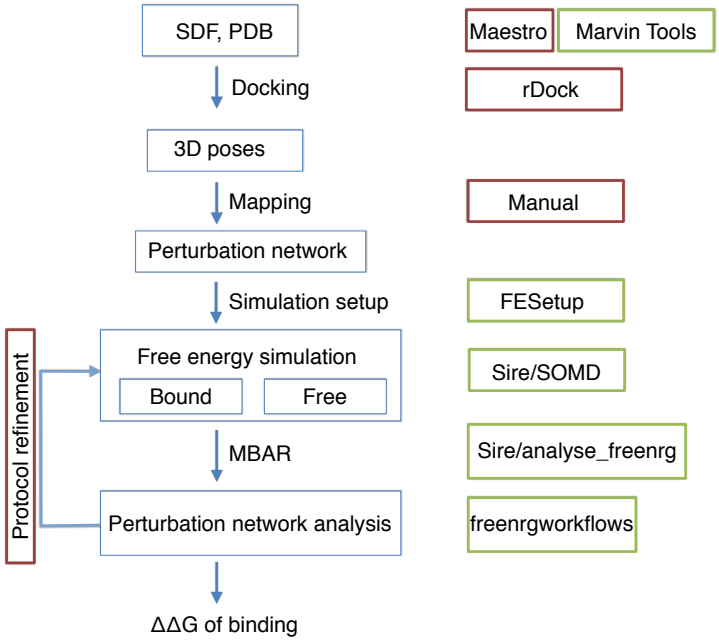

Fig. 1 Semi-automated workflow for predicting relative free energies of binding. Workflow operations are depicted by blue boxes. Green boxes denote software available for automated execution of the workflow step. Red boxes denote operations that require human intervention.

files provided, using MarvinTools scripts, from Marvin Sketch 15.3.30. No water molecules were retained for the docking calculations.

2.2.2 Docking

No automated docking calculations were performed for the literature data. Instead compounds were manually build by overlay with the binding mode of compound GW4064 as observed in the X-ray crystal structure 3DTC. Docking calculations were performed for compounds **91**, **101** and **102**. Docking was performed with rDock [38], generating the cavity using the two sphere method available in the program, centering a 15 Å cavity within residues M294, I356, S336 and Y373 using a 1.5 and 4.0 Å for the radius of small and large spheres respectively. An aromatic ring pharmacophoric restraint was enforced within 4 Å of the the center of the cavity. Coincident binding modes were obtained for the three compounds. However,

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

to minimize the differences between binding modes of the different compounds, poses for all compounds in a set were derived from a common scaffold. To this end, the binding mode of the largest compound **102** was selected as a template and subsequently modified in Maestro to obtain poses that were energy minimized to avoid steric clashes.

For the alchemical free energy calculations coordinates for water molecules accompanying the X-ray structure provided by the organisers were superimposed with the coordinates of the poses. It was found that 1 water molecule could clash with the larger compounds (such as e.g. **102**). Consequently, it was manually displaced to a nearby position. The main difference between stage 1 and stage 2 is that once the crystal structure poses were released after stage 1 the additional knowledge gained from the crystal structures guided the preparation of new input files with different binding modes for the alchemical free energy calculations.

2.2.3 Alchemical calculations input preparation

Once a set of satisfactory 3D poses for both set1 and set2 was obtained, a relative free energy perturbation network was manually designed for both set1 and set2 ligands. set1 included one ambiguous binding mode for compound **47**. For set2 only three of the 18 compounds had a clearly preferred binding mode. Typically there was uncertainty in the position of ortho or meta substituents of a benzyl ring. Whenever there was ambiguity, the different binding modes were included in the perturbation map. With the perturbation networks defined, the FESetup [39] release 1.3dev, was used to parametrise set1 and set2 ligands, setup ligands in a water box as well as protein environment and create the needed input for the alchemical free energy simulations.

Ligands

Ligands were parametrised using the generalised amber force field 2 (GAFF2) [40], followed by solvation in a rectangular box of 12 Å length using TIP3P water [41, 42]. An energy minimisation using a steepest decent algorithm with 500 steps was

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

carried out on the water box, followed by an NVT simulation with the ligand restrained, during which the system was heated to 300 K over 1000 steps. Next an NPT equilibration at 1 atm was run for 5000 steps, followed by the release of the restraint on the ligand over 500 steps. FESetup used pmemd for this equilibration. For each perturbation a SOMD compatible perturbation file was then created from the perturbation map produced FESetup.

Protein ligand complex

For the protein and ligand complex the protein and previously parametrised ligand were combined and solvated in a rectangular box of 10 Å. The protein forcefield was the amber 14 SB forcefield [40]. An equivalent solvation and equilibration protocol was used as described for the solvated ligand only.

2.2.4 Alchemical free energy simulations

The alchemical free energy protocol used here is based on the SOMD software as available in the Sire 2016.1.0 release [43]. This version of SOMD is linked with OpenMM 7.0.1 [44] that provides a CUDA compatible integrator in order for simulations to be run on a cluster of GPUs.

Details about the theoretical background are available elsewhere [6, 45, 7, 4, 46–48, 10]. The main idea behind alchemical free energy calculation is to avoid direct computation of the free energy change associated with the reversible binding of a ligand to a protein. Instead one computes the free energy change for artificially morphing a ligand (L_1) into another ligand (L_2). Repeating this process for L_1 and L_2 in aqueous solution or bound to the protein of interest enables construction of a thermodynamic cycle that yield the relative binding free energy of the two ligands.

