

Absolute Free Energy of Binding Calculations for Macrophage Migration Inhibitory Factor in Complex with a Drug-like Inhibitor

Yue Qian, Israel Cabeza de Vaca, Jonah Z. Vilseck[‡], Daniel J. Cole[‡], Julian Tirado-Rives and William L. Jorgensen*

Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107, United States

ABSTRACT: Calculation of the absolute free energy of binding (ΔG_{bind}) for a complex in solution is challenging owing to the need for adequate configurational sampling and an accurate energetic description, typically with a force field (FF). In this study, Monte Carlo (MC) simulations with improved side-chain and backbone sampling are used to assess ΔG_{bind} for the complex of a drug-like inhibitor (MIF180) with the protein macrophage migration inhibitory factor (MIF) using free energy perturbation (FEP) calculations. For comparison, molecular dynamics (MD) simulations were employed as an alternative sampling method for the same system. With the OPLS-AA/M FF and CM5 atomic charges for the inhibitor, the ΔG_{bind} results from the MC/FEP and MD/FEP simulations, -8.80 ± 0.74 and -8.46 ± 0.85 kcal/mol, agree well with each other and with the experimental value of -8.98 ± 0.28 kcal/mol. The convergence of the results and analysis of the trajectories indicate that sufficient sampling was achieved for both approaches. Repeating the MD/FEP calculations using current versions of the CHARMM and AMBER FFs led to a 6-kcal/mol range of computed ΔG_{bind} . These results show that calculation of accurate ΔG_{bind} for large ligands is both feasible and numerically equivalent, within error limits, using either methodology.

[‡] Current address: Indiana University, School of Medicine, Indianapolis Indiana 46202-5126

[‡] Current address: School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne NE1 7RU, UK.

This is the author's manuscript of the article published in final edited form as:

Qian, Y., Cabeza de Vaca, I., Vilseck, J. Z., Cole, D. J., Tirado-Rives, J., & Jorgensen, W. L. (2019). Absolute Free Energy of Binding Calculations for Macrophage Migration Inhibitory Factor in Complex with a Drug-like Inhibitor. *The Journal of Physical Chemistry B*. <https://doi.org/10.1021/acs.jpcc.9b07588>

INTRODUCTION

Alchemical binding free energy calculations have been rapidly developing and are now being widely applied in structure-based drug design (SBDD).¹⁻⁶ Different statistical mechanics approaches have been explored to try to achieve accurate binding affinity predictions.^{1,7} Perturbative free energy methods such as thermodynamic integration (TI), free-energy perturbation (FEP) and Bennett's acceptance ratio (BAR) are based on the assumption that the configurational space of two different states is similar enough to obtain valid evaluations of the difference in free energies. To ensure this condition, the stratification technique splits the transformation path into a number of intermediate steps or "λ-windows" that yield adequate overlap of the configurational spaces. Relative binding free energy ($\Delta\Delta G_{\text{bind}}$) calculations, where the initial and final molecules are very similar, have been dominant in structure-based drug design (SBDD) studies.^{5,6} In contrast, absolute binding free energy calculations decouple energetically the ligand entirely from its environment, either the surrounding solvent molecules or a protein binding site.^{8,9} As the removal of the entire ligand molecule is performed, such calculations are computationally demanding and potentially sensitive to sampling and numerous setup issues for the protein. On the other hand, they do address the fundamental thermodynamic gauge of molecular recognition and the results can be directly compared to experimental binding data, after corrections for standard states are introduced.¹⁰⁻¹² The calculations when performed in a prospective manner provide a rigorous test of current methodologies and force fields.¹³ However, such calculations are still far from routine and, as considered here, further examination

1
2
3
4 of methodological issues and the impact of alternative force fields is needed.
5

6
7 Much work in the area has been done with molecular dynamics (MD) methods using
8
9 software packages such as GROMACS,¹⁴ AMBER,¹⁵ NAMD,¹⁶ CHARMM,¹⁷ and OpenMM.¹⁸
10
11 Much less work has used Monte Carlo statistical mechanics (MC), though it can be very efficient
12
13 compared to MD for liquid simulations.¹⁹ Unlike MD, where a new configuration is generated by
14
15 integrating equations of motion for all atoms, MC explores the configurational space by localized
16
17 random moves of solvent and solute molecules.¹⁹ It also permits enhanced sampling of
18
19 conformational changes and of local regions of interest, *e. g.*, near the protein binding pocket.
20
21 Moreover, NVT and NPT ensembles are readily implemented through the Metropolis sampling
22
23 without the need to apply thermostats and barostats. Recent improvements in the MC-based
24
25 software package MCPRO have resulted in enhanced sampling of protein side chains and
26
27 backbone atoms.²⁰ Similar sampling and absolute free energies of binding were obtained for
28
29 complexes of benzene and analogs with T4 lysozyme L99A using MD or MC.²⁰ It is of interest,
30
31 then, to extend this study to a more drug-relevant biomolecular system and further assess the
32
33 sampling performance of MC and MD. For a common force field, MC and MD are expected to
34
35 converge to the same ΔG_{bind} results, once sufficient configurational sampling is achieved.
36
37
38
39
40
41
42
43
44
45
46

47 Macrophage migration inhibitory factor (MIF) is both a keto-enol tautomerase and a
48
49 cytokine associated with inflammatory diseases and cancer.^{21,22} It was selected as the subject of
50
51 this study since it has many characteristics that make it a suitable benchmark system: (1) the
52
53 trimeric protein has moderate size with 342 residues; (2) multiple high-resolution crystal
54
55
56
57

1
2
3
4 structures of complexes of MIF with tautomerase inhibitors are available; (3) the crystal
5
6 structures for many inhibitors show modest conformational changes for binding-site residues;
7
8 and, (4) experimental binding data, K_i and K_d values, are available from inhibition and
9
10 fluorescence polarization assays. In particular, for this work, we have chosen to study the
11
12 complex of the inhibitor MIF180 with human MIF, as illustrated in Figure 1 from the crystal
13
14 structure obtained in our laboratory.²² As indicated, the complex features a combination of
15
16 hydrogen bonding, van der Waals, and aryl-aryl interactions, which is typical for protein-drug
17
18 complexes. In contrast, the widely used L99A T4-lysozyme system binds benzene analogs
19
20 primarily through the hydrophobic effect.²⁰

21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Once adequate sampling is achieved, the effects of the accuracy of the force field and
other methodological factors can be evaluated for the benchmark system *via* ΔG_{bind} calculations.

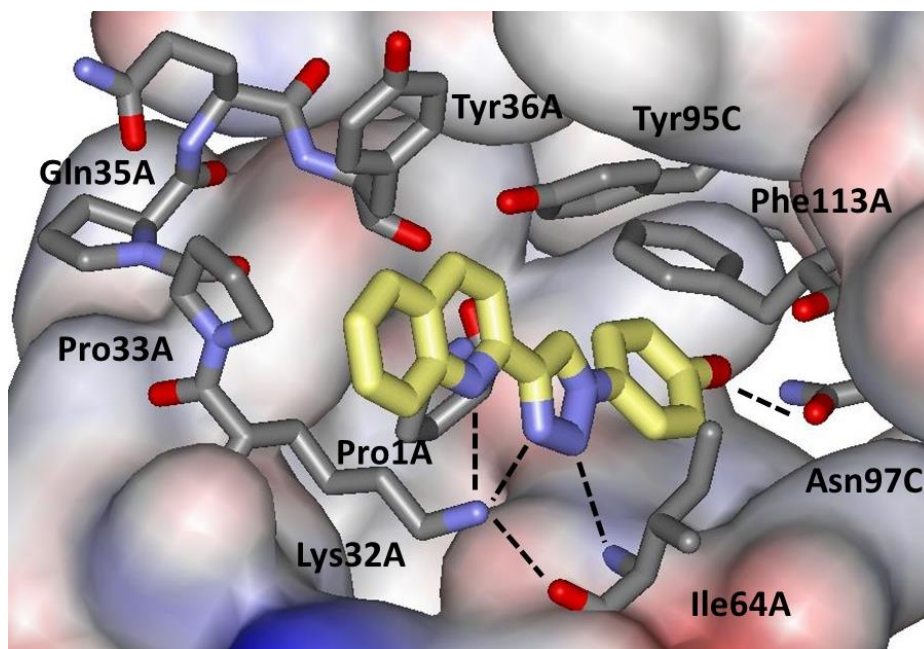


Figure 1. Rendering from the crystal structure of MIF180 bound to human MIF (PDB ID: 4WR8).²² Hydrogen bonds are indicated with dashed lines.

