Modelling Flexible Protein-Ligand Binding in p38α MAP Kinase using the QUBE Force Field

Joshua T. Horton,[†] Alice E. A. Allen,[‡] and Daniel J. Cole[*,†]

†*School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne NE1 7RU, United Kingdom*

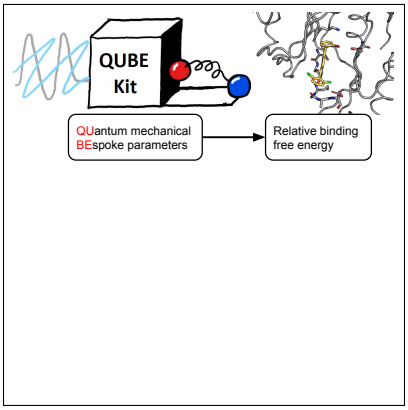
‡*Engineering Laboratory, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom*

E-mail: *daniel.cole@ncl.ac.uk

Abstract

The quantum mechanical bespoke (QUBE) force field is used to retrospectively calculate the relative binding free energy of a series of 17 flexible inhibitors of p38α MAP kinase. The size and flexibility of the chosen molecules represent a stringent test of the derivation of force field parameters from quantum mechanics, and enhanced sampling is required to reduce the dependence of the results on the starting structure. Competitive accuracy with a widely-used biological force field is achieved, indicating that quantum mechanics derived force fields are approaching the accuracy required to provide guidance in prospective drug discovery campaigns.

Graphical TOC Entry



Introduction

Free energy perturbation (FEP) based on molecular mechanics (MM) simulations can be an effective guide during the hit-to-lead stages of a drug design campaign as it provides a formally rigorous means to compute protein-ligand binding free energies.[1–3] In practice, the predictive ability of such simulations is effectively limited by two major factors, 1) the accuracy of the underlying MM force field that is used for the rapid calculation of the system energy, and 2) finite simulation times that can limit the conformational space explored.[4] In the expectation of making such calculations routinely reliable, the development of enhanced sampling methods is an active area of research,[5,6] yet virtually all FEP simulations employ transferable biological force fields, such as AMBER, OPLS, GROMOS and CHARMM, all with quite similar functional forms and parameter fitting strategies.[7] These biological force fields, alongside their small molecule counterparts, have had wide success to-date thanks to meticulous fitting of parameters to reproduce quantum mechanical (QM) and experimental properties of sets of small organic molecules. However, there is room for improvement.[8–10] It is widely acknowledged that atomic point charges are sensitive to their (local and long-ranged) environment, which is why small molecule force fields typically employ atomic charges that are fit to the molecular electrostatic properties (e.g. ESP or CMx charges[11]), on a case-by-case basis. Interestingly, this leads to a disconnect between protein and small molecule force fields, in which the former sets of atomic charges are read from a transferable library, and the latter are derived using methods that are not always consistent with the underlying biological force field. Also, standard libraries of parameters describing torsional rotation about flexible bonds are often blamed for observation of unphysical conformations in MM simulations, and these parameters are often re-derived specifically for the molecule under study.[12–15]

With regards to these issues, there has been recent interest in molecule-specific, or bespoke, force fields in which the parameters that govern the dynamics of the system are not assigned from a library based on predetermined atom types, but instead are inferred directly

from QM calculations specifically for the molecule of interest.[15–18] One such example is the QUantum mechanical BEspoke (QUBE) force field,[18] which has a particular focus on scalability to large system sizes and applications in the condensed phase.[19,20] The QUBE force field shares its functional form with OPLS, so that it retains the favorable computational efficiency of transferable force fields, but differs in that as many parameters as feasible are derived directly from routine, molecule-specific QM calculations. The ground state electron density of the molecule under study is first computed in a weak implicit solvent to simulate the effect of environmental polarization.[19] The density is then partitioned into a set of approximately spherical atom-centered basins via the density derived electrostatic and chemical (DDEC) atoms-in-molecule (AIM) approach,[21,22] from which we compute the environment-specific non-bonded parameters, including (atom-centered and off site) atomic charges and Lennard-Jones parameters.[18,19] Since the DDEC method is implemented in the linear-scaling density functional theory code, ONETEP,[23] we can derive these parameters consistently for both small molecules and also systems comprising thousands of atoms, such as proteins.[24,25] QUBE bond and angle force field parameters are derived directly from the QM Hessian matrix of small molecules, as described previously,[26] and flexible torsions may be parametrized by fitting to constrained one-dimensional QM dihedral scans.[18] Parameter assignment is automated by the QUBEKit software package.[18]

