This is an open access article published under a Creative Commons Attribution (CC-BY) License, which permits unrestricted use, distribution and reproduction in any medium, provided the author and source are cited.

JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

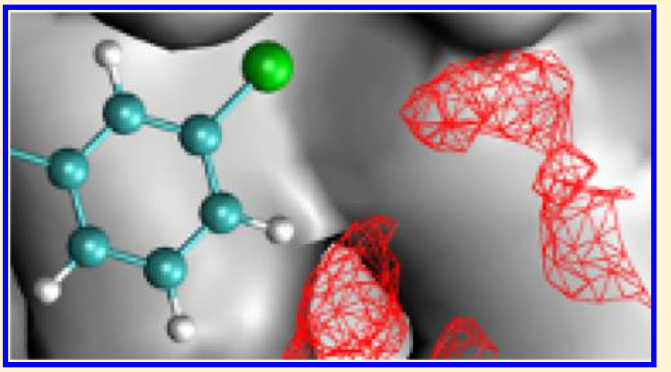
Rapid, Accurate, Precise, and Reliable Relative Free Energy Prediction Using Ensemble Based Thermodynamic Integration

Agastya P. Bhati, Shunzhou Wan, David W. Wright, and Peter V. Coveney*

Centre for Computational Science, Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, United Kingdom

Supporting Information

ABSTRACT: The accurate prediction of the binding affinities of ligands to proteins is a major goal in drug discovery and personalized medicine. The time taken to make such predictions is of similar importance to their accuracy, precision, and reliability. In the past few years, an ensemble based molecular dynamics approach has been proposed that provides a route to reliable predictions of free energies based on the molecular mechanics Poisson−Boltzmann surface area method which meets the requirements of speed, accuracy, precision, and reliability. Here, we describe an equivalent methodology based on thermodynamic integration to substantially improve the speed, accuracy, precision, and reliability of calculated relative binding free energies. We report the performance of the method when applied to a diverse set of protein targets and ligands. The results are in very good agreement with experimental data (90% of calculations agree to within 1 kcal/mol), while the method is reproducible by construction. Statistical uncertainties of the order of 0.5 kcal/mol or less are achieved. We present a systematic account of how the uncertainty in the predictions may be estimated.



1. INTRODUCTION

The free energy change associated with the binding of a lead compound or drug with a protein target, also referred to as the binding affinity, is a key property of interest in drug discovery as it correlates with the potency of possible drug candidates. The reliable prediction of binding affinities is important in drug discovery as well as in personalized medicine where it can be used to support clinical decision making. The free energy and its associated error need to be predicted within a span of a few hours in order to be useful in such real world applications.

Several *in silico* methods are available in the literature for calculating binding affinities. All these methods are based on classical molecular dynamics which is used to determine free energy and other macroscopic properties. Among them are so-called "exact" free energy methods which, through the use of a thermodynamic cycle, allow the determination of the relative binding affinity of two ligands. It is common in these methods to refer to a variable, λ, which describes the (unphysical) path taken to transform one ligand into another. Two categories of such have been developed; λ dynamics[1,2] in which λ is a dynamic variable and thermodynamic properties for multiple states are simultaneously evaluated in a single simulation and methods where separate simulations are run at fixed λ values and analyzed using formalisms such as free energy perturbation (FEP),[3] thermodynamic integration (TI),[4,5] Bennett acceptance ratio (BAR),[6] or multistate Bennett acceptance ratio (MBAR).[7] The double decoupling method[8,9] can, in principle, be used to calculate absolute binding affinities. In this method the free energy changes of decoupling the ligand from the solvent in one case and the binding site of the solvated receptor in the other are computed with their difference providing the absolute binding affinity. Another set of methods is referred to as "approximate" as they depend on underlying approximations. The empirically based linear interaction method (LIE)[10−12] and the molecular mechanics Poisson−Boltzmann surface area (MMPBSA)[13] methods fall in the latter category.

The first macromolecular free energy calculations were performed about three decades ago.[14−16] Thereafter, with the increase in speed of computers, the methods have been increasingly applied for calculating protein−ligand binding affinities within the academic community. Unfortunately, however, as we have emphasized in recent years,[17] "one-off" molecular dynamics simulations are not reproducible so most reported results lack scientific credibility. The lack of reproducibility of such an approach based on calculating free energy by performing just one MD simulation was further emphasized by Wright et al.[18] who showed that two independent MD simulations of HIV-1 protease bound to inhibitors with identical initial structure and force field parameters can produce binding affinities varying by up to 12

Received: October 5, 2016
Published: December 8, 2016

 © XXXX American Chemical Society

DOI: 10.1021/acs.jctc.6b00979
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

kcal/mol. Such variations can be much larger for more flexible ligands.[19]

The free energy is a thermodynamic property. Statistical mechanics provides the prescription for calculating such macroscopic quantities as ensemble averages of microscopic states. A theoretical discussion has been provided by Coveney and Wan.[20] Unfortunately, essentially all approaches described in textbooks and the research literature calculate these macroscopic properties from the time average of a single "long" duration trajectory. The key point is that, for systems which exhibit an equilibrium thermodynamic state, the microscopic dynamics must be at least mixing in the language of ergodic theory, hence chaotic.[20] No single trajectory can describe the behavior adequately. Various recent published works indeed demonstrate compellingly that multiple short simulations yield much more accurate binding affinities than a single long simulation.[19,21−25] A new method, namely "enhanced sampling of molecular dynamics with approximation of continuum solvent (ESMACS)", based on ensemble averages, has been shown to produce precise and reproducible absolute binding affinity predictions.[17−19,26−30]

There has been very limited application of "exact" free energy simulation methods in industrially based drug discovery. Recently, the FEP/REST method of binding affinity calculation has been proposed by Wang et al.[31] and applied to a large set of protein−ligand combinations, attracting interest within the pharmaceutical industry. Aldeghi et al.[32] described a method to calculate absolute binding affinities based on alchemical transformations, albeit with scope restricted to a small number of molecules binding to a rigid active site. An independent trajectory TI (IT-TI) method[25,33] has been reported where multiple independent trajectories were found to improve the accuracy of free energy changes calculated using TI.