1
2
3
4 In our laboratory, much effort has been devoted to steady improvements of the OPLS force
5
6 fields. Recently, the OPLS-AA force field for proteins and nucleic acids has been improved
7
8 through extensive reoptimization of the torsional parameters using high-level quantum
9
10 mechanical calculations and MC and MD simulations of series of peptides, proteins, nucleotides
11
12 and polynucleotides to yield OPLS-AA/M.^{23,24} In addition, OPLS parameters for general small-
13
14 molecule ligands are now available with atomic charges from QM calculations, after
15
16 optimization through studies of properties of pure liquids and free energies of hydration.²⁵
17
18 CHARMM²⁶ and AMBER²⁷ are two other popular force fields initially parameterized for
19
20 proteins and later extended to nucleic acids, lipids, and small molecules (CGenFF²⁸ and
21
22 GAFF²⁹). In the present work, four combinations of protein-ligand force fields are utilized,
23
24 namely OPLS-AA/M with OPLS-AA/CM5, OPLS-AA/M with OPLS-AA/CM1A, CHARMM
25
26 36 with CGenFF, and AMBER ff14sb with GAFF. These will be referred to as OPLS/CM5,
27
28 OPLS/CM1A, CHARMM/CGenFF, and AMBER/GAFF.
29
30
31
32
33
34
35
36
37
38

39 In this work, ΔG_{bind} results for the MIF180/MIF complex have been obtained from Monte
40
41 Carlo free energy perturbation (MC/FEP) and MD/FEP calculations using the OPLS/CM5 force
42
43 field for comparison with each other and with the K_d measurement from a fluorescence
44
45 polarization assay ($\Delta G_{\text{bind}} = RT \ln K_d$).³⁰ Six-degree-of-freedom (6DoF) restraints were adopted
46
47 for all simulations of the complex.^{12,13,31} In addition, the remaining force field combinations,
48
49 OPLS/CM1A, CHARMM/CGenFF and AMBER/GAFF, have been applied using the same
50
51 MD/FEP protocol to evaluate the sensitivity of the ΔG_{bind} results to these alternative choices.
52
53
54
55
56
57
58
59
60

METHODS

Absolute Binding Free Energy Calculations. Absolute binding free energy calculations were conducted via the double decoupling method (DDM) following the thermodynamic cycle depicted in Figure 2 and using eqs 1 and 2.⁸⁻¹² The difference between the two sides of the cycle, which effectively transfers the ligand from aqueous solution to the binding site, represents the binding affinity of the protein-ligand complex. The ligand intermolecular interactions are turned off (decoupled) from the water solvent in the unbound simulation to yield $\Delta G_{\text{unbound}}$. All calculations of ΔG_{bound} and $\Delta G_{\text{unbound}}$ were done in two stages with scaling of the atomic charges

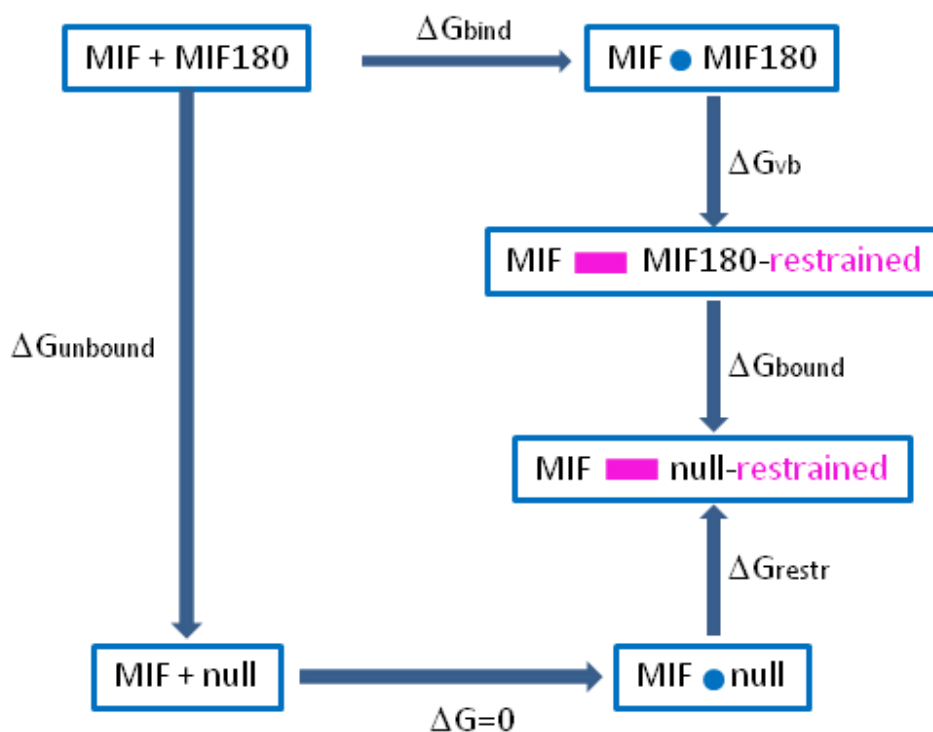


Figure 2. Thermodynamic cycle for computing ΔG_{bind} .

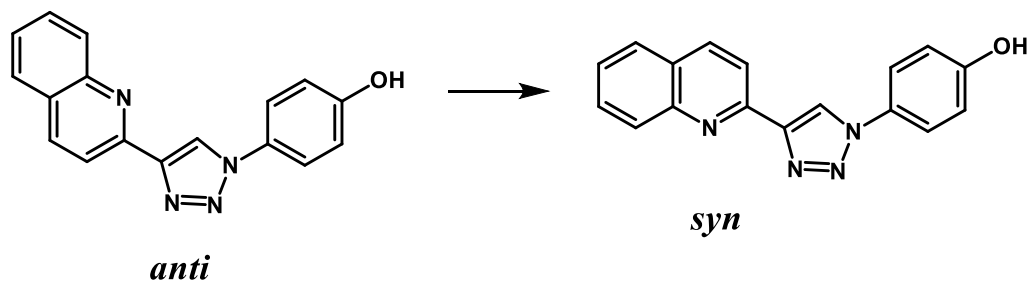
$$\Delta G_{bind} = \Delta G_{unbound} - \Delta G_{bound} + \Delta G_{restr} - \Delta G_{vb} \quad (1)$$

$$\Delta G_{restr} = -kT \ln \left[\frac{8\pi^2 V (K_r K_{\theta A} K_{\theta B} K_{\phi A} K_{\phi B} K_{\phi C})^{1/2}}{r_{a,A,0}^2 \sin \theta_{A,0} \sin \theta_{B,0} (2\pi kT)^3} \right] \quad (2)$$

to zero followed by removal of the intermolecular Lennard-Jones interactions. In the bound simulations, the decoupling from the solvent and the protein is done with the use of geometric restraints. They are introduced to keep the disappearing ligand in the observed position and orientation in the binding site (Figure 1) yielding ΔG_{vb} , and their effect on the free energy is corrected analytically via ΔG_{restr} . The latter term is calculated using eq 14 from the paper by Boresh *et al.*¹³ The equation is reproduced here as eq 2 and reflects imposition of restraints for six degrees of freedom (6DoF) that keep the ligand translationally and rotationally stable in the binding site.³²⁻³⁴ The specific algorithm from Wang *et al.*³⁵ is implemented in the colvars module of the utilized MD program NAMD³⁶ to control the six variables.^{31,37} The same algorithm is also used in our MC program, MCPRO. In the MC implementation, the restraints are turned on simultaneously with the removal of the atomic charges so ΔG_{vb} is included in the electrostatic portion of ΔG_{bound} , but it requires a separate simulation in the MD calculations using NAMD.

The sum of terms in eq 1 is sufficient to calculate ΔG_{bind} of most ligands that are either conformationally rigid or freely interconverting. The ligand in this study, MIF180, occurs in two different, non-interconverting conformations during the simulations bound in the protein and in the unbound aqueous phase. Therefore a correction term, ΔG_{conf} , needs to be added to the result

1
2
3
4 of eq 1 as a penalty for conversion of the ligand from the most stable conformation in aqueous
5
6
7 solution or in the gas phase with the triazole and quinoline nitrogen atoms *anti* to the *syn*
8
9 conformer observed in the complex (Scheme 1). This term was estimated *via* potential of mean



30
31
32
33
34
35
36
37

Scheme 1. Two different conformations of MIF180.

38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

force (PMF) calculations for rotation about the bond connecting the rings. The decoupling was only for interactions with the environment; intramolecular energy terms did not contribute to the FEP results, so the conformational change required separate assessment.