To date, the first generation of the QUBE force field has undergone extensive benchmarking against established performance metrics, such as the prediction of the condensed phase thermodynamic properties (density, heat of vaporization and free energy of hydration) of over 100 small organic molecules.[18] A custom library of bonded parameters for protein simulations has been developed and validated via the comparison of molecular dynamics trajectories with NMR observables.[20] In all of these cases, QUBE performed to a similar standard as established and optimized transferable force fields. In the context of FEP calculations, QUBE has been applied to the study of the benchmark L99A mutant of T4 lysozyme, achieving a mean unsigned error (MUE) of 0.85 kcal/mol in the prediction of the absolute

binding free energies of six benzene derivatives. However, typical hit-to-lead studies in drug discovery scenarios are significantly more complex than the above study in terms of the sizes of the ligands, the nature of their interactions, and their conformational flexibility.[2,27,28] In this letter, we therefore retrospectively calculate the relative binding free energies of a series of 17 drug-like inhibitors of p38α MAP kinase (Figure 1). This represents a typical optimization scenario involving both polar and non-polar substitutions around a benzene ring, with activities that span 2–3 orders of magnitude (Table 1). As we shall discuss, the binding pose is determined to a large extent by two flexible dihedral angles (ϕ_1 and ϕ_2, Figure 1), which impose complex sampling requirements on the simulations. This set of transformations has been the target for a range of activity prediction methods including FEP calculations, which were used to demonstrate the importance of the initial water placement during Monte Carlo (MC) simulations using the OPLS force field.[29]

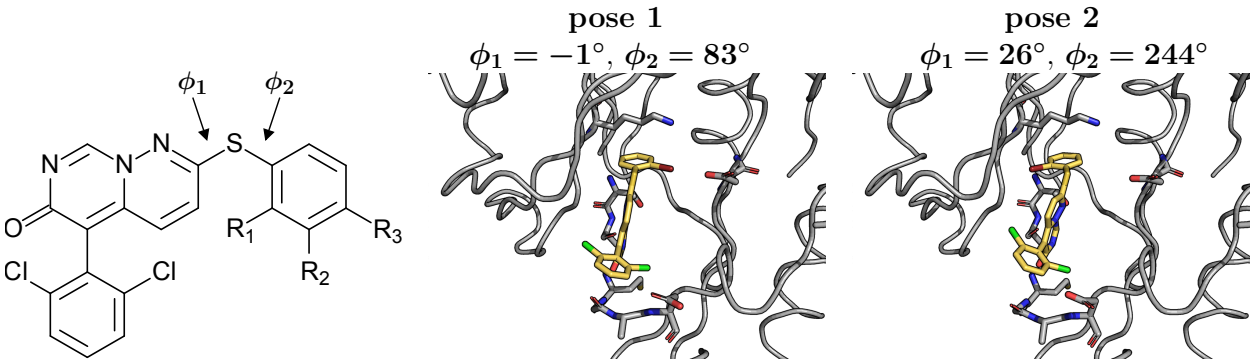


Figure 1: (left) Core structure of the p38α MAP kinase inhibitors studied here. Key flexible dihedrals (ϕ_1 and ϕ_2) are labelled. (right) Snapshots from FEP MC simulations of ligand **12** (yellow) highlighting binding poses 1 and 2.

Here, we begin by parametrizing molecule-specific force fields for the 17 p38 kinase inhibitors (plus compound **18**, which does not have experimental data for comparison but is a useful FEP intermediate) using the QUBEKit software.[18] Namely, non-bonded (charge and Lennard-Jones) parameters are derived using atoms-in-molecule partitioning of the ground state electron density as described previously.[18,19] Parametrization of the protein non-bonded parameters is performed using the same protocols, while bonded parameters are read in from

Table 1: List of p38α MAP kinase inhibitors with their experimentally measured IC_{50} activities.[29]