The purpose of the present paper is to describe a new method called thermodynamic integration with enhanced sampling (TIES) which yields reproducible, accurate, and precise relative binding affinities. Based on the direct calculation of ensemble averages, it allows us to determine statistically meaningful results along with complete control of errors. It should be mentioned that the accuracy of the potential parametrizations used may also have substantial impact on the accuracy of results.[34] However, in the present paper, the systems studied are known to be well described by the potential parametrizations chosen.

For successful uptake in drug design and discovery, reliable predictions of binding affinities need to be made on time scales which influence experimental programmes. For applications in personalized medicine, the selection of suitable drugs needs to be made within a few hours to influence clinical decision making.[35] Therefore, speed is of the essence if we wish to use free energy based calculation methods in any of these areas. With TIES, it is possible to make predictions on time scales which meet these requirements. With sufficient computer resources available, TIES is currently able to calculate accurate, precise, and reproducible binding affinities for a set of alchemical transformations within about 8 h of wall clock time.

This paper is organized as follows. Section 2 contains the underlying theory. In section 3, the methodology, a detailed error analysis of the distribution of free energies produced by our ensemble simulations, and the required automation tools used are described. The results are reported in section 4 followed by the discussion of a few special cases in section 5; section 6 presents our conclusions.

2. THEORY

Thermodynamic integration (TI) is well-known in the literature.[4,5,36,37] The relative binding affinity of two ligands L1 and L2 is calculated by considering an alchemical transformation between them connected through intermediate states defined by introducing a coupling parameter λ, such that at $\lambda = 0$ the system corresponds to ligand L1 (initial state) and at $\lambda = 1$ the system corresponds to ligand L2 (final state). The total energy of the system is taken to be its potential energy (V). The energy of the system can be defined as

$$V(\lambda, \mathbf{x}) = (1 - \lambda)V_1(\lambda, \mathbf{x}) + \lambda V_2(\lambda, \mathbf{x}) \tag{1}$$

where V_1 and V_2 are the potential energies of ligands L1 and L2 calculated using a chosen molecular mechanics force field. The derivative of the energy with respect to λ is used to compute the free energy difference as follows

$$\Delta G_{alch} = \int_0^1 \left\langle \frac{\partial V(\lambda, \mathbf{x})}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda \tag{2}$$

where $\langle \cdots \rangle_\lambda$ denotes an ensemble average in state λ. Such an ensemble at each λ window is generated using a single MD trajectory in most cases. By contrast, TIES uses an ensemble MD simulation approach as described in detail in the following section.

The thermodynamic cycle approach is employed to calculate the relative binding affinities $\Delta\Delta G$ between these two ligands associating with a protein using the following equation

$$\Delta\Delta G = \Delta G_1 - \Delta G_2 = \Delta G_{alch}^{aq} - \Delta G_{alch}^{bound} \tag{3}$$

where ΔG_1 and ΔG_2 are the binding free energies of ligands L1 and L2, respectively. ΔG_{alch}^{aq} and ΔG_{alch}^{bound} are the free energy differences associated with the alchemical transformation of ligand L1 into L2 in free and bound states, respectively.

2.1. Ensemble Based Thermodynamic Integration. As noted, most published TI calculations have been based on performing a single MD simulation at each λ-window. The derivative of the potential energy with respect to λ, $\partial V/\partial\lambda$, lies at the heart of the binding affinity so computed by numerical quadrature (eq 2). However, to calculate a reliable free energy change the derivative at each λ-window must be accurate, precise, and reproducible. An intrinsic problem with approaches based on single MD trajectories is that they sample but one instance from a Gaussian random process and are not as such reproducible; most authors have overlooked the need to compute macroscopic averages by ensemble averaging (or they assume that their one-off MD trajectories provide the equivalent via the ergodic theorem).[20] Some authors have recognized the advantage of using ensembles over single trajectory calculations in such free energy calculations,[21−25,33,38−42] but a systematic approach has been lacking.

At the same time, the problem of unreproducible results is removed by computing ensemble averages. Ensemble averaging provides us with the means to quantify uncertainty and hence to control errors. In TIES, multiple "replica" MD simulations are performed at each λ-window, where all replicas have identical initial conditions other than their initial velocities, which are drawn randomly from a Maxwell−Boltzmann distribution. The set of multiple replica simulations constitute an ensemble simulation whose size is the number of replica simulations performed. Thus, we can compute an ensemble average of $\partial V/\partial\lambda$ values for each λ-window. Examples of the frequency distribution of this ensemble of potential energy

DOI: 10.1021/acs.jctc.6b00979
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