Each alchemical free energy calculation for a pair of ligands L_1 and L_2 consisted minimally of one forward (L_1 to L_2) an one backward (L_2 to L_1) computation. Ligand pairs that showed poor agreement between forward and backwards simulation were repeated up to three times. Mean free energies and standard er-

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

ror were estimated from the resulting distributions of computed relative binding free energies. Further details are provided in the SI [49]. All simulations shared the following common set of parameters. Each simulation box was treated with periodic boundary conditions and simulations were run for 4 ns each using a 2 fs integration timestep with a Leap-Frog-Verlet integrator. Bonds involving hydrogens were constrained, except if the hydrogen atom was morphed to a heavy atom in the perturbation. The temperature was maintained at 298 K using an Anderson thermostat and a collision frequency of 10 ps^{-1} with velocities initially drawn from a Maxwell-Boltzmann distribution of that temperature. Pressure was kept at 1 atm using the Monte Carlo Barostat implemented in OpenMM with an update frequency of 25 MD steps. For non-bonded interactions an atom-based shifted Barker-Watts reaction field scheme was used with a cutoff of 10 Å and the reaction field dielectric constant $\epsilon = 82.0$. The number of λ windows for each simulation varied for different perturbations and a summary, as well as complete simulation parameters can be found in the SI. All input files are available on a github repository [50].

2.2.5 Free energy analysis

Free energy changes were estimated with the multi state Bennet's acceptance ratio (MBAR) estimator, as implemented in pymbar (v 3.0.0 beta 2) [51], which was integrated into the Sire app *analyse_freenrg*. The first 5% of the trajectories were discarded to allow for equilibration before MBAR analysis. The individually estimated free energy differences were then read into a networkx (v 1.11) digraph [52]. Perturbation for morphing L_1 to L_2 and L_2 to L_1 were both simulated and resulting binding free energies were averaged for the forward and (reversed) backward perturbations. When available, averages were calculated across multiple independent repeats. The error estimated between repeated runs of backwards/forwards simulations served as the estimated error for each averaged network edge. Binding free energies relative of a ligand L_i to a reference compound L_0 were then

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

estimated by enumerating all possible paths connecting L_i to L_0 in the network. The relative binding free energy and its uncertainty along a given path was obtained by summing relative binding free energies along each edge of the path and propagating errors. Next, a weighted average of all paths was produced, with the contribution of each path weighted by its estimated statistical error. Thus paths that have smaller statistical errors contribute more than paths that show larger statistical errors.

2.2.6 Charge scaling correction

Initial analysis of literature datasets (see results) suggested that polarisation effects may play a significant role in FXR ligand binding energetics. While no polarisable force-field was readily available to test this hypothesis, there has been some success in capturing polarisation effects in protein-ligand binding by QM/MM reweighting of trajectories computed with a classical potential energy function [53]. Given the time constraints posed by the competition, no such methodologies were used here. Rather, an *ad hoc* protocol based on empirical scaling of ligand partial charges was implemented.

Thus the corrected free energies were given by:

$$\Delta\Delta G_{\mathrm{bind,scale}}(L_1, L_2) = \Delta\Delta G_{\mathrm{bind}}(L_1, L_2) + \Delta\Delta G_{\mathrm{scaled}}(L_1, L_2), \tag{2}$$

where $\Delta\Delta G_{\mathrm{scaled}}$ is given by:

$$\begin{aligned}\Delta\Delta G_{\mathrm{scaled}}(L_1, L_2) = \Delta G_{\mathrm{scaled}}(L_1, \mathrm{free}) + \Delta G_{\mathrm{scaled}}(L_2, \mathrm{bound}) - \\ \Delta G_{\mathrm{scaled}}(L_1, \mathrm{bound}) - \Delta G_{\mathrm{scaled}}(L_2, \mathrm{free}),\end{aligned} \tag{3}$$

where the ΔG_{scaled} values are the free energy changes for scaling the partial charges of a ligand L_1 or L_2 in water, or bound to FXR. Such quantities were evaluated via MBAR analysis of trajectories for 5 evenly spaced λ windows sam-

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

pled for 1 ns each. Scaling factors ranged from 1 (no scaling) to 0.5 (50% decrease in magnitude of partial charges).

2.3 Errors analysis

For the comparison of computed binding free energies and experimental binding free energies two measures are mainly used in the analysis, Pearson R and mean unsigned error (MUE). To obtain an error estimate for both these measures a bootstrapping approach is used in which the mean and standard error of each of the computed free energy estimates serve as the mean and standard deviation of a Gaussian distribution. For each estimate a new value from this Gaussian distribution is drawn until a new artificial distribution of computed free energies is drawn. This resampled distribution is then correlated to the experimental data. Repeating the process 10,000 times gives rise to a distribution of MUE and R, for which a mean and 1 σ confidence interval can be computed. This was the default protocol used by our group to estimate metric errors. The organisers, however, chose a different way of estimating errors in the data sets to facilitate comparison between different submissions. This approach uses bootstrapping of the dataset, for which data points (both experimental and computed) are resampled with replacement until an artificial dataset of the same size is created. This is repeated 1,000 times, leading to a distribution for Pearson R with 1 σ confidence intervals. All error bars in figure 5 have been generated in this fashion.

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

3 Results

3.1 Literature datasets

The robustness of the computational protocol was first tested with the two literature datasets LitSet1 and LitSet2. Supplementary figures 1 and 2 depict the perturbation network used for LitSet1 and LitSet2 respectively [49]. A summary of the results of the comparison of the calculated and measured binding free energies is given in table 1. While the correlation between LitSet1 computational and experimental data with R=0.84±0.05 was deemed satisfactory, the mean unsigned error (MUE) at 3.0±0.2 kcal·mol^{-1} was judged unexpectedly large. For the second dataset LitSet2 the overall correlation R=0.56±0.03 is lower, however the MUE is significantly lower at 1.7±0.1 kcal·mol^{-1}.

Analysis of the pairwise alchemical free energy calculations in the two datasets suggested that calculated binding free energy changes for perturbations that involve substitution of a non polar group by a polar group were overly exaggerated with respect to experimental data. Also, LitSet2 contained one negatively charged compound (**1R**, carboxylic acid) which was predicted to be ca. 30 kcal·mol^{-1} less potent than its -H counterpart, whereas experimental data suggests weaker binding of the acid by ca. +2.6 kcal·mol^{-1}.