System Preparation. All structures were initially built starting from crystal structures of the complexes for MIF180 and a close analog (PDB IDs: 4WR8 and 4WRB)²² using the MCPRO *clu* utility. The full structure with 342 residues was retained and relaxed via a conjugate-gradient optimization using MCPRO with a dielectric constant of 2.0. For the ligand, two different OPLS-AA charge models, OPLS-AA/CM1A and OPLS-AA/CM5 were considered with the usual scaling factors for neutral-molecule partial charges of 1.14 for CM1A and 1.20 for CM5.³⁸ In addition, CHARMM general force field (CGenFF) parameters were obtained from its

1
2
3
4 webserver.²⁸ It should be noted that the output included warnings about low quality for torsional
5
6
7 parameters for several dihedral angles. Similarly, the ligand parameters were assigned for the
8
9
10 general AMBER force field (GAFF)²⁹ with the Antechamber package including AM1-BCC
11
12 atomic charges.³⁹ The protonation states for the protein residues were determined by H⁺⁺^{40,41}
13
14 and PropKa3.^{42,43} In both cases, the N-terminal proline was predicted to be neutral. Pro1 is the
15
16
17 putative catalytic base for the tautomerase reaction.²¹ However, no neutral N-terminal proline
18
19
20 parameters were available for the CHARMM36²⁶ and the AMBER ff14SB²⁷ force fields. In order
21
22
23 to obtain the necessary parameters, the C-terminal-capped PRO-MET dipeptide was processed
24
25
26 using CGenFF and GAFF. The parameters thus obtained were then mapped into the CHARMM
27
28
29 and AMBER protein force fields for the neutral terminal proline. Complexes for MIF with both
30
31
32 the neutral and protonated Pro1 were prepared for all four force field combinations.

33
34 **Monte Carlo Simulations.** The OPLS-AA/M force field²³ was used for the MIF protein,
35
36
37 while OPLS-AA with the 1.20*CM5 charge model³⁸ was used to represent the inhibitor, MIF180.
38
39
40 The calculations of absolute free energies were carried out following the double-decoupling
41
42
43 scheme (Figure 2). The ligand electrostatic and Lennard-Jones (LJ) interactions were decoupled
44
45
46 consecutively with simple overlap sampling (SOS).⁴⁴ The charges were first scaled to zero
47
48
49 linearly with the λ parameter, then the intermolecular LJ interactions were turned off using 1-1-6
50
51
52 softcore potentials.⁴⁵ The charge and LJ removals were split into 15 and 18 windows for the
53
54
55 unbound state and 15 and 41 windows in for the bound state respectively. Each window
56
57
58 comprised 80 million (80M) configurations of equilibration and 180M configurations of

1
2
3
4 averaging (80M/180M) for the unbound state; the bound-state calculations used 80M/240M
5
6 configurations for the electrostatic and 320M/240M for the LJ calculations. The PMF
7
8 calculations for the rotation around the bond connecting the triazole and quinoline rings were
9
10 conducted using the same system as prepared for the unbound simulations using 1M
11
12 configurations of equilibration and 5M averaging at 10° intervals.
13
14
15
16

17
18 During the bound state annihilation, the ligand was restrained to its initial position relative to
19
20 the protein using the six degree-of-freedom (6DoF) restraints.^{13,33} The required coordinate
21
22 system was constructed by choosing three sites from the ligand and three from nearby protein
23
24 residues. Specifically, the hydroxyl oxygen atom, the midpoint of N2 and N3 in the triazole ring,
25
26 and the quinoline N atom were selected as the three groups on the ligand; and the geometric
27
28 center of the heavy atoms for the Tyr36C side-chain, the Ile64A backbone and the Lys32A side-
29
30 chain were used for the protein. Analytical corrections for the restraints to the fully interacting
31
32 ligand and standard state were included to obtain the absolute binding free energy.^{13,35,46,47} The
33
34 force constants for the six restraints were 10 kcal/mol-Å² and 0.1 kcal/mol-deg² for distances and
35
36 angles, respectively. All restraint terms were gradually increased starting from zero during the
37
38 electrostatic decoupling and then kept constant.
39
40
41
42
43
44
45
46

47
48 For MC/FEP simulations, the unbound ligand was solvated in a 40-Å periodic cube containing
49
50 2100 TIP4P water molecules. For the complexes, a 25-Å radius cap with ca. 2000 TIP4P water
51
52 molecules centered on the ligand was used. In all cases, a residue-based cutoff of 10 Å was
53
54 applied to maintain consistency with the MD simulations.
55
56
57

1
2
3
4 **Molecular Dynamics Simulations.** All molecular dynamics (MD) simulations were
5
6 performed using the NAMD program version 2.11.¹⁶ The ligand was solvated in a 40-Å periodic
7
8 cube of TIP3P water,⁴⁸ the default model for NAMD, and the complex was solvated in TIP3P
9
10 water with 12-Å padding in all dimensions. The systems with unprotonated Pro1 were
11
12 electrically neutral, and the ones with the three copies of the proline protonated were neutralized
13
14 by addition of three chloride ions.
15
16
17
18
19

20 Langevin dynamics⁴⁹ was applied to enforce a temperature of 300 K and a pressure of 1 atm.
21
22 The time step was set at 2 fs using the SHAKE algorithm⁵⁰ to constrain all bonds to hydrogen
23
24 atoms. Coulombic interactions were truncated at 10 Å, and the Particle-Mesh-Ewald (PME)
25
26 method was used to include long-range electrostatic interactions.^{51,52} The LJ interactions were
27
28 smoothly switched off between 8 and 10 Å. For all of the MD simulations, the equilibration
29
30 protocol consisted of 50,000 steps of conjugate-gradient minimization, followed by 2 ns of
31
32 isothermal-isobaric dynamics for equilibration. The free energy rotational profiles were
33
34 calculated from the unbound ligand setup using 1 ns equilibration and 5 ns averaging at 1°
35
36 intervals.
37
38
39
40
41
42
43

44 The production FEP calculations were performed bi-directionally using 43 lambda windows.
45
46 Each λ -window featured another 1 ns of equilibration and 5 ns of averaging for both the unbound
47
48 and bound state, for a total 215 ns in each direction. The 6DoF restraints were applied to the
49
50 ligand in binding site in the same manner as with MCPRO. The calculations of the absolute free
51
52 energy of binding with protonated Pro1 using the AMBER/GAFF force field required an
53
54
55
56
57

1
2
3
4 additional conformational restraint to maintain the bound *syn* conformation. All of the MD
5
6
7 simulations were run in triplicate with small initial changes to generate independent trajectories.
8
9
10

11 **RESULTS AND DISCUSSION**

12 **Absolute Free Energy of Binding with MC**

13
14
15
16
17 **Calculations for the Unbound Ligand; Comparison of Water Models.** Since the MC
18
19
20 simulations used TIP4P water, while the MD simulations used TIP3P, the effect of the choice
21
22
23 was first examined for the unbound state by evaluating free energies of hydration (ΔG_{hyd}) of the
24
25
26 ligand. For this purpose, full annihilations of the ligand in aqueous solution and in the gas phase
27
28
29 were carried out with sequential removal of the Coulombic and LJ interactions. The net MC
30
31
32 results for ΔG_{hyd} were similar, specifically, -14.92 ± 0.10 kcal/mol for TIP4P and -14.23 ± 0.05
33
34 kcal/mol for TIP3P water. The results for the aqueous FEP calculations were taken as $\Delta G_{\text{unbound}}$.
35

36
37 **Calculations for the Complex.** Bound-state MC/FEP calculations were carried out
38
39 following similar protocols as in the lysozyme study.²⁰ However, a few methodological
40
41
42 adjustments were necessary due to the larger size and asymmetry of the MIF180 ligand. First, the
43
44
45 hard wall (HW) restraint was replaced by the 6DoF restraints. While the HW restraint was
46
47
48 sufficient for benzene and analogs as ligands, it resulted in very slow convergence and large
49
50
51 numerical fluctuations in initial studies with MIF180. With the 6DoF restraints, the bound-state
52
53
54 results showed good numerical stability.

55
56 Next, 41 λ -windows were used to ensure sufficient configuration-space overlap for the FEP
57

calculations. Simple overlap sampling (SOS) was used, where the midpoint of the window λ_M is the end-point for each perturbation. The bound-state free energy changes for each window and the corresponding fluctuations are plotted in Figures 3 and 4 from the 80M/240M run lengths.

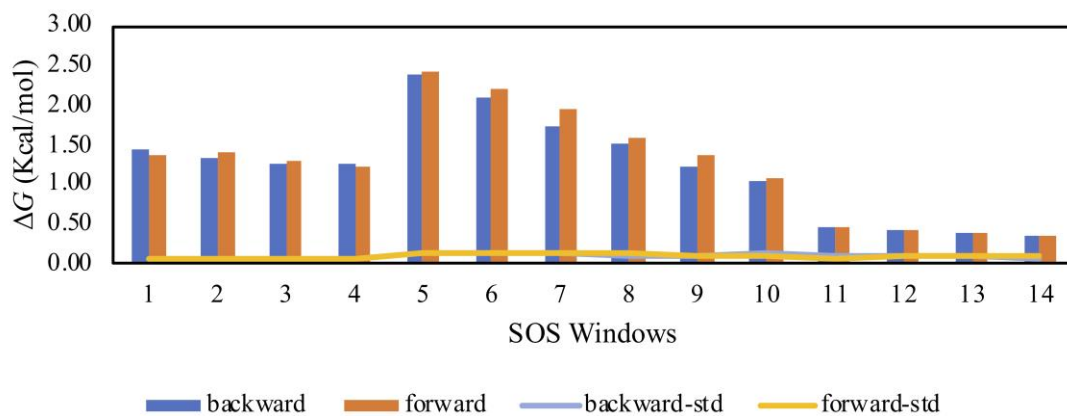


Figure 3: Bound-state free energy changes and fluctuations for each λ -window of the electrostatic decoupling. Bar plots: free energy change in each SOS window with forward referring to λ_1 to λ_M and backward λ_2 to λ_M , where λ_M refers to the intermediate point between each λ_1 and λ_2 pairs. Line plots: Standard deviation (1σ) with 80M/240M MC configurations.