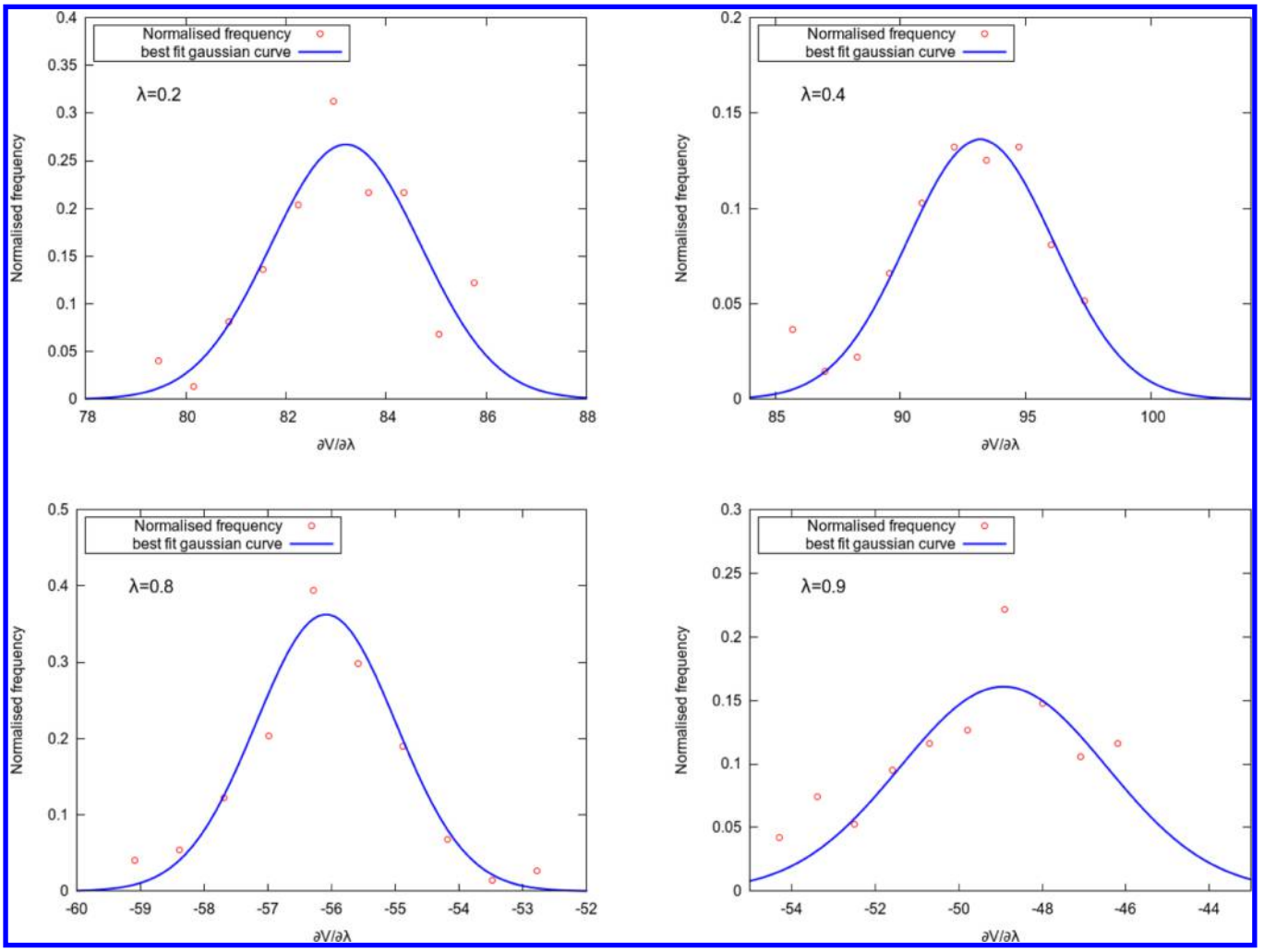


Figure 1. Normalized frequency distribution of ensembles of $\partial V/\partial\lambda$ values for 4 different intermediate alchemical states (λ = 0.2, 0.4, 0.8, 0.9) for the transformation from ligand L1Q to ligand LI9 binding to CDK2 shown as open red circles. Gaussian distributions with the same mean and standard deviation are shown in blue.

derivatives are shown in Figure 1; these are all well described by a Gaussian distribution, just as we have reported for all ensembles of MD simulations in recent publications.[17−19,26−30] This means that the integral in eq 2 should properly be interpreted as a stochastic integral. The average of all the potential energy derivatives at each λ-window is the value which is used for calculating the integral of $\partial V/\partial\lambda$ with respect to λ using a numerical method (e.g., the trapezoidal rule) to finally arrive at the alchemical free energy change ΔG_{alch} (eq 2).

3. METHOD

In this section, we describe the biomolecular systems studied and the standard TIES protocol. The ensemble size, intermediate λ values, and simulation length in this protocol have been arrived at after careful statistical analyses which are described in section 3.4. The complete end-to-end execution of our protocol is facilitated using tools described in section 3.5. As we shall see, the standard protocol is applicable to most of the ligand−protein systems studied here, although there are situations in which it needs to be modified.

3.1. Ligand−Protein Systems Studied. In this work, we have applied TIES to calculate the relative free energies of a large set of ligands bound to five different target proteins (Figures S1−S5 of the Supporting Information). They belong to different classes of protein including kinase and phosphatase with diverse functions in the human body and are important targets for a wide range of therapies. Myeloid cell leukemia 1 (MCL1) is a member of the Bcl-2 family of proteins. It is overexpressed and amplified in various cancers and promotes the aberrant survival of tumor cells that otherwise would undergo apoptosis.[43] Protein tyrosine phosphatase 1B (PTP1B) is a negative regulator of the insulin and leptin receptor pathways and thus an attractive therapeutic target for diabetes and obesity.[44] The serine protease thrombin is an established target for the prevention of cardiovascular diseases.[45] Selective targeting of tyrosine kinase 2 (TYK2) is considered as a potential treatment of inflammatory diseases, such as psoriasis and inflammatory bowel diseases (IBD).[46,47] Aberrant control of cyclin-dependent kinase 2 (CDK2) is a central feature of the molecular pathology of cancer.[48] The ligand transformations studied include chemical group modifications in which up to 10 heavy atoms are changed. In other recent work, we have applied TIES to studies of bromodomain and pan-TrkA inhibitors.[29,30]

3.2. Model Building and Simulation Setup. In the present paper, we study five different target systems (see Figure 2) using TIES. The starting structures for each of them were downloaded from the Protein Data Bank (PDB)[49] with the

DOI: 10.1021/acs.jctc.6b00979
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