The binding site of FXR is rather apolar (see figure 2A), and it was hypothesized that changes in ligand polarisation upon transfer from bulk to the FXR binding site may play a significant role. This prompted the development of an *ad hoc* protocol in an attempt to capture polarisation effects as described in the methods section via introduction of a set of charge scaling factors. The resulting correlation coefficient and MUE for scaling factors between 1 to 0.5 are also displayed in table 1. Figure 2A in the SI displays the correlation between the computed and experimental results, and figure 2B of the SI summarises the effect of changing the scaling corrections from 0.9 to 0.5 of the original charge. It was found that a scaling factor of 0.7 was the best tradeoff to minimize MUE whilst

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Dataset	scaling factor	R	MUE / kcal∘mol^{-1}
LitSet1	1.0	0.84± 0.05	3.0 ± 0.2
	0.9	0.83± 0.04	2.45± 0.18
	0.8	0.81 ± 0.04	2.24± 0.23
	0.7	0.78 ± 0.08	1.8± 0.15
	0.6	0.56 ± 0.09	2.2 ± 0.2
	0.5	0.51 ± 0.1	1.4 ± 0.2
LitSet2*	1.0	0.56 ± 0.05	1.77 ± 0.08
	0.9	0.54 ± 0.05	1.54 ± 0.09
	0.8	0.51 ± 0.05	1.46 ± 0.08
	0.7	0.44 ± 0.06	1.47 ± 0.09
	0.6	0.37 ± 0.06	1.61 ± 1.7
	0.5	0.23 ± 0.07	1.75 ± 0.08

Table 1 Summary of test dataset based on GW4064 compounds. *Charged compound **1R** has been omitted from the analysis.

maintaing a reasonable Pearson correlation coefficient. The effects are more pronounced for LitSet1. The one exception is the charged compound **1R** in LitSet2, for which reasonable agreement with experimental data required a scaling factor of 0.5.

Given time-constraints no further efforts were devoted to the literature datasets, and subsequent blinded submissions were made for protocols without charge scaling correction, or with a charge scaling correction of 0.7 for free energy perturbations that maintain net-charge, and 0.5 if the net-charge varies in the perturbation.

3.2 Blinded dataset

For the first stage of the competition binding modes for the FXR ligands in set1 and set2 had to be predicted by analysis of available crystal structures, or docking calculations as described in the methods section. Figure 2B top panel shows the structure of set1 ligand **17**, and the bottom panel depicts predicted and later disclosed binding modes. The RMSD is 0.9$\AA$ only, and the binding mode prediction can be considered successful. For set2 the X-ray crystal structure of **10** was later disclosed. The predicted binding mode deviates more, whereas at 2.5$\AA$ the RMSD

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

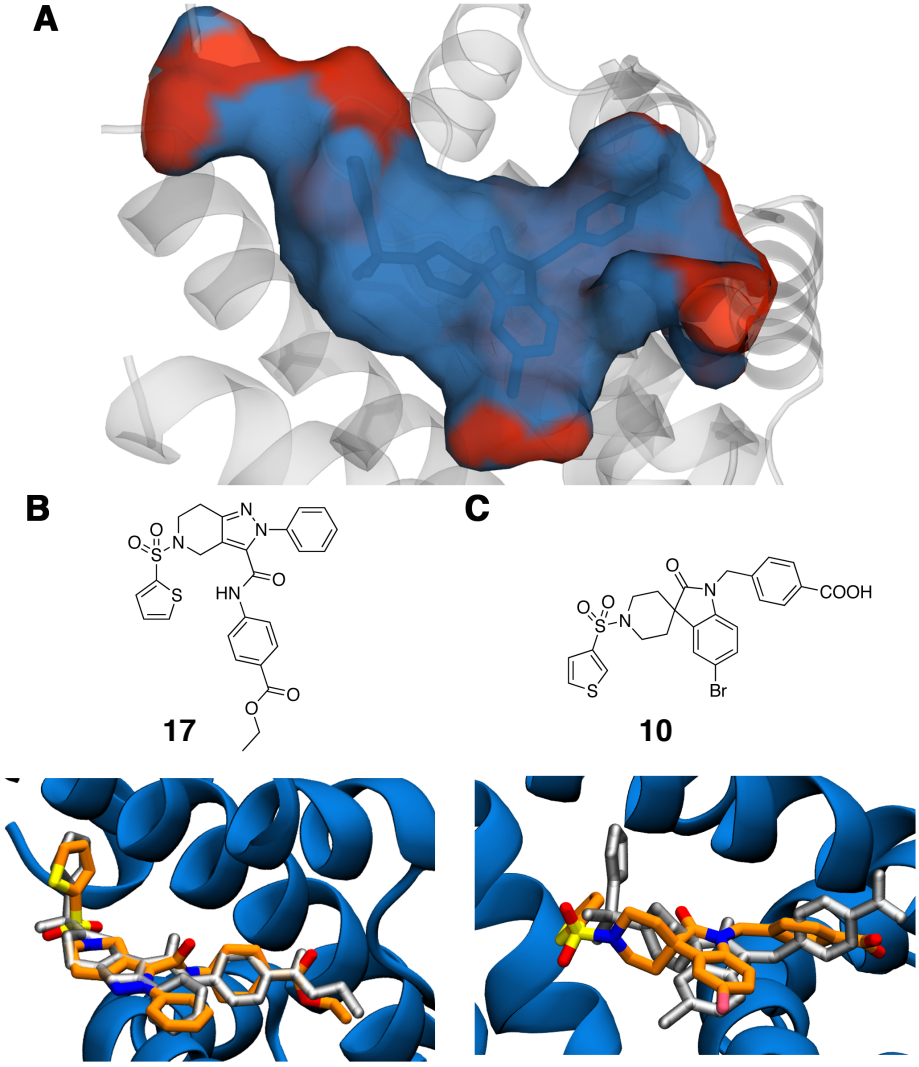

Fig. 2 A: Depiction of the FXR binding site, with hydrophilic residues shown in red and hydrophobic residues shown in blue. B: Compound **17** from set1, and predicted (orange sticks) versus observed (grey sticks) binding modes. C: Compound **10** from set2, and predicted (orange sticks) versus observed (grey sticks) binding modes.

is not exceptionally high, the thiophene ring has been positioned differently to the X-ray pose. This was of concern as many of the set2 compounds feature variations in aryl sulfonamide groups.