Compound	R_1	R_2	R_3	pIC_{50}
1	H	H	H	6.602
2	H	H	F	7.000
3	H	H	CH_3	5.854
4	H	Cl	Cl	6.097
5	H	CH_3	H	5.854
6	H	CH_3	CH_3	5.721
7	H	F	H	6.347
8	CH_3	H	H	6.699
9	H	Cl	F	6.301
10	H	Cl	H	6.553
11	CH_3	H	Cl	6.745
12	Br	H	H	6.602
13	CH_3	H	CH_3	6.577
14	OH	H	H	6.444
15	NH_2	H	F	6.658
16	Cl	H	F	7.444
17	F	F	F	8.046
18	F	H	H	N/A

a custom library.[20] Bond and angle parameters of the small molecules (**1**–**18**) are derived using the modified Seminario method computed using the QM Hessian matrix at the optimized geometries.[26] Finally, parameters describing rotation about the two flexible dihedral angles ϕ_1 and ϕ_2 are fit to constrained QM potential energy scans. Figure 2 shows the results of the torsion parameter optimization for ligand **1**. The fit to the underlying QM data is very good with an average root mean square deviation between sampled QM and QUBE torsional scans of 0.07 kcal/mol. For comparison, typical errors in excess of 1.5 kcal/mol are observed using small molecule transferable force fields.[30]

By deriving the QUBE force field directly from QM, our goal is to provide accurate and automated molecule-specific parameters that reproduce as closely as possible the full QM potential energy surface. Figure 3 shows the correlation between QUBE and QM relative energies of structures **3** and **10** extracted from Monte Carlo simulations (see later). The correlation between QUBE and QM energetics is similar to that previously reported,[18] and

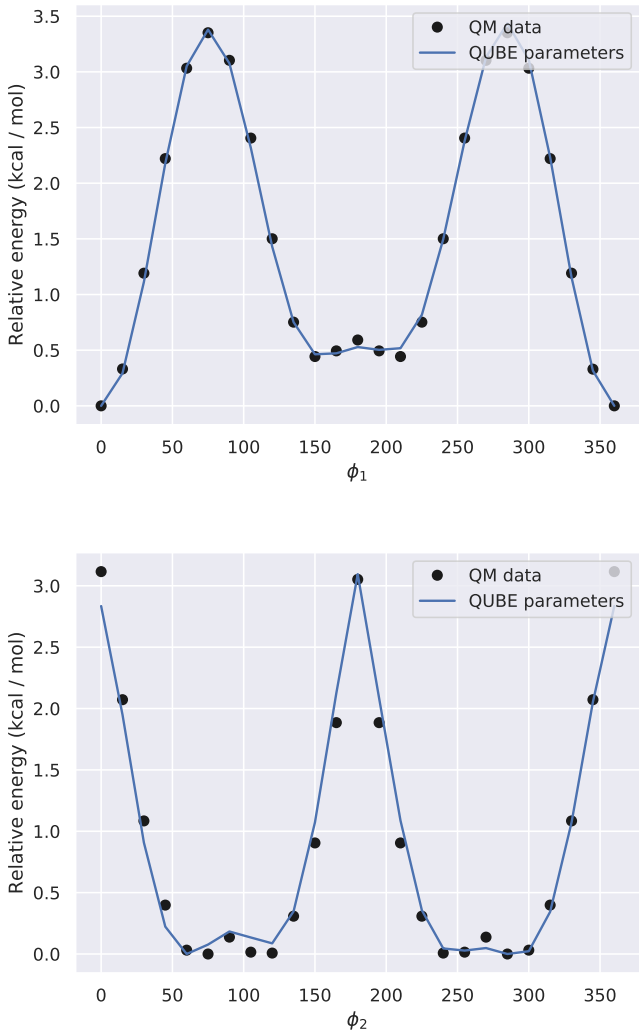


Figure 2: QUBE and QM potential energy surfaces for ligand **1** upon rotation of flexible dihedrals ϕ_1 and ϕ_2.

significantly QUBE does not predict any physically unreasonable structures (either bound to the protein or in water) whilst retaining the fixed MM functional form that provides us with a practical method for deployment in free energy predictions. Additional analysis of torsion scans and correlations between QM and QUBE energetics for the remaining ligands may be found in the Supporting Information.