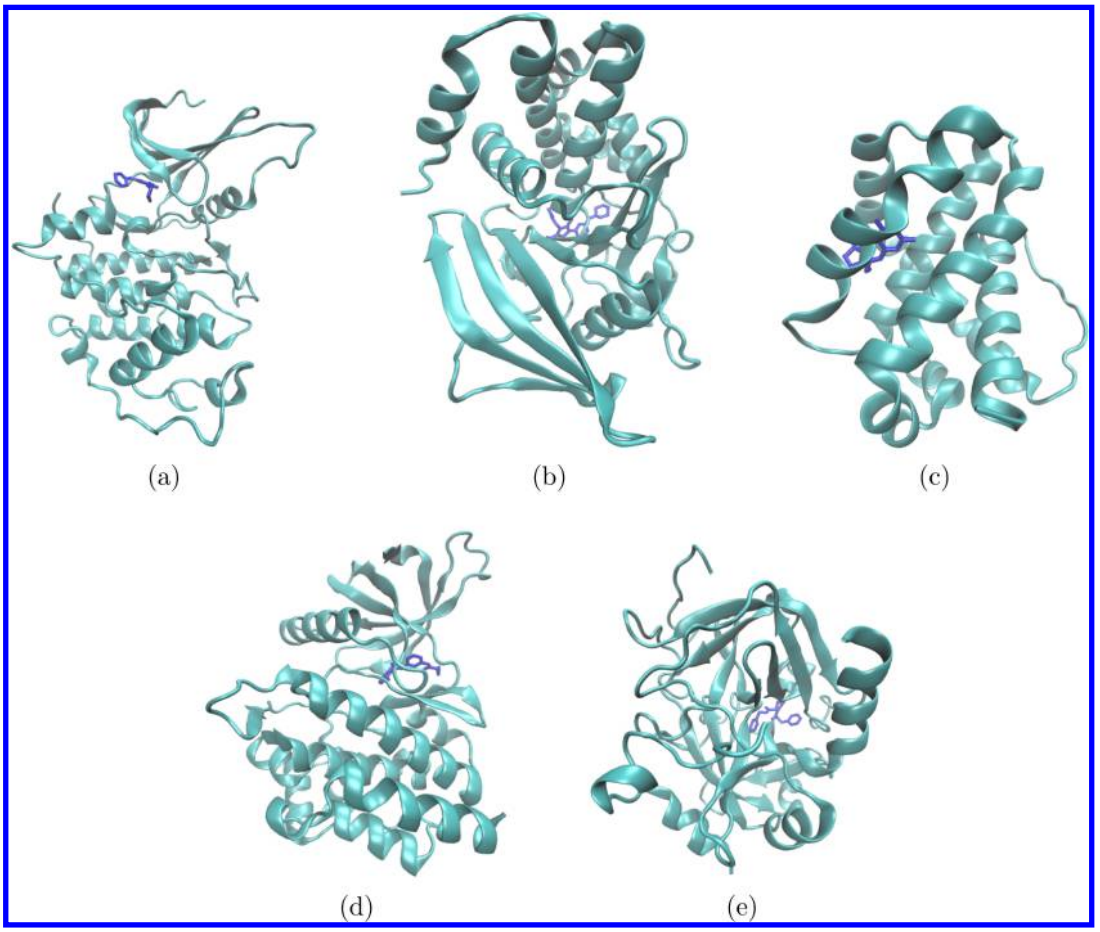


Figure 2. Structures of all five target proteins (ribbon representation) studied in each case shown bound to a ligand (blue, in stick representation): (a) cyclin-dependent kinase 2 (CDK2); PDB: 1H1Q, (b) protein tyrosine phosphatase 1B (PTP1B); PDB: 2QBS, (c) myeloid cell leukemia 1 (MCL1); PDB: 4HW3, (d) tyrosine kinase 2 (TYK2); PDB: 4GIH, and (e) thrombin; PDB: 2ZFF. Structures of all ligands (which are all drawn from congeneric series) are provided in the Supporting Information.

corresponding PDB IDs as 1H1Q for CDK2, 2ZFF for thrombin, 2QBS for PTP1B, 4HW3 for MCL1, and 4GIH for TYK2. For thrombin, 2ZC9 PDB was chosen as the starting structure to study the transformation of unsubstituted benzylamine to *meta*-substituted benzylamide where a water molecule initially present in the S1 pocket of thrombin is displaced by the appearance of a *meta*-chlorine atom (see section 5 for details). For each of these biomolecular systems, a set of around 10 ligand pairs was chosen. The structures of the ligands were downloaded from the Supporting Information of an open access article.[31]

TIES currently uses a dual topology scheme.[50] The disappearing part of the system (which exclusively belongs to the initial state) and the appearing part of the system (which exclusively belongs to the final state) are defined simultaneously. The hybrid potential energy function (eq 1) is set in such a way that the disappearing and appearing parts do not interact with each other. For an alchemical transformation from ligand L1 to ligand L2, the hybrid structure file is prepared by appending the coordinates of the appearing set of atoms from L2 to the structure file of L1 after making sure that their common parts are well aligned. The partial atomic charges for the hybrid ligand are derived from the partial atomic charges on the individual ligands such that the common atoms have identical charges, taken to be the average of their charges in the individual ligands. The charges on disappearing and appearing parts are adapted accordingly by reparametrizing the ligands after constraining the charges on the common atoms to their new values. The set of atoms to be considered as part of the common region in the hybrid ligand is selected such that their overall charges in the individual ligands do not differ by more than 0.1 *e* and, in addition, each atom satisfies this criterion. These criteria ensure that the common region of the two ligands is indeed chemically identical in both the ligands within the convergence limit of 0.1 *e*. It should be noted that one or both of these criteria may need to be relaxed in the presence of highly charged or polar groups in the ligands.

The complexes of target protein and hybrid ligand were solvated in an orthorhombic water box with buffer width of 14 Å. Counterions were added to neutralize the system electrostatically. The AMBER ff99SB-ILDN force field[51] was employed in all our simulations for protein parameters. Ligand parameters were produced using the general AMBER force field (GAFF).[52] The restrained electrostatic potential (RESP) method was used to calculate partial atomic charges using *Antechamber* (AmberTools 12) after geometry optimization by Gaussian03 for all ligands. The package NAMD 2.9[53] was used for all the molecular dynamics simulations with three-dimensional periodic boundary conditions. The systems were maintained at a temperature of 300 K and a pressure of 1 bar in an NPT ensemble using the standard NAMD protocol of Langevin dynamics (with a damping coefficient of 5 ps^{-1}) and a

DOI: 10.1021/acs.jctc.6b00979
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

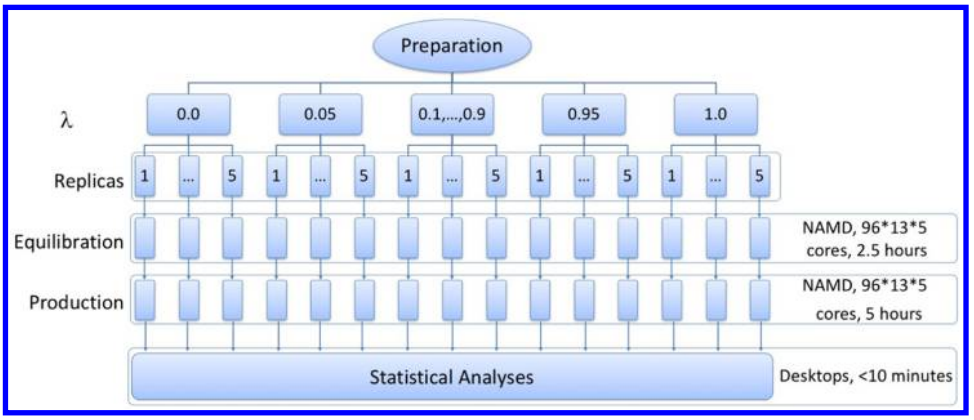


Figure 3. TIES protocol requiring 5 replica simulations at each λ window. For a single alchemical transmutation, 13 λ-windows are used amounting to 65 molecular dynamics simulations in total. The number of cores and wall clock time employed on a Cray XC30 supercomputer are displayed on the right side of the figure.