Table 2 shows results for the protocols submitted at stage 1 of the competition. The *expert opinion full* protocol was a submission where binding energies were

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Dataset	Protocol	ID	R	MUE / kcal·mol^{-1}
Set1	expert opinion full	pbjwu	0.16±0.04	1.71±0.08
	expert opinion same charge	N/A	0.17±0.04	1.76±0.09
	full	a3c8k	0.16±0.1	2.6±0.1
	full guided	bolbu	0.14±0.03	1.9±0.04
	same charge	0psyy	0.29±0.04	1.9±0.1
Set 2	expert opinion full	fxtpq	0.23±0.05	1.79±0.09
	expert opinion same charge	N/A	0.03±0.05	1.72±0.1
	full	qvnq5	-0.57±0.04	3.18±0.22
	same charge	f6een	0.2±0.1	1.35±0.12

Table 2 Performance of the protocols submitted at stage 1 of the D3R competition.

predicted by one of the authors (JM) by analysis of literature structure-activity relationships and visualisation of predicted binding modes. The *expert opinion same charge* protocol was not submitted but is presented to facilitate comparison with other protocols. The *full* protocol was a submission on the full dataset analysed as described in methods. The *full guided* protocol was a submission where only a small number of pathways in the perturbations network were hand-picked by one of us (JM) to evaluate binding free energies for the dataset. This was only done for set1. Finally the *same charge* protocol was a submission of alchemical free energy predictions restricted to the largest subset of compounds with the same net-charge.

For the second stage of the competition, calculations were repeated from a new set of poses for set2 compounds. set1 poses were the same as in stage 1. Additionally individual perturbations were categorised as 'easy', 'medium' or 'difficult' on the basis of the precision of the calculated relative binding free energies obtained at stage 1, and this led to lambda schedule protocols with less, the same amount of, or more, windows as in stage 1 (see SI for details). The time left in the competition was used to carry out multiple repeats of the perturbations that showed higher statistical errors. Additionally, the optimisation of charge scaling factors on the literature datasets had completed then, and scaling factor corrections were also applied to set1 and set2 datasets. Table 3 shows the results for protocols

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Dataset	Protocol	ID	R	MUE / kcal· mol^{-1}
Set1	full	07tpe	0.13 ± 0.02	2.20 ± 0.08
	same charge	olv52	0.3 ±0.02	1.41 ± 0.08
	full scaled	inspj	0.12 ±0.02	2.32 ± 0.09
	same charge scaled	4botu	0.13 ± 0.02	2.4 ± 0.1
Set2	full	qt771	-0.44 ± 0.02	3.79 ± 0.1
	same charge	0jz8u	0.54 ± 0.03	1.67 ± 0.08
	full scaled	jzrt5	0.41 ± 0.06	1.65 ±0.09
	same charge scaled	c1nbt	0.41 ± 0.05	1.56±0.1

Table 3 Performance of the protocols submitted at stage 2 of the D3R competition.

submitted at stage 2 of the competition. Only *full* and *same charge* protocols, and their scaled variants, were submitted.

At stage 1 the *expert opinion* protocol shows R values for both set1 and set2 of ca. 0.2, and MUE values ca. 1.7 kcal· mol^{-1}. The performance is similar or worse for the *expert opinion same charge* protocol. Alchemical free energy based protocols on the full dataset fare poorly with similar or lower R values, and higher MUE values. Submissions that only considered compounds with the same net charge show better performance (R ca. 0.2-0.3, MUE ca. 1.4-1.9 kcal· mol^{-1}). Overall none of the protocols show satisfactory correlation with experiment.

At stage 2 of the competition, the *full* and *same charge* submissions show lower statistical errors because the additional repeats calculations on the noisier perturbations have improved convergence. For set1 the MUE decreases, but the R metric is no different from stage 1 submissions. The scaled submissions for the full dataset and the same charge dataset achieve similar R values, but the MUE has worsened. For set2 lower statistical errors are also observed with respect to stage 1. The *full* submission produces similarly low R values and high MUE values. However the *full scaled* submission significantly increases R from ca. -0.4 to +0.4, while decreasing MUE from ca. 3.8 to 1.6 kcal· mol^{-1}. The *same charge* submission shows a significant increase in R with respect to stage 1 (from ca. 0.2 to ca. 0.5), but the MUE increases from 1.3 to 1.7 kcal· mol^{-1}. Finally, the *same charge scaled* protocol achieves a poorer R value (ca. 0.4) and similar MUE value.

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Overall the most significant improvement at stage 2 is observed for the set2 same charge dataset. This could be because the predicted binding modes at stage 1 were not in agreement with the subsequently disclosed X-ray structures. The scaling protocol appears to yield large improvements on set2, but this actually comes at the expense of decreased predictive power for the subset of compounds that carry the same net-charge (see below).