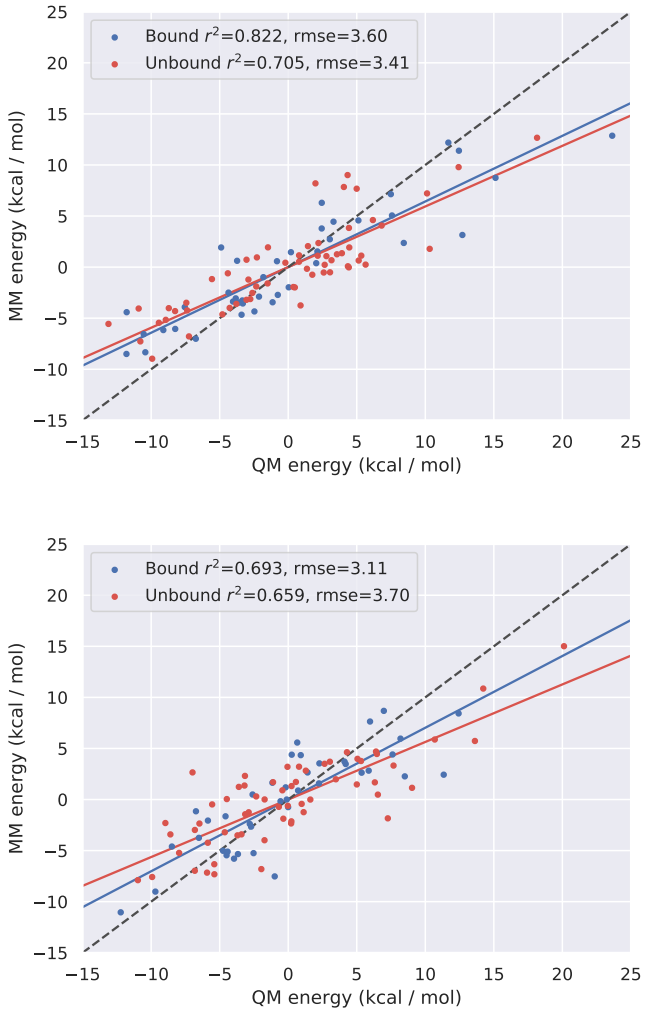


Figure 3: Comparison between QUBE and QM single point energies of structures of **3** (top) and **10** (bottom) extracted from bound and unbound (in water) MC simulations. The mean energies of each distribution have been shifted to zero. Also shown are the correlation (r^2) and root mean square errors (rmse, kcal/mol) between the two distributions.

Having parametrized the 18 inhibitors, we turn now to the computation of their relative binding free energies to p38 kinase. Free energy calculations were performed using the

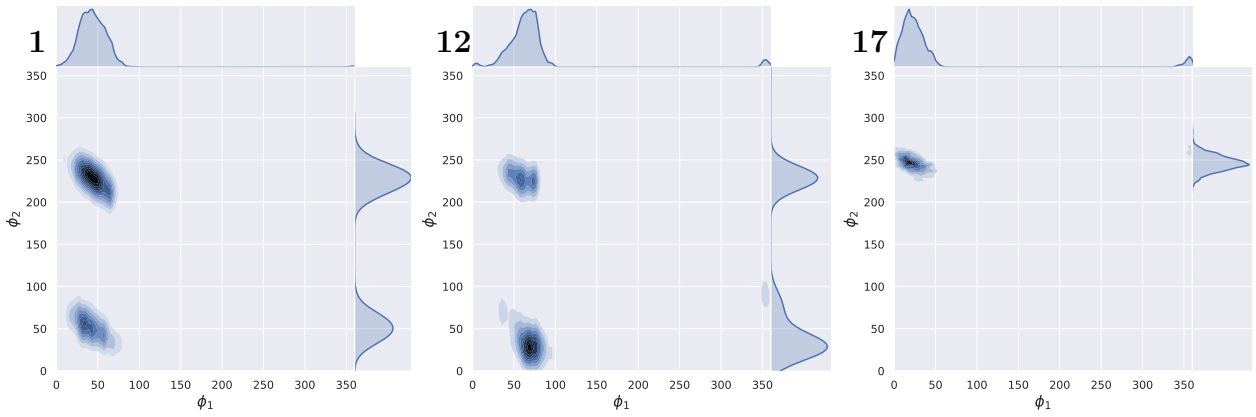


Figure 4: Two-dimensional dihedral distributions observed during the protein-ligand complex simulations of ligands **1**, **12** and **17**. See also Figure 1 for indicative poses.