Berendsen barostat (compressibility of 4.57×10^{-5} bar^{-1} and a relaxation time of 100 fs). A time step of 2 fs was used. van der Waals contributions were perturbed using linearly varying λ across the full range (0 to 1). A soft core potential[54,55] was used for the van der Waals interactions to avoid divergent potential energy due to the sudden appearance of atoms close to the end points of the alchemical transformation, often called "end point catastrophes". Moreover, the electrostatic interactions of the disappearing atoms were linearly decoupled from the simulations between λ values of 0 and 0.55 and completely turned off beyond that, while those of the appearing atoms were linearly coupled to the simulations from λ value 0.45 to 1 and completely extinguished otherwise.

A total of 13 λ-windows were used for all TIES calculations. An ensemble simulation of size 5 was performed at each λ-window. For each replica, energy minimization followed by 2 ns long equilibration was performed using the protocol defined in a previous publication.[56] Production runs for each replica were 4 ns long. While the coordinates were recorded every 10 ps, $\partial V/\partial\lambda$ values were recorded every 2 ps. The choice of 4 ns for the simulation length and 5 for the ensemble size is based on the uncertainty quantification and error analysis discussed in section 3.4. The protocol mentioned here is a general recommendation, but there may be cases where one may wish or need to adjust it. The key point is that the ensemble size should be chosen such that the results are converged, and, hence, one may need to increase the ensemble size at different λ-windows to handle particular cases. Figure 3 exhibits the TIES workflow indicating the core counts and wall clock time required for a typical TIES calculation. Given a sufficiently powerful machine, it is possible to complete a TIES calculation using the protocol described in this work (2 ns equilibration followed by 4 ns of production simulation) within 8 h or less. For any given system the turnaround time will depend on both the number of atoms in the system of interest and the available number of cores (in many cases GPUs can provide a further increase in performance). The scalability of TIES calculations allows us to calculate relative free energies for multiple alchemical transformations concurrently. These require the same wall clock time as that for a single calculation simply by scaling up the required resources accordingly.[57]

3.3. Stochastic Integration and Error Propagation. The relative binding affinity of two ligands, $\Delta\Delta G$, is given by the difference between free energy changes associated with the alchemical transformation of one ligand into another in free and bound states, ΔG^{aq}_{alch} and ΔG^{bound}_{alch}, respectively, as per eq 3. These two terms are given by the integral of the ensemble average of $\partial V/\partial\lambda$, eq 2. As earlier discussed, the potential energy and hence its λ-derivative behave like Gaussian random variables (Figure 1), and so each point in the integrand has an underlying Gaussian distribution. The integral of a Gaussian random process is itself a Gaussian random process with variance given by the convolution of the variance of all the points used to evaluate it.[58] Therefore, we interpret the integral in eq 2 in terms of stochastic calculus. We calculate the ensemble average of the potential derivative as the average of its values from all five replicas of our ensemble simulation, where the individual value for each replica is taken to be the average potential derivative over the whole simulation length. The integral is calculated numerically using the mean potential derivative at each λ value. The error, σ_λ, for a λ window is taken to be the bootstrapped standard error of the mean of the λ-derivatives from all replicas. The variance of alchemical free energy changes ΔG^{aq}_{alch} and ΔG^{bound}_{alch} and the variance of the final relative binding affinity are calculated as

$$\sigma^2_{1/2} = \sum \sigma^2_\lambda (\Delta\lambda)^2 \tag{4a}$$

$$\sigma^2 = \sigma^2_1 + \sigma^2_2 \tag{4b}$$

The sum of the $\Delta\Delta G$ predictions for a set of ligand transformations forming a closed thermodynamic cycle is often referred to as the "hysteresis" of the cycle. Owing to the finite uncertainties associated with such $\Delta\Delta G$ predictions, the hysteresis always deviates from its theoretical value of zero and hence is one way of assessing the accuracy of the predictions.[27] It is noteworthy that on repeating all the calculations one may get a different value for the hysteresis, and hence the value of such hysteresis has an accompanying uncertainty similar to that of the $\Delta\Delta G$ predictions. Wang et al.'s recently published FEP/REST method attempts to artificially correct the $\Delta\Delta G$ predictions made by shifting their values so as to remove the associated hysteresis.[31] However, the benefit of doing this is debatable given that in their methodology there is an uncontrolled uncertainty associated with each prediction; furthermore the approach may distribute a large error arising in one prediction over the entire thermodynamic cycle, thereby distorting other correct predictions. TIES based predictions of free energy changes do not exhibit significant hysteresis as adequate sampling at

DOI: 10.1021/acs.jctc.6b00979
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