Figure 3 depicts detailed results for the full dataset of set1 compounds at different stages of the competition. Figure 3A shows at stage 1 the relative binding free energy of charged compound **101** is significantly overestimated with respect to all other neutral compounds . Compound **45** is also a significant outlier. Figure 3B shows that at stage 2 there is a trend towards better agreement with experiment, apart from **101** and **45** that remain significantly off. Figure 3C shows that the *full scaled* submission considerably improves free energy estimates for compound **101**, but also drastically decreases the accuracy of estimates for neutral compounds **102, 48, 95, 96**. It is not clear why predictions for **45** consistently perform so poorly.

Some highlights for the binding free energy estimations of set2 are shown in figure 4. The stage 1 *full* submission is depicted in figure 4A. The poor estimation of the relative binding free energy of neutral compounds (**38, 41, 73, 75**) with respect to other compounds in the dataset that carry one negative charge is very apparent. Figure 4B, shows the *full* submission at stage 2 of the competition. Estimates for the charged compounds improve, however the binding free energies of the neutral compounds are still poorly estimated. This is more apparent by inspection of figure 4C, which shows the *same charge* submission where only charged compounds were considered. This is the best performing protocol overall in terms of correlation coefficient of $R = 0.54 \pm 0.03$ and $MUE = 1.67 \pm 0.08$ kcal$\cdot$ mol^{-1}.

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

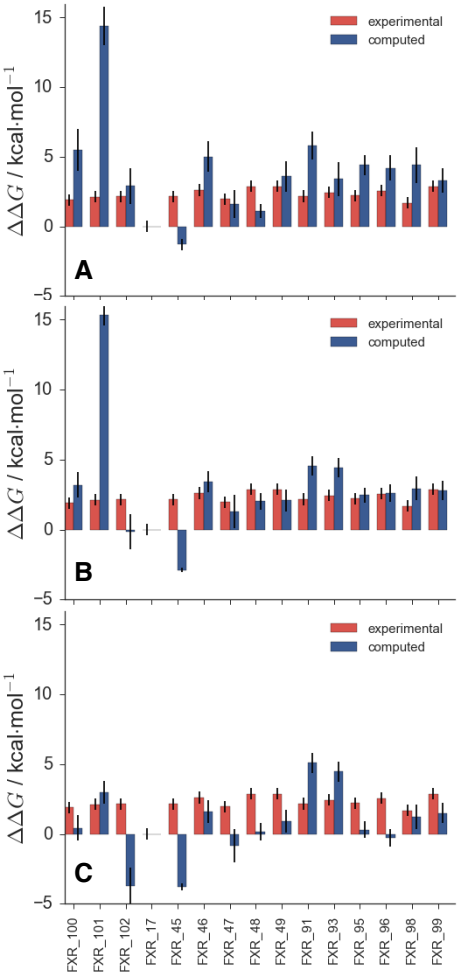

Fig. 3 A: Stage1 *full* submission for set1(ID a3c8k) showing clear overestimation of the relative binding free energy of charged compound **101**. B: Stage 2 *full* submission (ID 07tpe). C: Stage 2 *full scaled* submission (ID inspj).

3.3 Comparison to other submissions

The organisers released data for all binding free energy prediction submissions shortly after the end of stage2, and a summary of the correlation coefficients can be seen in figure 5. Figure 5A and B are stage 1 submissions for set1 and set2 respectively. Results for stage 2 set1 and set2 are shown in figure 5 C and D. The authors submissions are shown in green and can also be identified by their submission ID listed in table 2 and table 3. It should be noted that the shown

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

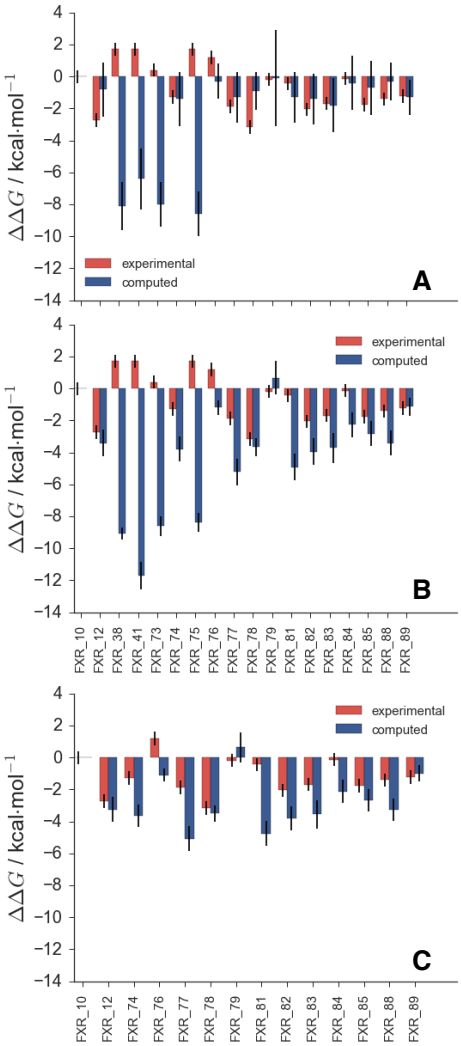

Fig. 4 A: Stage 1 *full* submission for set2 compounds (ID qvnq5). B: Stage 2 *full* submission for set2 compounds (ID qt771). C: Stage 2 *same charge* submission (ID 0jz8u).

correlation coefficients are slightly different to the ones reported in the tables 2 and 3. This is down to the use of different error analysis methods between the authors and the organisers, as discussed in the methods section 2.2. What becomes apparent however, from figure 5 is that for set1 both at stage 1 and stage 2 there is no protocol that obviously outperforms another protocol and no statistically significant ranking is possible. For set2 and particularly stage 2 there are four protocols that perform better than the rest, which are a mix of alchemical free