MCPRO software.[31] The ligand binding site is expected to be hydrated, and so the JAWS water placement algorithm[32] was used to optimize the initial solvent distribution. As reported previously,[29] the majority of the ligands **1**-**17** are expected to bind in pose 1 (Figure 1). Hence, we set up the ligands initially in pose 1, but employed the replica exchange with solute tempering (REST) enhanced sampling method with the goal of reducing the dependence of the computed binding free energies on the starting conditions. Importantly, in MCPRO, the REST algorithm may be employed alongside the 'flip' protocol, in which selected dihedral angles (here, ϕ_1 and ϕ_2) undergo Monte Carlo moves that are much larger than typical. Figure 4 illustrates the effects of this sampling procedure. Ligand **1** is symmetric under 180° flips in ϕ_2, and indeed approximately equal distributions of the two conformers are observed at $\phi_2 = 40°$ (pose 1) and $\phi_2 = 220°$ (pose 2). Interestingly, despite starting in pose 1, **17** shows a single peak at $\phi_2 = 250°$, indicating a strong preference for pose 2. This agrees with previous observations using the OPLS force field,[29] and x-ray crystal structures of similar ligands[33] (Figure 5). Of note, in that former study, MC simulations were required starting from both poses 1 and 2 since interconversion between the two is not expected during these simulations using either standard MC or molecular dynamics. In contrast, the use of the REST/flip algorithm facilitates binding mode determination and free energy prediction from a single MC run. Despite being asymmetric, **12** shows similar behavior to **1**, with peaks

around $\phi_2 = 30°$ and $\phi_2 = 210°$ (Figure 4). This is perhaps reasonable, since **12** is similar in chemistry to **17**, but the bulkier Br atom may hinder full inclusion into the pose 2 binding pocket. Overall, we conclude that using the QUBE force field and REST enhanced sampling algorithm described here, the asymmetric ligands **4**–**11** and **13**–**15** bind in pose 1, ligands **16**–**18** bind in pose 2, and ligand **12** is intermediate between the two.

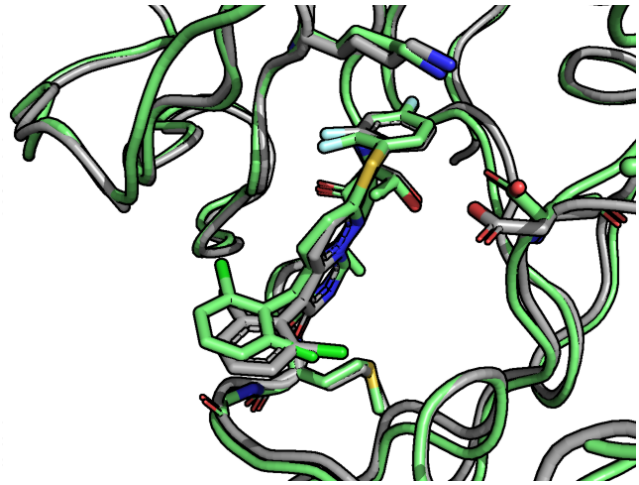


Figure 5: Overlay of the crystal structure (PDBID: 3FC1, gray) with the last snapshot (green) of the MC simulation of **17** bound to p38α MAP kinase.

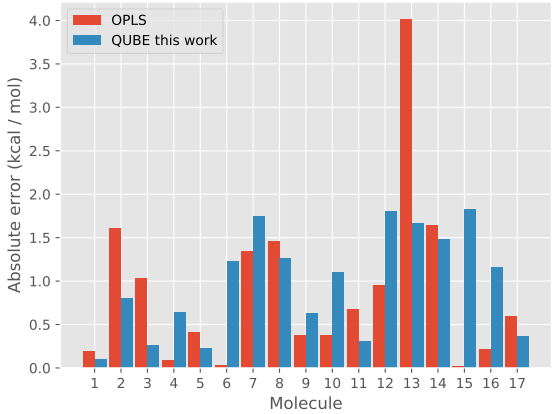


Figure 6: Absolute errors in predicted relative binding free energies computed using the QUBE and OPLS[29] force fields, compared to experiment.

Having elucidated the preferred binding poses of the 17 inhibitors, we turn now to the