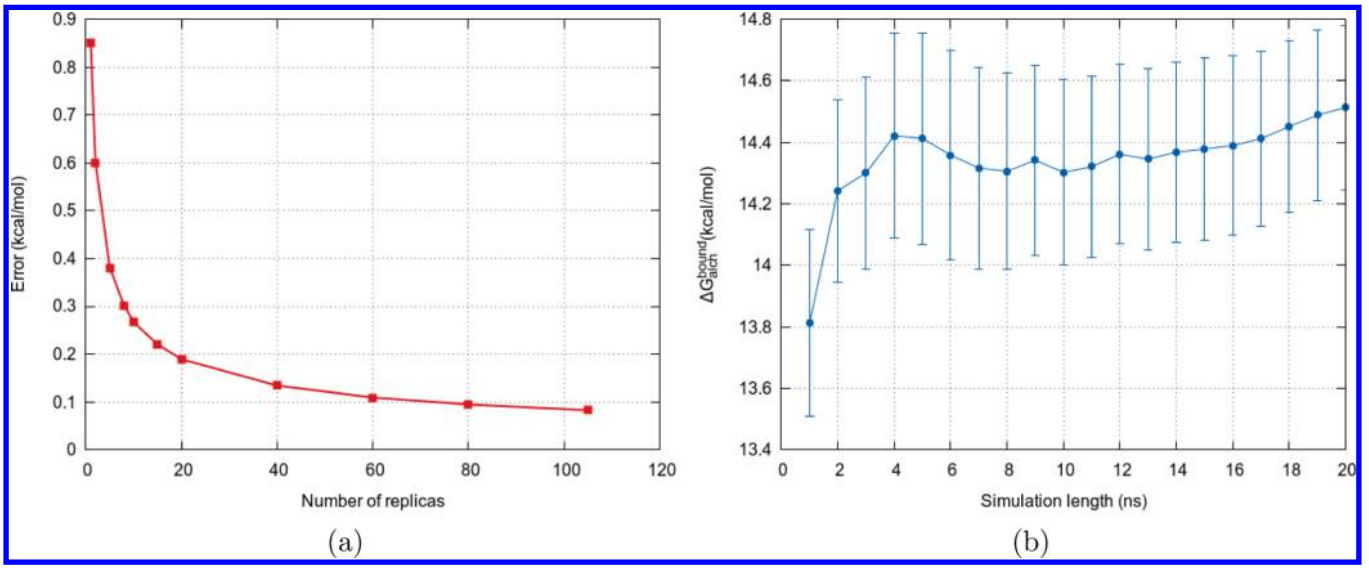


Figure 4. (a) Variation of error with the ensemble size per λ window and (b) the variation of ΔG_{alch}^{bound} with simulation length for the transformation of ligand L1Q to ligand LI9 bound to CDK2. The above plots form a basis for our choice of simulation length as 4 ns and the ensemble size as 5.

each λ-window in combination with stochastic integration minimizes it. Table S6 of the Supporting Information lists the sets of ligands forming closed cycles in the transformations studied here along with the corresponding values of hysteresis. All but one of them have hysteresis values lying within the uncertainty range. The set of ligands L32,L42,L38 binding to MCL1 as listed in Table S6 of the Supporting Information is an example where the large error in just one prediction (for the transformation L38-L42) leads to the large value of hysteresis in the closed cycle formed by them.

3.4. Uncertainty Quantification and Error Analysis.

The TIES protocol has three adjustable parameters: the ensemble size, individual replica simulation time, and the set of λ values to be selected. In this work, we have chosen the ensemble size to be 5 and the simulation length to be 4 ns. Here we provide the justification for these choices which are made on the basis of an analysis of the variation of error with both of these parameters. Figure 4a shows the variation of TIES error with the ensemble size in each λ window keeping the simulation length fixed to 4 ns at each lambda window. We performed an ensemble of 105 replicas at each λ-window for the transformation of ligand L1Q to ligand LI9 bound to CDK2. It is evident that the $\Delta\Delta G$ converges fully after the ensemble size 30 or so, where the error falls below 0.15 kcal/mol. We find that the error at the ensemble size 5 is a little below 0.4 kcal/mol, reducing to about 0.3 kcal/mol on increasing the ensemble size to 10 and to about 0.2 kcal/mol on increasing it to 20. Thus, a 4-fold increase in the computational cost yields marginal improvements. Therefore, here we take an ensemble size 5 as a good balance between computational cost and precision of results. However, this analysis clearly shows that a very high level of precision can be achieved given a sufficient amount of computation. It should be noted here that in ESMACS[26] a similar analysis was used to arrive at the appropriate ensemble size of 25, which is larger than that employed for each λ-window in the standard TIES protocol. This is because $\Delta\Delta G_{TIES}$ is evaluated from the integral in eq 2, the error associated with which is in turn calculated as per eq 4. Thus, there is more smoothing of errors as compared to that in ESMACS where there is a single ensemble computed at the end points. Figure 4b shows the variation of ΔG_{alch}^{bound} with simulation time for the fixed ensemble size (in this case 5) per λ window. It is evident that the value of ΔG_{alch}^{bound} does not vary by more than 0.1 kcal/mol after 4 ns. A similar argument based on minimizing the cost-benefit ratio on increasing the simulation length accounts for our choice of simulation length of 4 ns per replica simulation. Extensive studies of diverse ligand−protein systems all confirm that the cumulative averages of the derivatives $\partial V/\partial\lambda$ converge within 2 ns for all λ windows.[26,27,59] In the present case, we show the convergence in Figure S7 of the Supporting Information.

In this study all systems under investigation contain a relatively rigid ligand bound to a small globular protein, justifying the use of the protocol established for CDK2 in all cases. Should the system of interest differ, for example in containing a flexible ligand, it is trivial to increase replica number and/or simulation length as necessary to improve sampling and reduce statistical uncertainty.

3.5. Automation and High Performance Computing.

The TIES protocol would be very lengthy, tedious, and error-prone to perform manually. Its execution is much faster and more error-proof when performed in an automated fashion. We automated the implementation of TIES by a judicious combination of our software tools known as the Binding Affinity Calculator (BAC)[56] and FabSim[60] as we have also done for ESMACS. BAC incorporates the entire sequence of steps to be performed to produce the final results as shown in Figure 3. *Inter alia* it automatically builds the input files required for the dual topology scheme. The final step of statistical analysis of the data from the ensemble of replicas can be performed on a desktop (or remote machine) yielding the final results including error estimation and uncertainty quantification. A user-friendly version of BAC, namely *uf*-BAC, has been developed to extend its accessibility to nontechnical users.[26]

FabSim is a Python-based toolkit developed to simplify a range of computational tasks for researchers in diverse disciplines. It comes into play in the TIES workflow during the production phase. Since TIES involves performing ensemble simulations, it is highly desirable to have a well-defined scheme for data management and curation. The

DOI: 10.1021/acs.jctc.6b00979
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 1. Summary of TIES Results for the Five Different Target Proteins Studied[b]