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

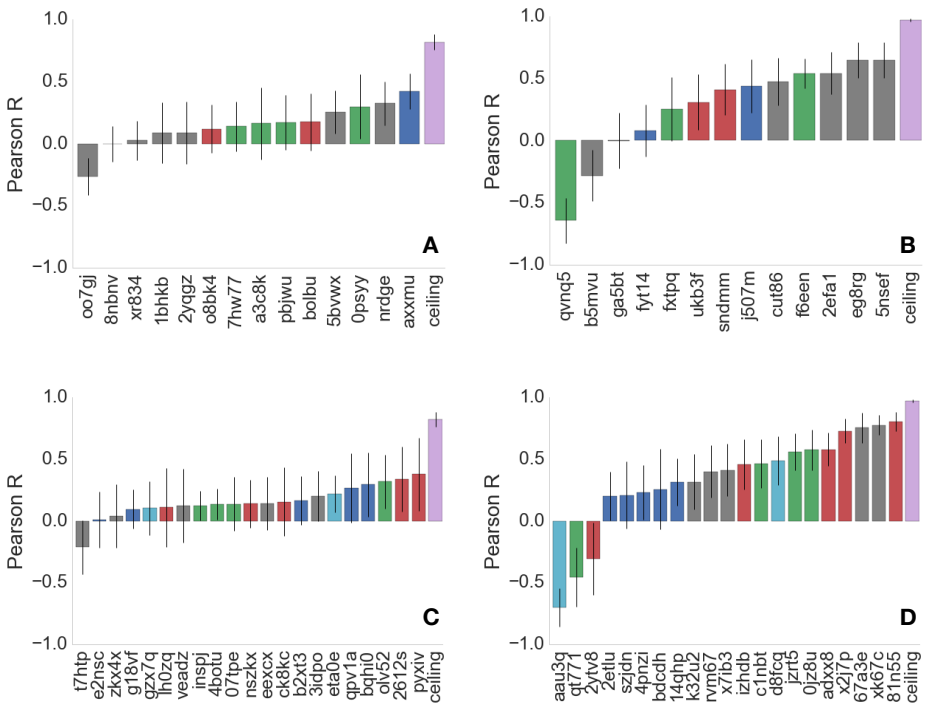

Fig. 5 Summary of all submitted protocols. A: Stage 1 for set1. B: Stage 1 for set2. C: Stage 2 for set1. D: Stage 2 for set2. Green colours denote the authors submissions and protocol IDs can be identified in table 2 and table 3. Red colours denote other alchemical methods, blue colours denote MMPBSA based methods, the light blue colour denote quantum mechanical based methods and grey denotes any other methods. The ceiling entry is discussed in the text and shown in purple. All method descriptions are made available by the organisers and can be found on the www.drugdesigndata.org website.

energy and other protocols. Submission **81n55** is an alchemical method based on FEP+, submission **xk67c** uses a non-equilibrium pulling approach using Gromacs as the simulation framework, submission **67a3e** uses the software MIX to perform energy minimisation of protein-ligand complexes, and submission **x2j7p** also used FEP+.

Figure 5 suggests, that there is no clear trend indicating a method that consistently outperforms others. Furthermore, the overall correlation between predicted free energies and experimental values is poor and unreliable for set1.

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

4 Conclusions

The blinded predictions on FXR ligands highlighted difficulties in reliably estimating relative binding energies between compounds that differ in their net-charge with the current workflow. This was anticipated in light of past experience and until methodological advances lift this limitation, alternative ad-hoc protocols may prove more reliable. For instance, other groups submitted in this competition alchemical binding energy predictions for ligands modelled as protonated acids in order to maintain the same net-charge across the full D3R datasets. While this is an unlikely chemical state for the unbound or bound ligands given the assay conditions, this setup did lead to superior predictions for the full D3R datasets.

The relatively reasonable correlations obtained retrospectively on LitSet1 and LitSet2 ligand series were encouraging, but the high mean-unsigned error on the LitSet1 dataset prompted the development of an approximate charge scaling protocol to account for potential neglect of polarisation effects. This had no beneficial effect on the accuracy of the blinded predictions and this protocol is not recommended for further use. Overall this indicates difficulties in reliably anticipating the robustness and transferability of the protocol across different ligand series, let alone different binding sites. In spite of the difficulties encountered it is useful to note that expert opinion based on analysis of literature SARs proved no more predictive on set1 and worse on set2 (excluding charged compounds). While this observation lacks statistical relevance – presumably there would also be variability in different expert opinions – it does highlight the difficulty of the problem. By contrast, expert opinion often fares well for poses prediction when compared against automated software workflows [27,54]. It was also encouraging that the correlation for set2 same-charge subset increased once experimental data about the binding mode of a representative set2 ligand could be taken into account. By contrast no significant variation was observed for set1 upon repeating the calculations, presumably because the binding modes had been well predicted at stage 1 of the competition.

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

The D3R Grand Challenge 2016 free energy datasets were markedly larger than those used in the 2015 competition. This enabled a more reliable comparison of the performance of different methodologies. Nevertheless, is apparent that both set1 and set2 are still too small to reliably rank a large number of submissions made by different groups. A general trend for alchemical free energy protocols can be observed, establishing them typically in the top 33% of submission, in particular in set2, for both correlation coefficient and root mean square error (RMSE), as shown in the SI. It is noteworthy that the features of the distribution of experimental binding energies for set1 (shorter span and uneven density) contribute to making predictions intrinsically more difficult than for set2. In addition, the precision of the experimental data was not determined. Assuming a ca. 0.4 kcal·mol^{-1} uncertainty in experimental measurements [55] together with bootstrapping suggests ceiling values for R of ca. 0.82±0.06 and 0.97±0.01 for set1 and set2 respectively. Thus the best performing methods are far from achieving high accuracy R values on set1, but show respectable correlation on set2. While it may be difficult to source significantly larger datasets amenable to alchemical free energy calculations, it may be useful to assess their intrinsic difficulty for the design of future competitions.

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Acknowledgements J.M. is supported by a Royal Society University Research Fellowship. The research leading to these results has received funding from the European Research Council under the European Unions Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement No. 336289. J. J-J is supported by a Marie Sklodowska-Curie fellowship grant agreement No 655677.