prediction of protein-ligand relative binding free energies. Figure 6 compares the errors in the relative binding free energies computed using the QUBE protein/ligand force field with experiment. For comparison, the corresponding quantities are also displayed for the OPLS force field from previous work.[29] Full details of the transformations and computed free energies are given in the Supporting Information. Overall, the mean unsigned error (MUE) is 0.98 kcal/mol, which is competitive with the generally accepted accuracy of standard biological force fields for transformations of this type[2] and, in particular, with previous calculations using the OPLS force field on this system (0.88 kcal/mol). The largest errors, using the QUBE force field are for ligands **12**–**15**, which all include bulky and/or polar substituents at the R_1 position (Figure 1), as well as ligand **7**. The torsional profiles of these ligands are all reasonable, and so it seems likely that non-bonded interactions and/or sampling errors are to blame. We have found previously that QUBE can underestimate hydration free energies of some molecules containing bulky hydrophobic and hydroxyl functional groups by up to around 2 kcal/mol.[18] Other possibilities, highlighted by Luccarelli et al.,[29] are that changes in solvent distribution in the binding pocket and/or protein side chain conformational changes are not properly sampled during alchemical perturbation. To investigate the adequacy of the REST method for sampling the complex binding mode of **12**, we re-ran the **12**→**18** and **18**→**1** transformations starting with the ligand in pose 2. However, the error in the relative binding free energy of **12** fell only from 1.8 kcal/mol to 1.5 kcal/mol, indicating that the binding mode is sufficiently sampled during our simulations.

In summary, we have benchmarked the accuracy of the QUBE force field against relative binding free energies of 17 drug-like inhibitors of p38α MAP kinase. The selected protein-ligand complex includes challenges due to sampling of protein-ligand binding modes and binding site hydration, and is therefore representative of typical hit-to-lead optimization projects. The mean unsigned error of 0.98 kcal/mol of the first generation of QUBE is competitive with widely-used biological force fields, and encouragingly the crystallographic binding pose of **17** was obtained despite starting from an alternative structure. More gen-

erally, the FEP/REST enhanced sampling protocol employed here allowed us to obtain all predictions starting from a single binding pose, in contrast to previous studies that required two.[29] One current disadvantage of QUBE is parameterization time, which can be of the same order of magnitude as the free energy calculation itself. However, there is future scope for the use of, for example, fragmentation schemes for reducing the computational expense of torsion scans and machine learning methods for non-bonded parameter assignment,[34] especially when employed in congeneric series of ligands such as this one. Meanwhile, a wide range of accuracy improvements are envisaged, from the use of off-site charges in relative binding free energy calculations to improve the description of electron density anisotropy,[18] to improved descriptions of polarization, van der Waals and short-range repulsion using advanced force field functional forms.[35,36] Future work will continue to improve the accuracy and throughput of the QUBE force field for binding free energy applications in prospective medicinal chemistry efforts.

Computational Methods

Input structures for the complexes between p38α MAP kinase and the 18 inhibitors were prepared starting from the crystal structure (PDB: 1OUY[33]) as described in the Supporting Information using the MCPRO 3.2[31] and BOMB[37] software packages. Ligand force fields were parametrized using the QUBEKit software package.[18] Quantum chemistry geometry optimizations and frequency calculations were performed in Gaussian09[38] using the ωB97XD functional and 6-311++G(d,p) basis set. Equilibrium bond lengths and angles were extracted from the QM optimized geometry, and the bond-stretching and angle-bending force constants were derived from the QM Hessian matrix via the modified Seminario method with a vibrational scaling factor of 0.957.[26] Constrained one-dimensional torsional optimizations were also performed using Gaussian09, with the same level of theory and basis set, in 15° increments from 0° to 360°. Torsion parameter optimizations of dihedrals ϕ_1 and

ϕ_2 were performed for each ligand separately using QUBEKit with no Boltzmann weighting or regularization.[18] OPLS atom types were retained during torsion fitting to reduce the parameter search space, while all remaining small molecule torsion parameters were taken from the OPLS force field. Non-bonded parameter assignment was performed for both small molecules and the protein (2961 atoms) using the ONETEP linear-scaling density functional theory code and DDEC AIM analysis (Supporting Information). All bonded parameters of the protein were assigned from a transferable library that has been specifically designed to be compatible with the QUBE FF.[20] Water molecules were described using the TIP4P water model. FEP/REST calculations were performed using the MCPRO software, version 3.2, which includes recent improvements to the efficiency of protein MC moves.[39] Protocols were similar to those used previously and are described in the Supporting Information.[40]

Acknowledgement

This research made use of the Rocket High Performance Computing service at Newcastle University. DJC acknowledges financial support from EPSRC grant EP/R010153/1.

Supporting Information Available

Additional computational methods, free energy pathways, torsion fitting and correlation analysis (PDF). Database of FEP results and small molecule and protein force field parameters (ZIP). This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Jorgensen, W. L. Efficient drug lead discovery and optimization. *Acc. Chem. Res.* **2009**, *42*, 724–733.

(2) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *J. Am. Chem. Soc.* **2015**, *137*, 2695–2703.