	CDK2	thrombin	TYK2	MCL1	PTP1B
no. of transformations	7	11	11	16	10
PDB	1H1Q	2ZFF	4GIH	4HW3	2QBS
exp metrics	IC_{50}	ITC	K_i	K_i	K_i
reference	48	45	46, 47	43	44
range of uncertainty (kcal/mol)	0.2−0.3	0.2−0.4	0.1−0.2	0.2−0.5[a]	0.2−0.5
RMSE (kcal/mol)	0.8	0.8	0.5	1.4	0.5
MAE (kcal/mol)	0.7	0.7	0.4	1.2	0.4
Pearson's r	0.87	0.90	0.88	0.80	0.84
Spearman's ρ	0.86	0.91	0.88	0.80	0.82

[a]Five of the 16 transformations have uncertainties between 0.7 and 0.8 kcal/mol due to the presence of a charged carboxylate group in the transforming part of the ligands. See details in section 5. [b]The number of alchemical transformations, crystal structures used, original publications reporting the experimental binding affinities, and the experimental method used to determine the binding affinities are provided. Isothermal titration calorimetry is abbreviated as ITC, IC_{50} stands for half maximal inhibitory concentration, and K_i denotes inhibition constant. The range of uncertainty obtained is mentioned such that all the predictions have uncertainties lying within it. Values of several statistical parameters - root mean squared error (RMSE) and mean absolute error (MAE) for all TIES predictions as well as Pearson's r and Spearman's ρ between $\Delta\Delta G_{TIES}$ and experimental results - are also reported in order to assess the quality of TIES results.

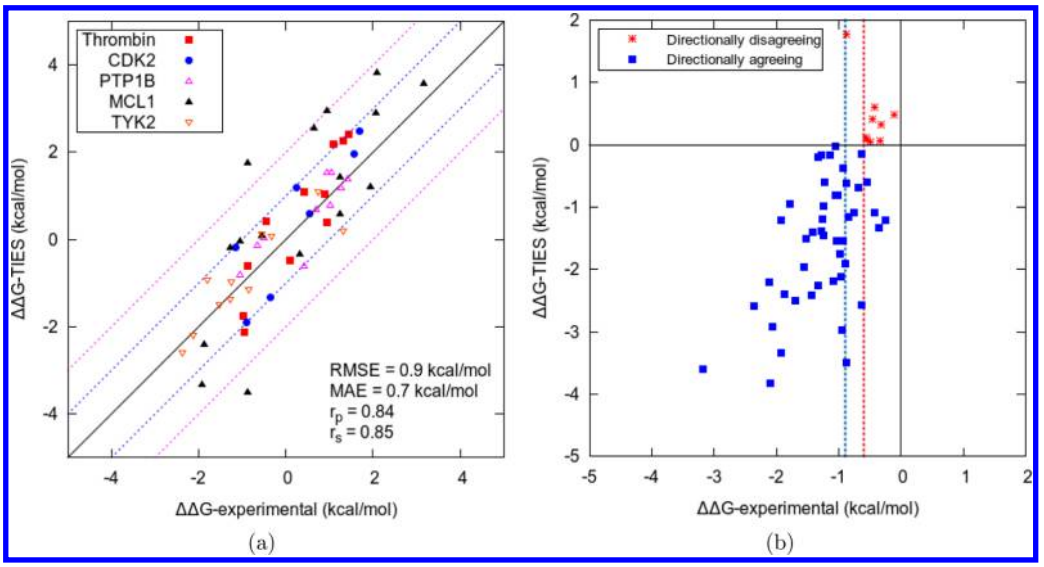


Figure 5. (a) Correlation between TIES-predicted relative binding affinities and experimental data for all five protein targets studied. The black line is the perfect correlation line. Blue and pink dotted lines show the ±1 kcal/mol and ±2 kcal/mol ranges, respectively. The majority of points lie within the ±1 kcal/mol band, a few points lie in the ±1 kcal/mol to ±2 kcal/mol band, and only two points lie outside the ±2 kcal/mol range. (b) An alternative representation of the same data such that all the experimental values are negative. Blue squares are the directionally agreeing predictions, while red stars are the directionally disagreeing ones. Red and blue dotted lines show the boundary of experimental values equal to −0.6 kcal/mol and −0.9 kcal/mol. All the predictions with $\Delta\Delta G < -0.9$ kcal/mol are in directional agreement, while all directionally disagreeing predictions lie on the right side of the red dotted line except one.

arrangement of input and output files in a standard way makes portability between different users and environments very convenient. FabSim is configured to automatically transfer data and stage the equilibration and production jobs on to the desired supercomputer using a set of simple one-line commands.[60]

4. RESULTS

In this study we applied our TIES methodology to carry out relative free energy calculations for 55 alchemical transformations between ligands binding to five different target proteins (Table 1). We were interested in evaluating the accuracy and precision of our predictions when applied to diverse sets of protein−ligand systems encapsulating different types of physicochemical interactions. Table 1 summarizes the results obtained from this study. The majority of calculations agree extremely well with the experimentally determined values (Figure 5). The overall mean absolute error (MAE) is 0.7 kcal/mol, and the root-mean-square error (RMSE) is 0.9 kcal/mol (Figure 5a). The calculated relative free energies strongly correlate with the experimental results with a Pearson's r of 0.84 and also rank the ligands well based on their relative binding affinities with a Spearman's ρ of 0.85 as shown in Figure 5a. Our calculations have a high level of precision with a range of uncertainty from 0.1 to 0.5 kcal/mol (Table 1).

Table 1 provides the evaluation of the performance of the TIES methodology for individual target proteins. For each of the target proteins, TIES predictions have the same level of accuracy and precision as for the overall results. The RMSE and MAE are below 0.8 and 0.7 kcal/mol, respectively, for all targets except MCL1, for which the respective values are 1.4 and 1.2 kcal/mol. The Pearson coefficient is no less than 0.80 for any of the targets and reaches 0.90 for thrombin. Similarly, the Spearman coefficient is always higher than 0.80, reaching 0.91

DOI: 10.1021/acs.jctc.6b00979
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

for thrombin. The range of uncertainties for each target protein, as mentioned in Table 1, shows the high level of precision achieved with our methodology for the individual biomolecular systems studied. It should be noted that 5 out of 16 transformations of MCL1 ligands have uncertainties in the range of 0.7−0.8 kcal/mol. This is due to the presence of a charged carboxylate group in the mutating part of the ligands. More details are provided in section 5.