References

1. Yuqing Deng and Benot Roux. *J. Chem. Theo. Comput.*, 2(5):1255–1273, 2006.
2. Chia-En Chang and Michael K. Gilson. *J. Am. Chem. Soc.*, 126(40):13156–13164, 2004.
3. Julien Michel. *Phys. Chem. Chem. Phys.*, 16:4465–4477, 2014.
4. John D Chodera, David L Mobley, Michael R Shirts, Richard W Dixon, Kim Branson, and Vijay S Pande. *Curr. Opin. Struct. Biol.*, 21(2):150 – 160, 2011.
5. Wei Jiang and Benot Roux. *J. Chem. Theo. Comput.*, 6(9):2559–2565, 2010.
6. Julien Michel, Nicolas Foloppe, and Jonathan W. Essex. *Mol. Inform.*, 29(8-9):570–578, 2010.
7. Julien Michel and Jonathan W. Essex. *J. Comput.-Aided Mol.Des.*, 24(8):639–658, 2010.
8. William L. Jorgensen and C. Ravimohan. *J. Chem. Phys.*, 83(6):3050–3054, 1985.
9. Bhalachandra L. Tembre and J.Andrew Mc Cammon. *Computers & Chemistry*, 8(4):281 – 283, 1984.
10. Lingle Wang, Yujie Wu, Yuqing Deng, Byungchan Kim, Levi Pierce, Goran Krilov, Dmitry Lupyan, Shaughnessy Robinson, Markus K. Dahlgren, Jeremy Greenwood, Donna L. Romero, Craig Masse, Jennifer L. Knight, Thomas Steinbrecher, Thijs Beuming, Wolfgang Damm, Ed Harder, Woody Sherman, Mark Brewer, Ron Wester, Mark Murcko, Leah Frye, Ramy Farid, Teng Lin, David L. Mobley, William L. Jorgensen, Bruce J. Berne, Richard A. Friesner, and Robert Abel. *J. Am. Chem. Soc.*, 137(7):2695–2703, 2015.
11. Gaetano Calabr, Christopher J. Woods, Francis Powlesland, Antonia S. J. S. Mey, Adrian J. Mulholland, and Julien Michel. *J. Phys. Chem. B*, 120(24):5340–5350, 2016.
12. Nathan M. Lim, Lingle Wang, Robert Abel, and David L. Mobley. *Journal of Chemical Theory and Computation*, 12(9):4620–4631, 2016.
13. Julien Michel, Marcel L. Verdonk, and Jonathan W. Essex. *Journal of Chemical Theory and Computation*, 3(5):1645–1655, 2007.
14. Julien Michel and Jonathan W. Essex. *Journal of Medicinal Chemistry*, 51(21):6654–6664, 2008.
15. Yuan Hu, Brad Sherborne, Tai-Sung Lee, David A. Case, Darrin M. York, and Zhuyan Guo. *Journal of Computer-Aided Molecular Design*, 30(7):533–539, 2016.

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

16. Stefania Evoli, David L. Mobley, Rita Guzzi, and Bruno Rizzuti. *Phys. Chem. Chem. Phys.*, 18:32358–32368, 2016.

17. James Luccarelli, Julien Michel, Julian Tirado-Rives, and William L. Jorgensen. *Journal of Chemical Theory and Computation*, 6(12):3850–3856, 2010.

18. Julien Michel, Julian Tirado-Rives, and William L. Jorgensen. *The Journal of Physical Chemistry B*, 113(40):13337–13346, 2009.

19. Julien Michel, Julian Tirado-Rives, and William L. Jorgensen. *Journal of the American Chemical Society*, 131(42):15403–15411, 2009.

20. Sushil K. Mishra, Gaetano Calabr, Hannes H. Loeffler, Julien Michel, and Jaroslav Koa. *J. Chem. Theo. Comput.*, 11(7):3333–3345, 2015.

21. Julien Michel, Marcel L. Verdonk, and Jonathan W. Essex. *Journal of Medicinal Chemistry*, 49(25):7427–7439, 2006.

22. Bradley Sherborne, Veerabahu Shanmugasundaram, Alan C. Cheng, Clara D. Christ, Renee L. DesJarlais, Jose S. Duca, Richard A. Lewis, Deborah A. Loughney, Eric S. Manas, Georgia B. McGaughey, Catherine E. Peishoff, and Herman van Vlijmen. *Journal of Computer-Aided Molecular Design*, 30(12):1139–1141, 2016.

23. Andriy Kryshtafovych, Krzysztof Fidelis, and John Moult. *Proteins: Structure, Function, and Bioinformatics*, 82:164–174, 2014.

24. Shoshana J Wodak and Ral Mndez. *Current Opinion in Structural Biology*, 14(2):242 – 249, 2004.

25. Jian Yin, Niel M. Henriksen, David R. Slochower, Michael R. Shirts, Michael W. Chiu, David L. Mobley, and Michael K. Gilson. *Journal of Computer-Aided Molecular Design*, 31(1):1–19, 2017.

26. Richard D. Smith, James B. Dunbar, Peter Man-Un Ung, Emilio X. Esposito, Chao-Yie Yang, Shaomeng Wang, and Heather A. Carlson. *Journal of Chemical Information and Modeling*, 51(9):2115–2131, 2011.

27. Antonia S.J.S. Mey, Jordi Jurez-Jimnez, Alexis Hennessy, and Julien Michel. *Bioorg. Med. Chem.*, 24(20):4890 – 4899, 2016.

28. Stefano Bosisio, Antonia S. J. S. Mey, and Julien Michel. *Journal of Computer-Aided Molecular Design*, 30(11):1101–1114, 2016.