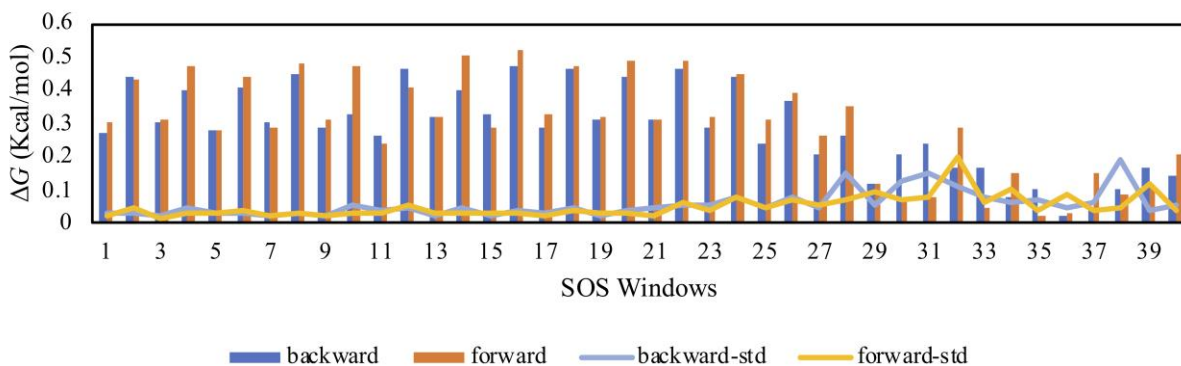


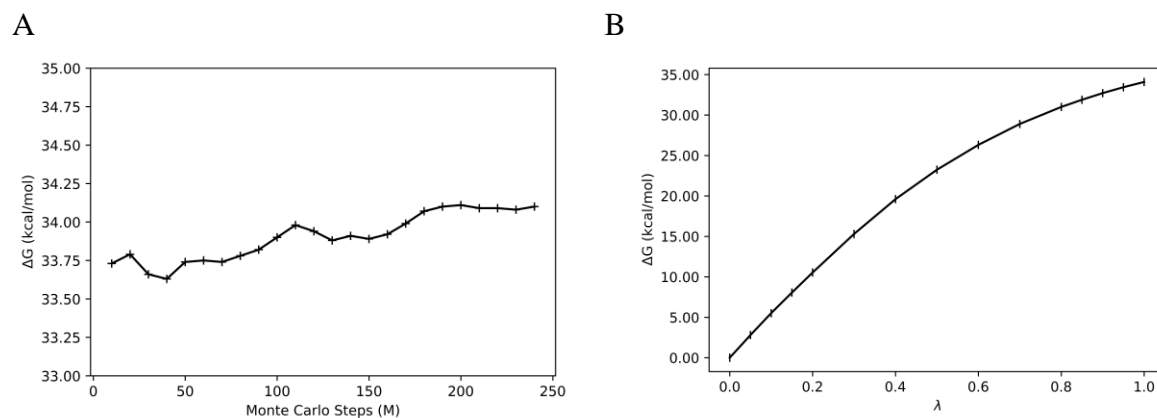
Figure 4: Bound-state free-energy changes and fluctuations for each λ -window of the LJ decoupling as in Figure 3.

While the electrostatic free-energy fluctuations were all below 0.13 kcal/mol, the LJ free-energy changes exhibit two behaviors. Low fluctuations are exhibited in the first two-thirds of the

1
2
3
4 calculations, up to window 26 ($0.00 \leq \lambda \leq 0.67$), and then higher fluctuations occurred for the
5
6 remaining windows ($0.67 \leq \lambda \leq 1.00$). In the latter region, all the electrostatic interactions and a
7
8 large portion of the Lennard-Jones ones have been eliminated. Thus, the protein backbone and
9
10 side chains should be free to relax, and water molecules migrate into the emptying binding site.
11
12 These changes can be expected to be accompanied by larger energy fluctuations. As a result,
13
14 increase of the equilibration stage for the LJ decoupling was explored yielding the results in
15
16 Figures 5 and 6. The electrostatic decoupling reached convergence after the 80M configurations
17
18 of equilibration and ca. 200M of averaging (Figure 5A). The LJ decoupling, however, still
19
20 showed a slight downward drift after 240M MC steps of averaging following the 80M
21
22 equilibration (Figure 6A); the corresponding total free-energy change for 80M/240M is shown in
23
24 Figure 6B. Restarting the averaging at this point corresponds to an equilibration of 320M
25
26 configurations. This was done and followed by another 240M configurations of averaging to
27
28 yield the results in Figure 6C, which show good convergence after ca. 100M configurations of
29
30 averaging. The total free-energy change as a function of λ for the LJ component is then shown in
31
32 Figure 6D for the 320M/240M run lengths and reflects a change of about 2 kcal/mol from Figure
33
34 6B. All the LJ windows are well converged with the 320M/240M protocol, as shown in
35
36 Supplementary Figure S4.
37
38
39
40
41
42
43
44
45
46
47
48
49

50 The free energy changes for the unbound inhibitor were 27.56 ± 0.03 kcal/mol for the
51
52 Coulombic term and 1.59 ± 0.10 kcal/mol for the LJ term, while the corresponding values for the
53
54 bound state were 34.09 ± 0.44 kcal/mol and 16.53 ± 0.60 kcal/mol with 320M/240M. The 6DoF
55
56
57

1
2
3
4 restraints correction, ΔG_{restr} , was 12.01 kcal/mol and the free energy penalty for the ligand
5
6 conformational change in TIP4P water was 0.66 kcal/mol (see below). Thus, the absolute free
7
8 energy of binding for the MIF180 complex via eq 1 is -8.80 ± 0.74 kcal/mol. The statistical
9
10 uncertainty mainly arises from the bound-state calculations and places significant limits on the
11
12 precision. The accord with the experimental value of -8.98 ± 0.28 kcal/mol³⁰ is notable; it is in
13
14 the 1-kcal/mol error range reported for relative free-energy results.^{6,53} However, many more
15
16 examples are needed with different ligands and methodological variations before general
17
18 conclusions can be reached. The following results provide some insights along these lines.
19
20
21
22
23
24
25
26
27



41
42 **Figure 5.** Electrostatic decoupling for the bound state after 80M configurations of equilibration.
43 (A) Evolution of the free energy change with simulation length. (B) Total free energy change as a
44 function of λ .
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

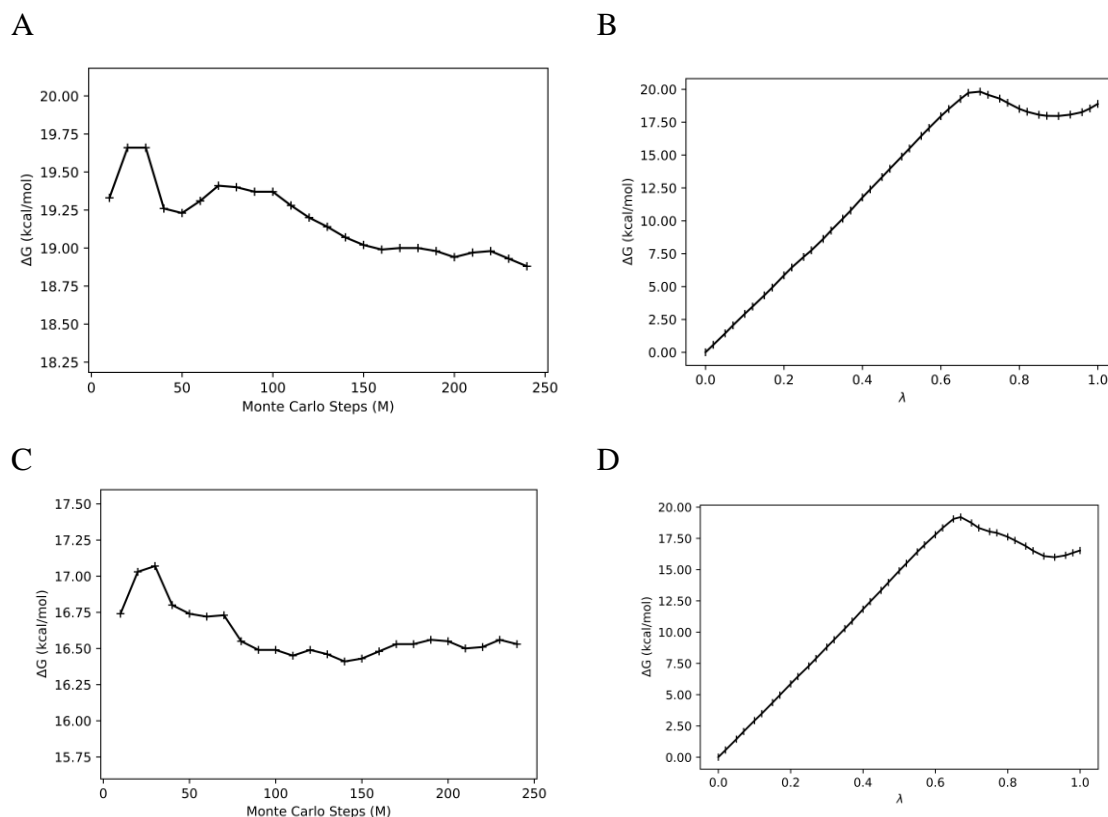


Figure 6. LJ decoupling for the bound state. (A) Evolution of the free energy change with simulation length after 80M configurations of averaging; (B) Free-energy change as a function of λ with 80M/240M; (C) Evolution of the free-energy change with simulation length after 320M configurations of equilibration; (D) Free energy change as a function of λ with 320M/240M.