(3) Mobley, D. L.; Gilson, M. K. Predicting binding free energies: frontiers and benchmarks. *Ann. Rev. Biophys.* **2017**, *46*, 531–558.

(4) 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. Struc. Biol.* **2011**, *21*, 150–160.

(5) Bernardi, R. C.; Melo, M. C.; Schulten, K. Enhanced sampling techniques in molecular dynamics simulations of biological systems. *BBA-Gen. Subjects* **2015**, *1850*, 872–877.

(6) Spiwok, V.; Sucur, Z.; Hosek, P. Enhanced sampling techniques in biomolecular simulations. *Biotechnol. Adv.* **2015**, *33*, 1130–1140.

(7) Riniker, S. Fixed-charge atomistic force fields for molecular dynamics simulations in the condensed phase: An overview. *J. Chem. Inf. Model.* **2018**, *58*, 565–578.

(8) Cole, D. J.; Horton, J. T.; Nelson, L.; Kurdekar, V. The future of force fields in computer-aided drug design. *Future Med. Chem.* **2019**, *11*, 2359–2363.

(9) Yin, J.; Henriksen, N. M.; Slochower, D. R.; Shirts, M. R.; Chiu, M. W.; Mobley, D. L.; Gilson, M. K. Overview of the SAMPL5 host–guest challenge: Are we doing better? *J. Comput. Aid. Mol. Des.* **2017**, *31*, 1–19.

(10) Nerenberg, P. S.; Head-Gordon, T. New developments in force fields for biomolecular simulations. *Curr. Opin. Struc. Biol.* **2018**, *49*, 129–138.

(11) Dodda, L. S.; Vilseck, J. Z.; Tirado-Rives, J.; Jorgensen, W. L. 1.14* CM1A-LBCC: localized bond-charge corrected CM1A charges for condensed-phase simulations. *J. Phys. Chem. B* **2017**, *121*, 3864–3870.

(12) Huang, L.; Roux, B. Automated force field parameterization for nonpolarizable and polarizable atomic models based on ab initio target data. *J. Chem. Theory Comput.* **2013**, *9*, 3543–3556.

(13) Wang, L.-P.; Martinez, T. J.; Pande, V. S. Building force fields: An automatic, systematic, and reproducible approach. *J. Phys. Chem. Lett.* **2014**, *5*, 1885–1891.

(14) Betz, R. M.; Walker, R. C. Paramfit: automated optimization of force field parameters for molecular dynamics simulations. *J. Comput. Chem.* **2015**, *36*, 79–87.

(15) Prampolini, G.; Campetella, M.; De Mitri, N.; Livotto, P. R.; Cacelli, I. Systematic and automated development of quantum mechanically derived force fields: The challenging case of halogenated hydrocarbons. *J. Chem. Theory Comput.* **2016**, *12*, 5525–5540.

(16) Grimme, S. A general quantum mechanically derived force field (QMDFF) for molecules and condensed phase simulations. *J. Chem. Theory Comput.* **2014**, *10*, 4497–4514.

(17) Cacelli, I.; Cimoli, A.; Livotto, P. R.; Prampolini, G. An automated approach for the parameterization of accurate intermolecular force-fields: Pyridine as a case study. *J. Comput. Chem.* **2012**, *33*, 1055–1067.

(18) Horton, J. T.; Allen, A. E.; Dodda, L. S.; Cole, D. J. QUBEKit: Automating the Derivation of Force Field Parameters from Quantum Mechanics. *J. Chem. Inf. Model.* **2019**, *59*, 1366–1381.

(19) Cole, D. J.; Vilseck, J. Z.; Tirado-Rives, J.; Payne, M. C.; Jorgensen, W. L. Biomolecular force field parameterization via atoms-in-molecule electron density partitioning. *J. Chem. Theory Comput.* **2016**, *12*, 2312–2323.

(20) Allen, A. E.; Robertson, M. J.; Payne, M. C.; Cole, D. J. Development and Validation of the Quantum Mechanical Bespoke Protein Force Field. *ACS Omega* **2019**, *4*, 14537–14550.

(21) Manz, T. A.; Sholl, D. S. Chemically meaningful atomic charges that reproduce the electrostatic potential in periodic and nonperiodic materials. *J. Chem. Theory Comput.* **2010**, *6*, 2455–2468.

(22) Manz, T. A.; Sholl, D. S. Improved atoms-in-molecule charge partitioning functional for simultaneously reproducing the electrostatic potential and chemical states in periodic and nonperiodic materials. *J. Chem. Theory Comput.* **2012**, *8*, 2844–2867.