29. Stefano Bosisio, Antonia S. J. S. Mey, and Julien Michel. *Journal of Computer-Aided Molecular Design*, 31(1):61–70, 2017.

30. Symon Gathiaka, Shuai Liu, Michael Chiu, Huanwang Yang, Jeanne A. Stuckey, You Na Kang, Jim Delproposto, Ginger Kubish, James B. Dunbar, Heather A. Carlson, Stephen K. Burley, W. Patrick Walters, Rommie E. Amaro, Victoria A. Feher, and Michael K. Gilson. *J. Comput.-Aided Mol. Des.*, 30(9):651–668, 2016.

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

31. Michael R. Shirts, Christoph Klein, Jason M. Swails, Jian Yin, Michael K. Gilson, David L. Mobley, David A. Case, and Ellen D. Zhong. *Journal of Computer-Aided Molecular Design*, 31(1):147–161, 2017.

32. V.; Gilson M.K.; Burley S.K. Drug Design Data Resource: Amaro, R.; Feher. An open resource to advance computer-aided drug design, (n.d.).

33. James S. Nichols, Derek J. Parks, Thomas G. Consler, and Steven G. Blanchard. *Anal. Biochem.*, 257(2):112 – 119, 1998.

34. Song Feng, Minmin Yang, Zhenshan Zhang, Zhanguo Wang, Di Hong, Hans Richter, Gregory Martin Benson, Konrad Bleicher, Uwe Grether, Rainer E. Martin, Jean-Marc Plancher, Bernd Kuhn, Markus Georg Rudolph, and Li Chen. *Bioorg. Med. Chem.*, 19(9):2595 – 2598, 2009.

35. Hans G.F. Richter, Gregory M. Benson, Denise Blum, Evelyne Chaput, Song Feng, Christophe Gardes, Uwe Grether, Peter Hartman, Bernd Kuhn, Rainer E. Martin, Jean-Marc Plancher, Markus G. Rudolph, Franz Schuler, Sven Taylor, and Konrad H. Bleicher. *Bioorg. Med. Chem. Lett.*, 21(1):191 – 194, 2011.

36. Hans G.F. Richter, G.M. Benson, K.H. Bleicher, D. Blum, E. Chaput, N. Clemann, S. Feng, C. Gardes, U. Grether, P. Hartman, B. Kuhn, R.E. Martin, J.-M. Plancher, M.G. Rudolph, F. Schuler, and S. Taylor. *Bioorg. Med. Chem. Lett.*, 21(4):1134 – 1140, 2011.

37. Schrödinger release 2015-2: Maestro, version 10.2, schrdinger, llc, new york, ny, 2015.

38. Sergio Ruiz-Carmona, Daniel Alvarez-Garcia, Nicolas Foloppe, A. Beatriz Garmendia-Doval, Szilveszter Juhos, Peter Schmidtke, Xavier Barril, Roderick E. Hubbard, and S. David Morley. *PLOS Computational Biology*, 10(4):1–7, 04 2014.

39. Hannes H. Loeffler, Julien Michel, and Christopher Woods. *J Chem. Inf. Comput. Sci.*, 55(12):2485–2490, 2015.

40. D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Goetz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M. J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P. A. Kollman, 2014.

41. William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. *The Journal of Chemical Physics*, 79(2):926–935, 1983.

42. Eyal Neria, Stefan Fischer, and Martin Karplus. *The Journal of Chemical Physics*, 105(5):1902–1921, 1996.

43. C. Woods, A. S. J. S. Mey, G. Calabro, S. Bosisio, and J. Michel. Sire molecular simulations framework, (http://siremol.org). (accessed May 31, 2016).

bioRxiv preprint doi: https://doi.org/10.1101/150474; this version posted June 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

44. Peter Eastman, Mark S. Friedrichs, John D. Chodera, Randall J. Radmer, Christopher M. Bruns, Joy P. Ku, Kyle A. Beauchamp, Thomas J. Lane, Lee-Ping Wang, Diwakar Shukla, Tony Tye, Mike Houston, Timo Stich, Christoph Klein, Michael R. Shirts, and Vijay S. Pande. *J. Chem. Theory Comput.*, 9(1):461–469, 2013.
45. David L. Mobley and Pavel V. Klimovich. *J. Chem. Phys.*, 137(23):230901, 2012.
46. Jacob G. Zeevaart, Ligong Wang, Vinay V. Thakur, Cheryl S. Leung, Julian Tirado-Rives, Christopher M. Bailey, Robert A. Domaoal, Karen S. Anderson, and William L. Jorgensen. *J. Am. Chem. Soc.*, 130(29):9492–9499, 2008.
47. Thomas Steinbrecher, Andrea Hrenn, Korinna L. Dormann, Irmgard Merfort, and Andreas Labahn. *Bioorg. Med. Chem.*, 16(5):2385 – 2390, 2008.
48. Jiyao Wang, Yuqing Deng, and Benot Roux. *Biophys. J.*, 91(8):2798 – 2814, 2006.
49. A. S. J. S. Mey, J. Juarez-Jimenez, and J. Michel. Supplementary information.
50. A. S. J. S. Mey and J. Michel. Github page for simulation data, 2017. (accessed May 31, 2016).
51. Michael R. Shirts and John D. Chodera. *J Chem. Phys.*, 129(12), 2008.
52. Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference (SciPy2008)*, pages 11–15, Pasadena, CA USA, August 2008.
53. Frank R. Beierlein, Julien Michel, and Jonathan W. Essex. *The Journal of Physical Chemistry B*, 115(17):4911–4926, 2011.
54. Charles H. Reynolds Kenneth M. Merz Jr., Dagmar Ringe, editor. *Drug Design Structure- and Ligand-Based Approach.* Cambridge University Press, 2010. Chapter 7.
55. Scott P. Brown, Steven W. Muchmore, and Philip J. Hajduk. *Drug Discovery Today*, 14(7):420 – 427, 2009.