Absolute Free Energy of Binding with MD

MD Calculations for the Unbound Ligand; Comparison of Small-Molecule

Force Fields. It is informative to compare results with the small-molecule force fields for computation of the absolute free energy of hydration of the MIF180 ligand. MD simulations in the CHARMM version of TIP3P water and in gas phase were used to calculate ΔG_{hyd} with the OPLS-AA/CM5, OPLS-AA/CM1A, CGenFF and GAFF force fields using NAMD. The resulting values for the electrostatic (ΔG_Q) and van der Waals (ΔG_{LJ}) components of the free

energy of hydration are listed in Table 1. The aqueous-phase results were also used to calculate ΔG_{bind} . From Table 1, the predictions for ΔG_{hyd} range over 2.7 kcal/mol, with ranges of 1.1 and 3.4 kcal/mol for the LJ and Coulombic components, respectively. Thus, the major differences likely arise from variations in the partial charges for the ligand. This is documented for the atoms with the largest differences in Figure 7 and Table 2. It can be seen that the largest variations are concentrated in the central quinolinyltriazole fragment, especially for the nitrogen atoms and C2 of the quinoline ring (C0I). The differences are substantial and show that current fixed-charge force fields are far from agreement on this important item, which has obvious implications for

Table 1: Free Energies of Hydration (kcal/mol) for MIF180 from MD Simulations

	1.20*CM5	1.14*CM1A	CGenFF	GAFF
ΔG_Q	-13.58±0.08	-14.34±0.04	-12.33±0.38	-15.76±0.08
ΔG_{LJ}	0.30±0.16	0.34±0.10	-0.78±0.27	-0.05±0.01
ΔG_{hyd}	-13.28±0.18	-14.00±0.11	-13.11±0.46	-15.81±0.08

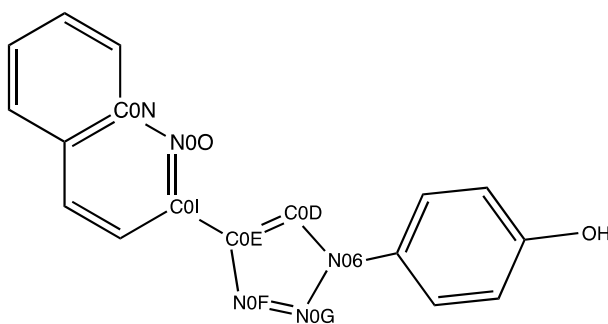


Figure 7. Atoms with the largest partial charge differences in MIF180.

Table 2: Atomic Charge Comparison for Selected Atoms

Atom Name	1.20*CM5	1.14*CM1A	CGenFF	GAFF
N06	-0.181	-0.293	0.499	0.117
C0D	0.032	-0.065	-0.443	-0.160
C0E	0.118	-0.106	0.095	0.185
N0F	-0.248	-0.113	-0.408	-0.338
N0G	-0.110	0.092	-0.332	-0.158
C0I	0.165	0.157	0.519	0.419
C0N	0.141	0.064	0.345	0.404
N0O	-0.418	-0.248	-0.625	-0.676

interactions with any surrounding water or biomolecules. In the present case, N2 and N3 of the triazole and the quinoline nitrogen all participate in hydrogen bonds with MIF (Figure 1). The magnitudes of the partial charges with CGenFF are particularly large, though this does not translate into a lower ΔG_{hyd} in Table 1. The complete list of atomic charges for the four force fields can be found in Supplementary Table S1 as well as a graphical representation of the electrostatic potentials in figure S5.

Conformation of the Unbound Ligand. To calculate the ΔG_{bind} with the present decoupling methodology, it is necessary to include a penalty for the conformational change of the ligand from *anti* in water to *syn* upon binding as reflected in Scheme 1. Thus, potential of mean force (PMF) calculations were run with NAMD using TIP3P water for each force field; the dihedral angle for the bond connecting the quinoline and triazole rings (N0F-C0E-C0I-N0O) was driven from 0° in the *syn* conformation to 180° in the *anti* conformation in both the aqueous and gas phases. The calculated PMFs are depicted in Figure 8.

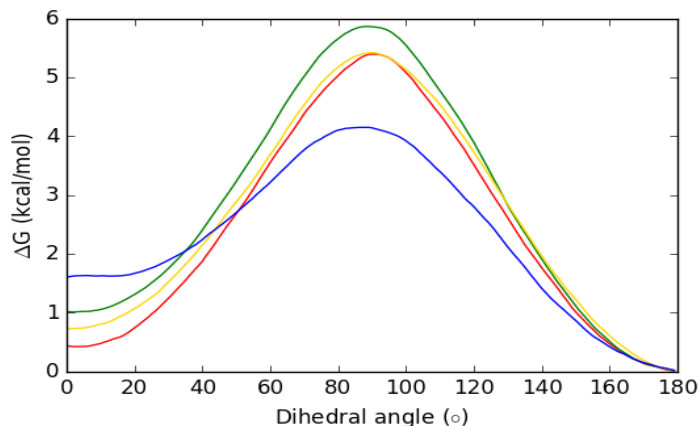
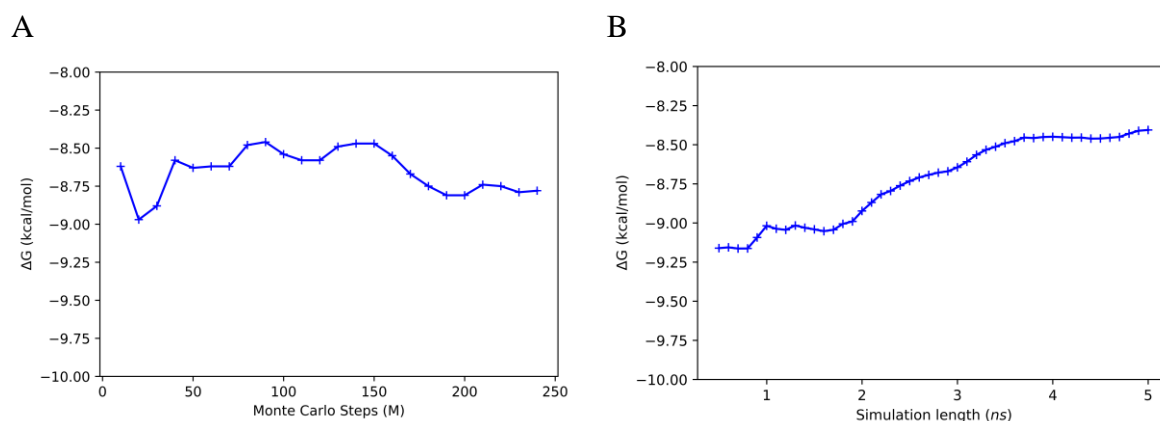


Figure 8. Free-energy profiles for the *syn* to *anti* conversion of MIF180 in TIP3P water: OPLS/CM5 (red curve), OPLS/CM1A (green), CHARMM/CGenFF (yellow) and AMBER/GAFF (blue).

In TIP3P water, the OPLS/CM5 force field gives the smallest free energy difference from *anti* to *syn* with a ΔG_{conf} of 0.44 kcal/mol. While OPLS/CM1A (1.00 kcal/mol) and CGenFF (0.71 kcal/mol) gave similar free-energy curves, GAFF showed a stronger preference for the *anti* conformer, 1.60 kcal/mol. An analogous MC/PMF calculation was done for MIF180 in TIP4P water with the OPLS/CM5 force field, which yielded a ΔG_{conf} of 0.66 kcal/mol. It may also be noted in Figure 8 that the GAFF force field gives a significantly lower *syn* to *anti* barrier. A consequence seemed to arise in the case with protonated Pro1 for which the *syn* conformation was not maintained in the protein binding site during initial MD runs using the GAFF force field. So, in this case, an additional restraint was applied to fix the *syn* preference for the complex.