(23) Skylaris, C.-K.; Haynes, P. D.; Mostofi, A. A.; Payne, M. C. Introducing ONETEP: Linear-scaling density functional simulations on parallel computers. *J. Chem. Phys.* **2005**, *122*, 084119.

(24) Lee, L. P.; Cole, D. J.; Skylaris, C.-K.; Jorgensen, W. L.; Payne, M. C. Polarized protein-specific charges from atoms-in-molecule electron density partitioning. *J. Chem. Theory Comput.* **2013**, *9*, 2981–2991.

(25) Lee, L. P.; Limas, N. G.; Cole, D. J.; Payne, M. C.; Skylaris, C.-K.; Manz, T. A. Expanding the scope of density derived electrostatic and chemical charge partitioning to thousands of atoms. *J. Chem. Theory Comput.* **2014**, *10*, 5377–5390.

(26) Allen, A. E.; Payne, M. C.; Cole, D. J. Harmonic force constants for molecular mechanics force fields via Hessian matrix projection. *J. Chem. Theory Comput.* **2017**, *14*, 274–281.

(27) Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L. Performance of Protein–Ligand Force Fields for the Flavodoxin–Flavin Mononucleotide System. *J. Phys. Chem. Lett.* **2016**, *7*, 3032–3036.

(28) Christ, C. D.; Fox, T. Accuracy assessment and automation of free energy calculations for drug design. *J. Chem. Inf. Model.* **2013**, *54*, 108–120.

(29) Luccarelli, J.; Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. Effects of water placement on predictions of binding affinities for p38α MAP kinase inhibitors. *J. Chem. Theory Comput.* **2010**, *6*, 3850–3856.

(30) Liu, Z.; Barigye, S. J.; Shahamat, M.; Labute, P.; Moitessier, N. Atom Types Independent Molecular Mechanics Method for Predicting the Conformational Energy of Small Molecules. *J. Chem. Inf. Model.* **2018**, *58*, 194–205.

(31) Jorgensen, W. L.; Tirado-Rives, J. Molecular modeling of organic and biomolecular systems using BOSS and MCPRO. *J. Comput. Chem.* **2005**, *26*, 1689–1700.

(32) Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. Prediction of the water content in protein binding sites. *J. Phys. Chem. B* **2009**, *113*, 13337–13346.

(33) Fitzgerald, C. E.; Patel, S. B.; Becker, J. W.; Cameron, P. M.; Zaller, D.; Pikounis, V. B.; O'Keefe, S. J.; Scapin, G. Structural basis for p38α MAP kinase quinazolinone and pyridol-pyrimidine inhibitor specificity. *Nat. Struct. Mol. Biol* **2003**, *10*, 764.

(34) Bleiziffer, P.; Schaller, K.; Riniker, S. Machine learning of partial charges derived from high-quality quantum-mechanical calculations. *J. Chem. Inf. Model.* **2018**, *58*, 579–590.

(35) Visscher, K. M.; Geerke, D. P. Deriving Force-Field Parameters from First Principles Using a Polarizable and Higher Order Dispersion Model. *J. Chem. Theory Comput.* **2019**, *15*, 1875–1883.

(36) Manz, T. A.; Chen, T.; Cole, D. J.; Limas, N. G.; Fiszbein, B. New scaling relations to compute atom-in-material polarizabilities and dispersion coefficients: part 1. Theory and accuracy. *RSC Advances* **2019**, *9*, 19297–19324.

(37) Barreiro, G.; Kim, J. T.; Guimarães, C. R.; Bailey, C. M.; Domaoal, R. A.; Wang, L.; Anderson, K. S.; Jorgensen, W. L. From docking false-positive to active anti-HIV agent. *J. Med. Chem.* **2007**, *50*, 5324–5329.

(38) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, .; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian09 Revision E.01. Gaussian Inc. Wallingford CT 2009.

(39) 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.

(40) Cole, D. J.; Janecek, M.; Stokes, J. E.; Rossmann, M.; Faver, J. C.; McKenzie, G. J.; Venkitaraman, A. R.; Hyvönen, M.; Spring, D. R.; Huggins, D. J.; Jorgensen, W. L.

Computationally-guided optimization of small-molecule inhibitors of the Aurora A kinase–TPX2 protein–protein interaction. *Chem. Commun.* **2017**, *53*, 9372–9375.