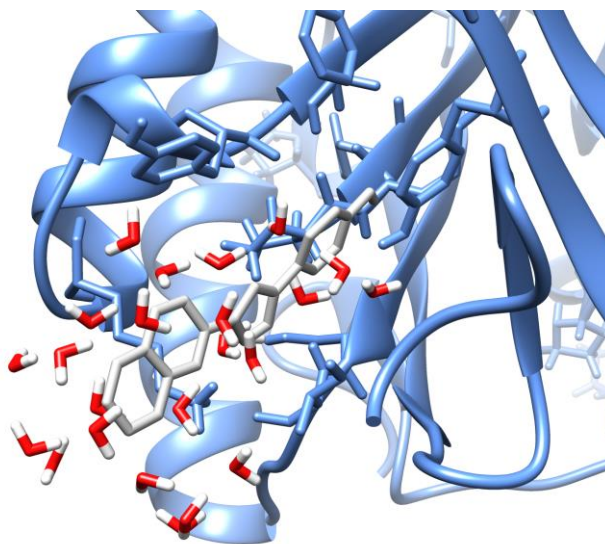
MD Results for the Bound Complex with the OPLS/CM5 Force Field. As noted above, the simulations for the bound complex were run in triplicate using 43 lambda windows

1
2
3
4 and averaging for 5 ns. With Pro1 unprotonated, the absolute free energies of binding (ΔG_{bind})
5
6
7 from the three runs were -8.35, -7.49 and -9.55 kcal/mol, to give an average of -8.46 ± 0.85
8
9 kcal/mol. The ΔG_{bind} obtained by MD/FEP matches well with the MC/FEP result, -8.80 ± 0.74
10
11 kcal/mol, and the experimental value at -8.98 ± 0.28 kcal/mol,³⁰ given the range of the
12
13 uncertainties and the methodological differences. The bound MC/FEP simulations used a TIP4P
14
15 water cap without electrostatic long-range corrections, whereas the MD/FEP runs used TIP3P
16
17 water in a periodic truncated octahedron with Ewald corrections for long-range electrostatic
18
19 interactions. Both simulation methods demonstrated a statistical uncertainty near 1 kcal/mol,
20
21 which is in line with their difference of 0.34 kcal/mol. The evolutions of the free energies of
22
23 binding are compared for MD/FEP and MC/FEP in Figure 9. On this basis, the results for both
24
25 methods appear to be converged to within ca. 0.3 kcal/mol using the 240M configurations of
26
27 averaging with MC and $3 \times 5 = 15$ ns of averaging per window with MD.
28
29
30
31
32
33
34
35
36
37
38
39



52
53 **Figure 9.** Evolution of the free energy of binding with simulation length in each window for
54 MIF180: (A) MC/FEP and (B) MD/FEP.
55
56
57

1
2
3
4 The simulations described here were designed to probe the convergence of each sampling
5
6 method and are therefore longer than typical production runs would be. Hence, the comparison
7
8 of timings and efficiency between these methods is at best qualitative at this point, and should be
9
10 used as such. In the bound calculations, a single MC run of 100M configurations required 120
11
12 core-hrs in a Xeon E5-2660 as compared to 375 core-hrs for 5 ns of MD. Since NAMD is well
13
14 parallelized, the wall-clock timing of a MD run can be easily reduced by using multiple cores. A
15
16 single complete calculation for the bound MIF-MIF180 complex utilized 33,300 and 38,700 cpu-
17
18 hrs in MC and MD simulations, respectively.
19
20
21
22
23
24



25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44 **Figure 10.** Water in binding site at the end of the MD simulation. Ghost of the ligand is in grey.

45
46
47 Concerning the migration of water into the binding site as the ligand disappears, Figure 10
48
49 shows the final configuration of the last λ window in the forward direction of the MD/FEP
50
51 calculation. Only water molecules within 5 Å of the ghost of the ligand are shown. It is clear that
52
53 water molecules penetrated well into the binding pocket during the decoupling of the bound
54
55
56
57

1
2
3
4 ligand. The average numbers of water molecules within 8 Å of N06 at the center of the ligand
5
6 were 3 and 18 and the beginning and end of the MC/FEP simulations, and 6 and 11 for MD/FEP.
7
8

9 **MD Results for the Four Force Fields and the Protonation State for Proline 1.** The

10 MD simulations for ΔG_b were performed in the same way with the other three force fields.
11
12

13 Furthermore, in order to investigate the preferred protonation state for Pro1 of MIF in the
14
15 complex, the FEP calculations were also performed with Pro1 protonated in all three MIF
16
17 monomers. The results of the ΔG_b calculations with the four force fields are summarized in Table
18
19
20
21
22

23 3 for MIF180/MIF with and without the protonated N-terminal proline. As mentioned above,
24
25 additional conformational restraint of the ligand to the *syn* conformation was needed for the
26
27 simulations with Pro1 protonated using the AMBER/GAFF force field. The results for ΔG_{bind} are
28
29
30

31 clearly very sensitive to both the choice of force field and protonation state for Pro1. In all cases,
32
33

34 ΔG_{bind} is more favorable with neutral Pro1 and closer to the experimental result of -8.98 ± 0.28
35
36

37 kcal/mol.³⁰ The OPLS/CM5, OPLS/CM1A, and AMBER/GAFF results strongly support the
38
39 assignment of Pro1 as neutral. The 2.3-kcal/mol spread of ΔG_{bind} results with neutral Pro1 from
40
41 these three alternative force fields is probably a reasonable reflection of the current state of the
42
43 art. Expecting 1 kcal/mol accuracy from any given fixed-charge force field on a specific complex
44
45 of a protein with a drug-like ligand is overly optimistic. The results with CHARMM/CGenFF are
46
47 an outlier. It is difficult to trace the origin of the problem, but it may reflect the strong variation
48
49 in partial atomic charges shown in Table 2 or uncertainties about the quality of some of the
50
51 torsional parameters noted by the on-line server.²⁸ In any event, it would be premature to draw
52
53
54
55
56
57

general conclusions about the force fields. For the OPLS force fields, there is no particular reason to favor the use of 1.20*CM5 over 1.14*CM1A charges based on results for pure organic liquids and free energies of hydration,³⁸ so the results with CM5 may just be fortunate and unique in this case.

Table 3. MD Results for Absolute Free Energies of Binding (ΔG_{bind} , kcal/mol) with Pro1 Neutral or Protonated

	OPLS/CM5	OPLS/CM1A	CHARMM/CGenFF	AMBER/GAFF
Neutral	-8.46 ± 0.85	-6.20 ± 0.41	-2.37 ± 1.27	-7.47 ± 0.99
Protonated	-3.53 ± 0.76	-0.17 ± 1.22	-2.04 ± 1.04	-2.92 ± 0.26

Structural Analyses. The crystal structure of the complex features multiple aryl-aryl interactions and hydrogen bonds as indicated in Figures 1 and 11. The phenolic hydroxyl group is hydrogen bonded with Asn97C ($r(\text{OO}) = 2.52 \text{ \AA}$), and the triazole N2 with Ile64A ($r(\text{NN}) = 2.90 \text{ \AA}$). Also, the quinoline N, triazole N3 and backbone O of Ile64A ($r(\text{NN}) = 3.33 \text{ \AA}$, $r(\text{NN}) = 2.95 \text{ \AA}$, $r(\text{NO}) = 2.81 \text{ \AA}$) are hydrogen-bonded to the ammonium nitrogen of Lys32A.²² For comparison with the MD results in TIP3P water, the five hydrogen-bond distances were averaged

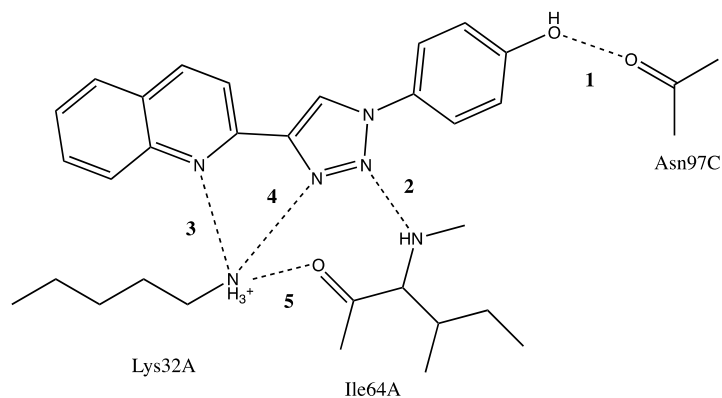


Figure 11. Intermolecular hydrogen bonds in the MIF-180 complex.

Table 4. Average Hydrogen-Bond Distances Computed Using Different Force Fields (Å)

	Crystal ^a	OPLS/ CM5	OPLS/ CM1A	CHARMM/ CGenFF	AMBER/ GAFF
1-r(OO)	2.52	2.72	2.81	2.84	2.69
2-r(NN)	2.90	3.12	4.88	4.56	3.06
3-r(NN)	3.33	4.61	5.02	7.02	3.50
4-r(NN)	2.95	4.47	4.67	5.83	3.47
5-r(NO)	2.81	3.02	3.34	4.02	2.92

^a Ref. 22.

over the final 10 ns of the MD trajectories for the fully formed complexes. The average intermolecular hydrogen-bond distances designated in Figure 11 are compared in Table 4 from the simulations with Pro1 unprotonated. Histograms for the hydrogen-bond distances showing the sampled ranges with the four force fields are provided in Supplementary Figure S2.

The shortest average protein-ligand contacts are found for OPLS/CM5 and AMBER/GAFF, which is consistent with their most favorable ΔG_{bind} results (Table 3). All of the force fields yield a hydrogen bond between the phenolic hydroxyl group and Asn97C, and a short contact for the Lys32A ammonium group and the oxygen atom of Ile64A, in agreement with the crystal structure. However, the predictions for the three N...N hydrogen bonds are varied. While AMBER/GAFF retains the three hydrogen bonds, the other three force fields do not to different degrees. The most separated structure is found for CHARMM/GenFF, which is consistent with its overly weak ΔG_{bind} in Table 3. OPLS/CM5 retains the hydrogen bond between the triazole N2 and the backbone NH of Ile64A, while the coordination of the Lys32A ammonium group with N3 of the triazole and the quinoline nitrogen atom is weakened. This may be reasonable since

1
2
3
4 Lys32A is at the entrance to the binding site and is largely solvent-exposed. There are also
5
6 interprotein contacts in this region in the crystal structures,²² which provide for some exclusion
7
8 of water and may lead to differences in structure for the crystal and dilute aqueous solution.
9
10

11
12 The corresponding results with Pro1 protonated are provided in Supplementary Table S3 and
13
14 Figure S2. In this case, the ligand is much more separated from the protein. Basically, all of the
15
16 hydrogen bonds are broken except for the one with the phenolic oxygen atom with
17
18 OPLS/CM1A, CHARMM/CGenFF, and AMBER/GAFF.
19
20
21
22
23

24 CONCLUSION

25
26
27 Computation of the absolute free energy of binding (ΔG_{bind}) for the complex of a drug-like
28
29 ligand, MIF180, and human MIF has been investigated with both Monte Carlo statistical
30
31 mechanics and molecular dynamics using double decoupling and four current fixed-charge force
32
33 fields. Both MC and MD protocols were devised that yielded well converged ΔG_{bind} values,
34
35 though more efficient protocols using fewer λ -windows may be possible.⁵⁴ The MC/FEP protocol
36
37 with improved sampling techniques²⁰ and the OPLS/CM5 force field performed well and gave an
38
39 accurate estimate for ΔG_{bind} in comparison to the experimental data. It was confirmed that a very
40
41 similar result is obtained using the same force field in molecular dynamics simulations with the
42
43 NAMD program. The MD/FEP calculations were then carried out for three additional force fields
44
45 OPLS/CM1A, CHARMM 36 with CGenFF, and AMBER ff14sb with GAFF. The results for
46
47 ΔG_{bind} notably cover a 6 kcal/mol range, though three of the results are within 2.2 kcal/mol
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 (Table 3). Significant differences in the computed structures for the complexes are also found
5
6 with the general observation that shorter average protein-ligand contacts do correlate with more
7
8 negative ΔG_{bind} values. Many additional studies of this type are needed to make such
9
10 computations more routine, to identify optimal protocols, and to reveal unambiguously any
11
12 problematic issues for current force fields and sampling methods. It is proposed that human MIF
13
14 is a good test system for such work owing to its moderate size and to the availability of multiple
15
16 high-resolution crystal structures as well as accurate binding data for numerous, diverse
17
18 inhibitors.^{21,22,30}
19
20
21
22
23
24
25
26
27

28 ASSOCIATED CONTENT

29
30
31 **Supporting Information.** This information is available free of charge on the ACS Publications
32
33 website at <http://pubs.acs.org>. It consists of a pdf file showing the atom label assignments for
34
35 MIF180, averages and histograms of hydrogen-bond distances for the complexes, free energy
36
37 evolution of the LJ MC windows, and full comparison of ligand atomic charges. In addition, a
38
39 zip archive containing the pdb, psf, and prm files needed to run the ligand MD calculations with
40
41 all four force fields is provided.
42
43
44
45
46
47
48
49
50

51 AUTHOR INFORMATION

52 53 54 Corresponding Author

55
56
57
58
59
60

*E-mail: william.jorgensen@yale.edu, phone: (203)432-6278

ORCID

Yue Qian: 0000-0003-1989-7586

Israel Cabeza de Vaca: 0000-0002-6208-1091

Daniel J. Cole: 0000-0003-2933-0719

Julian Tirado-Rives: 0000-0001-7330-189X

William L. Jorgensen: 0000-0002-3993-9520

Acknowledgements

This work was supported by the National Institutes of Health (GM32136). Computational resources were provided by the Yale University Faculty of Arts and Sciences High-Performance Computing Center.

Notes

The authors declare no competing financial interest.

REFERENCES:

1. Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. Alchemical Free Energy Methods for Drug Discovery: Progress and Challenges. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150–160.
2. Jorgensen, W. L. Computer-Aided Discovery of Anti-HIV Agents. *Bioorg. Med. Chem.* **2016**, *24*, 4768-4788.
3. Mobley, D. L.; Gilson, M. K. Predicting Binding Free Energies: Frontiers and Benchmarks. *Ann. Rev. Biophys.* **2017**, *46*, 531-558.
4. Abel, R.; Wang, L.; Mobley, D. L.; Friesner, R. A. A Critical Review of Validation, Blind Testing, and Real-World Use of Alchemical Protein-Ligand Binding Free Energy Calculations. *Curr. Top. Med. Chem.* **2017**, *17*, 1-9.
5. Cournia, Z.; Allen, B.; Sherman, W. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J. Chem. Inf. Model.* **2017**, *57*, 2911–2937.
6. Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; *et al.* Accurate and Reliable Prediction of Relative Protein-Ligand Binding Affinities via Free Energy Calculations: Validation in Prospective Drug Discovery. *J. Am. Chem. Soc.* **2015**, *137*, 2695-2703.
7. Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe Jr, E. W. Computational Methods in Drug Discovery. *Pharmacol. Rev.* **2014**, *66*, 334–395.
8. Jorgensen, W. L.; Buckner, J. K.; Boudon, S.; Tirado-Rives, J. Efficient Computation of Absolute Free Energies of Binding by Computer Simulations. Application to the Methane Dimer in Water. *J. Chem. Phys.* **1988**, *89*, 3742-3746.
9. Cieplak, P.; Kollman, P. A. Calculation of the Free Energy of Association of Nucleic Acid Bases in Vacuo and Water Solution. *J. Am. Chem. Soc.* **1988**, *110*, 3734–3739.
10. Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. The Statistical-

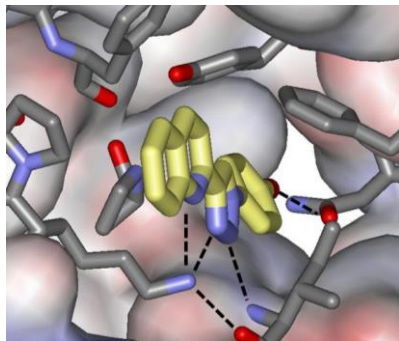
- 1
2
3
4 Thermodynamic Basis for Computation of Binding Affinities: A Critical Review. *Biophys.*
5 *J.* **1997**, *72*, 1047–1069.
6
7
8 11. Deng, Y.; Roux, B. Calculation of Standard Binding Free Energies: Aromatic Molecules in
9 the T4 Lysozyme L99A Mutant. *J. Chem. Theory Comput.* **2006**, *2*, 1255–1273.
10
11
12 12. Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. Absolute Binding Free Energies: A
13 Quantitative Approach for Their Calculation. *J. Phys. Chem. B* **2003**, *107*, 9535–9551.
14
15
16 13. Gaieb, Z.; Liu, S.; Gathiaka, S.; Chiu, M.; Yang, H.; Shao, C.; Feher, V. A.; Walters, W. P.;
17 Kuhn, B.; Rudolph, M. G.; *et al.* D3R Grand Challenge 2: Blind Prediction of Protein-
18 Ligand Poses, Affinity Rankings, and Relative binding free energies. *J. Comput. Aided*
19 *Mol. Design* **2018**, *32*, 1-20.
20
21
22
23
24 14. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindah, E.
25 Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from
26 Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.
27
28
29
30 15. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.;
31 Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J.*
32 *Comput. Chem.* **2005**, *26*, 1668–1688.
33
34
35
36 16. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.;
37 Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J.*
38 *Comput. Chem.* **2005**, *26*, 1781–1802.
39
40
41
42 17. Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux,
43 B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; *et al.* CHARMM: The Biomolecular
44 Simulation Program B. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
45
46
47
48 18. Eastman, P.; Friedrichs, M. S.; Chodera, J. D.; Radmer, R. J.; Bruns, C. M.; Ku, J. P.;
49 Beauchamp, K. A.; Lane, T. J.; Wang, L. P.; Shukla, D.; *et al.* OpenMM 4: A Reusable,
50 Extensible, Hardware Independent Library for High Performance Molecular Simulation. *J.*
51 *Chem. Theory Comput.* **2013**, *9*, 461–469.
52
53
54
55
56
57
58
59
60

19. Jorgensen, W. L.; Tirado-Rives, J. Monte Carlo vs Molecular Dynamics for Conformational Sampling. *J. Phys. Chem.* **1996**, *100*, 14508–14513.
20. Cabeza de Vaca, I.; Qian, Y.; Vilseck, J. Z.; Tirado-Rives, J.; Jorgensen, W. L. Enhanced Monte Carlo Methods for Modeling Proteins Including Computation of Absolute Free Energies of Binding. *J. Chem. Theory Comput.* **2018**, *14*, 3279–3288.
21. Trivedi-Parmar, V.; Robertson, M. J.; Cisneros, J. A.; Krimmer, S. G.; Jorgensen, W. L. Optimization of Pyrazoles as Phenol Surrogates to Yield Potent Inhibitors of Macrophage Migration Inhibitory Factor. *Chem. Med. Chem.* **2018**, *13*, 1092-1097.
22. Dziedzic, P.; Cisneros, J. A.; Robertson, M. J.; Hare, A. A.; Danford, N. E.; Baxter, R. H. G.; Jorgensen, W. L.; Nadia, E.; Baxter, R. H. G.; Jorgensen, W. L. Design, Synthesis, and Protein Crystallography of Biaryltriazoles as Potent Tautomerase Inhibitors of Macrophage Migration Inhibitory Factor. *J. Am. Chem. Soc.* **2015**, *137*, 2996–3003.
23. Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L. Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. *J. Chem. Theory Comput.* **2015**, *11*, 3499–3509.
24. Robertson, M. J.; Qian, Y.; Robinson, M. C.; Tirado-Rives, J.; Jorgensen, W. L. Development and Testing of the OPLS-AA/M Force Field for RNA. *J. Chem. Theory Comput.* **2019**, *15*, 2734-2742.
25. Dodda, L. S.; Cabeza de Vaca, I.; Tirado-Rives, J.; Jorgensen, W. L. LigParGen Web Server: An Automatic OPLS-AA Parameter Generator for Organic Molecules. *Nucl. Acids Res.* **2017**, *45*, W331-W336.
26. Best, R. B.; Xiao Zhu, X.; Shim, J.; Lopes, P.; Mittal, J.; Feig, M.; MacKerell, A. D. Jr.; Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone Phi, Psi, and Side Chain Chi1 and Chi2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257-3273.
27. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C.

- 1
2
3
4 Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from
5 Ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
- 6
7
8 28. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian,
9 E.; Guvench, O.; Lopes, P.; Vorobyov, I.; *et al.* CHARMM General Force Field: A Force
10 Field for Drug-like Molecules Compatible with the CHARMM All-Atom Additive
11 Biological Force Fields. *J. Comput. Chem.* **2009**, *31*, 671–690.
- 12
13
14
15
16 29. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and
17 Testing of a General Amber Force Field. *J Comput Chem* **2004**, *25*, 1157–1174.
- 18
19
20 30. Cisneros, J. A.; Robertson, M. J.; Valhondo, M.; Jorgensen, W. L. A Fluorescence
21 Polarization Assay for Binding to Macrophage Migration Inhibitory Factor and Crystal
22 Structures for Complexes of Two Potent Inhibitors. *J. Am. Chem. Soc.* **2016**, *138*, 8630–
23 8638.
- 24
25
26
27
28 31. Fu, H.; Cai, W.; Hémin, J.; Roux, B.; Chipot, C. New Coarse Variables for the Accurate
29 Determination of Standard Binding Free Energies. *J. Chem. Theory Comput.* **2017**, *13*,
30 5173–5178.
- 31
32
33
34 32. Woo, H. H.-J.; Roux, B. Calculation of Absolute Protein-Ligand Binding Free Energy
35 from Computer Simulations. *Proc. Natl. Acad. Sci.* **2005**, *102*, 6825–6830.
- 36
37
38 33. Gumbart, J. C.; Roux, B.; Chipot, C. Standard Binding Free Energies from Computer
39 Simulations: What Is the Best Strategy? *J. Chem. Theory Comput.* **2013**, *9*, 794–802.
- 40
41
42 34. Gumbart, J. C.; Roux, B.; Chipot, C. Efficient Determination of Protein–Protein Standard
43 Binding Free Energies from First Principles. *J. Chem. Theory Comput.* **2013**, *9*, 3789–
44 3798.
- 45
46
47
48 35. Wang, J.; Deng, Y.; Roux, B. Absolute Binding Free Energy Calculations Using Molecular
49 Dynamics Simulations with Restraining Potentials. *Biophys J* **2006**, *91*, 2798–2814.
- 50
51
52 36. Fiorin, G.; Klein, M. L.; Hémin, J. Using Collective Variables to Drive Molecular
53 Dynamics Simulations. *Mol. Phys.* **2013**, *111*, 3345–3362.
- 54
55
56
57
58
59
60

- 1
2
3
4 37. Fu, H.; Gumbart, J. C.; Chen, H.; Shao, X.; Cai, W.; Chipot, C. BFEE: A User-Friendly
5 Graphical Interface Facilitating Absolute Binding Free-Energy Calculations. *J. Chem. Inf.*
6 *Model.* **2018**, *58*, 556-560.
7
8
9
10 38. Dodda, L. S.; Vilseck, J. Z.; Cutrona, K. J.; Jorgensen, W. L. Evaluation of CM5 Charges
11 for Nonaqueous Condensed-Phase Modeling. *J. Chem. Theory Comput.* **2015**, *11*, 4273-
12 4282.
13
14
15
16 39. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type
17 Perception in Molecular Mechanical Calculations. *J. Mol. Graph. Model.* **2006**, *25*, 247-
18 260.
19
20
21
22 40. Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H++: A Server
23 for Estimating pK_as and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids*
24 *Res.* **2005**, *33*, W368–W371.
25
26
27
28 41. Anandkrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: Automating pK Prediction and
29 the Preparation of Biomolecular Structures for Atomistic Molecular Modeling and
30 Simulations. *Nucleic Acids Res.* **2012**, *40*, W537–W541.
31
32
33
34 42. Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. Improved Treatment
35 of Ligands and Coupling Effects in Empirical Calculation and Rationalization of p K a
36 Values. *J. Chem. Theory Comput.* **2011**, *7*, 2284–2295.
37
38
39
40 43. Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3:
41 Consistent Treatment of Internal and Surface Residues in Empirical p K a Predictions. *J.*
42 *Chem. Theory Comput.* **2011**, *7*, 525–537.
43
44
45
46 44. Jorgensen, W. L.; Thomas, L. L. Perspective on Free-Energy Perturbation Calculations for
47 Chemical Equilibria. *J. Chem. Theory Comput.* **2008**, *4*, 869–876.
48
49
50 45. Steinbrecher, T.; Mobley, D. L.; Case, D. A. Nonlinear Scaling Schemes for Lennard-
51 Jones Interactions in Free Energy Calculations. *J. Chem. Phys.* **2007**, *127*, 214108.
52
53
54 46. Mobley, D. L.; Chodera, J. D.; Dill, K. A. On the Use of Orientational Restraints and
55
56
57
58
59
60

- 1
2
3
4 Symmetry Corrections in Alchemical Free Energy Calculations. *J. Chem. Phys.* **2006**, *125*,
5 084902.
6
7
8 47. Deng, Y.; Roux, B. Computations of Standard Binding Free Energies with Molecular
9 Dynamics Simulations. *J. Phys. Chem. B* **2009**, *113*, 2234–2246.
10
11
12 48. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L.
13 Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.*
14 **1983**, *79*, 926–935.
15
16
17
18 49. van Gunsteren, W. F.; Berendsen, H. J. C. A Leap-Frog Algorithm for Stochastic
19 Dynamics. *Mol. Simul.* **1988**, *1*, 173–185.
20
21
22 50. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian
23 Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J.*
24 *Comput. Phys.* **1977**, *23*, 327–341.
25
26
27
28 51. Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An $N \cdot \log N$. Method for Ewald
29 Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
30
31
32 52. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth
33 Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
34
35
36 53. Perez, A.; Morrone, J. A.; Simmerling, C.; Dill, K. A. Advances in Free-Energy-Based
37 Simulations of Protein Folding and Ligand Binding. *Curr. Opin. Struct. Biol.* **2016**, *36*,
38 25–31.
39
40
41
42 54. Cabeza de Vaca, I.; Zarzuela, R.; Tirado-Rives, J.; Jorgensen, W. L. Robust FEP Protocols
43 for Creating Molecules in Solution. *J. Chem. Theory Comput.* **2019**, *15*, 3941-3948.
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



TOC Graphic.