Figure 5a is a scatter plot of predicted versus experimental relative binding affinities for all the transformations. The blue and pink dotted lines show the ±1 kcal/mol and ±2 kcal/mol bands, respectively, from the perfect correlation black line. All but two points lie within the pink lines, while a few points lie between the blue and pink lines. The majority of points is within the blue lines, serving to emphasize the high level of accuracy of our predictions. 45 out of 55 predictions (81.82%) deviate from the experimental values by less than 1.0 kcal/mol, while 53 out of 55 predictions (96.36%) do so by less than 2 kcal/mol (Table 2 (top)). More than half of our predictions have an error less than 0.6 kcal/mol. It is important to mention here that experimental binding affinities contain errors and are bound by significant uncertainties.[61,62] However, for the systems studied here the error bars on experimental results are unavailable. Figure S6 in the Supporting Information contains correlation plots for each biomolecular system separately along with the uncertainties in the TIES predictions shown as error bars.

Table 2. Summary of the Level of Accuracy Obtained for the Total Set of TIES Predictions[a]

MAE <	no. of predictions
0.4	19
0.6	30
0.8	36
1.0	45
1.2	49
1.4	50
$\lvert\Delta\Delta G_{exp}\rvert$ >	no. of directionally disagreeing predictions
0.0	9
0.3	8
0.4	6
0.5	3
0.6	1
0.9	0

[a]The number of predictions found to be accurate for a specified absolute error range (top) and the number of predictions found to be in directional agreement with the increasing absolute values of experimental results (bottom). The lack of experimental errors means that the entire discrepancy is assigned to the theoretical predictions. In practice, the agreement is likely to be even better than shown here.

Wang et al.[31] recently reported binding affinity predictions for the same protein targets as here using the FEP/REST method, and hence a direct comparison between their results and ours can be made. The RMSE and MAE for each protein target reported here (Table 1) are smaller than those given by FEP/REST, except for MCL1, for which they are the same in both studies. The correlation coefficients for each system reported here (Table 1) when compared to those given by the FEP/REST method are better for CDK2 (0.87 versus 0.48) and thrombin (0.90 versus 0.71), while being almost the same for the others. 81.82% and 96.36% of our predictions are accurate within the absolute error of 1.0 and 2.0 kcal/mol, respectively, while the same numbers from the FEP/REST study are 63.3% and 92.4%, respectively. A visual comparison of Figure 5a with Figure 3 of Wang et al.[31] shows that the number of points in the ±1 kcal/mol to ±2 kcal/mol band is much lower in the former as compared to the latter. Out of the 55 ligand transformations studied here, 18 are in common with the perturbations studied by Wang et al. using their FEP/REST methodology.[31] Figure S10 in the Supporting Information shows a direct comparison of these 18 predictions from both the methods. TIES exhibits marginally better accuracy with slightly smaller RMSE and MAE and slightly larger Pearson's r and Spearman's ρ. Only one of the 18 TIES predictions lies outside the 1 kcal/mol window from the experimental data, and one directionally disagrees with the experimental value. On the other hand, the corresponding numbers for the FEP/REST predictions are three and two, respectively. Furthermore, these authors' results are effectively obtained from running a single replica, meaning that they are liable to suffer from a lack of reproducibility.

Figure 5b provides a different representation of our results. A prediction is said to be directionally agreeing with experimental observations if $\Delta\Delta G_{TIES}$ has the same sign as that of $\Delta\Delta G_{exp}$, otherwise it is deemed to be directionally disagreeing. In Figure 5b we have flipped the signs of both $\Delta\Delta G_{exp}$ and $\Delta\Delta G_{TIES}$ for all the points which originally had positive $\Delta\Delta G_{exp}$. In other words, we rearranged eq 3 such that ligand L1 always has a more negative binding free energy than ligand L2. Such a representation is useful in order to show if the calculations exhibit the same trend as that of the experimental results. Thus, we are left with all the points on the left-hand side of the y-axis. Therefore, all the points which lie above the x-axis directionally disagree (red stars in Figure 5b), while those below the x-axis directionally agree (blue squares in Figure 5b). The red and blue dotted lines here are the boundaries of $\Delta\Delta G_{exp}$ equal to −0.6 kcal/mol and −0.9 kcal/mol, respectively. As shown in the figure, all the points in directional disagreement lie to the right of the red line except one, which lies between the red and the blue lines. This means that all our results predict the direction of the change in the binding affinities correctly if the absolute change in the corresponding experimental values is greater than or equal to 0.9 kcal/mol, with the proviso that there are no error bars on the experimental values that are available in this study, as mentioned earlier. Wang et al. suggest that for high quality measurements, the uncertainties on the experimental relative binding affinities $\Delta\Delta G_{exp}$ are of the order of 0.4−0.7 kcal/mol.[31] More generally, Chodera and Mobley state that published affinities may have errors of around 24%, reported errors near universally underestimating the error obtained from interlaboratory variation by one to 2 orders of magnitude.[62] All but one TIES prediction which directionally disagree lie within this range of experimental noise as shown in Table 2 (bottom). Therefore, when the overlap between the error bars of both the experimental and the theoretical $\Delta\Delta G$ predictions is taken into account, even better agreement can be achieved than that reported here.

4.1. Reproducibility of the Predictions. Another strength of TIES is its ability to yield reproducible predictions. By reproducibility, we mean that if a TIES calculation is repeated for a particular transformation, the new prediction of the relative binding affinity for that transformation would lie within $\pm\sigma$ and $\pm2\sigma$ range of the original prediction with a probability of 0.68 and 0.95, respectively. The typical

DOI: 10.1021/acs.jctc.6b00979
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

uncertainty on our predictions is 0.4 kcal/mol. Thus, in principle, any repeated TIES calculation for a typical transformation should yield a new prediction within ±0.8 kcal/mol of the original prediction 95% of the time.

To illustrate this property of our methodology, we chose six of our relative free energy calculations between ligands binding to the CDK2 target protein. An ensemble of 20 MD simulations at each λ-window was performed, and the final relative free energies were calculated (represented as open black circles along with their $\pm\sigma$ and $\pm 2\sigma$ ranges in Figure 6).

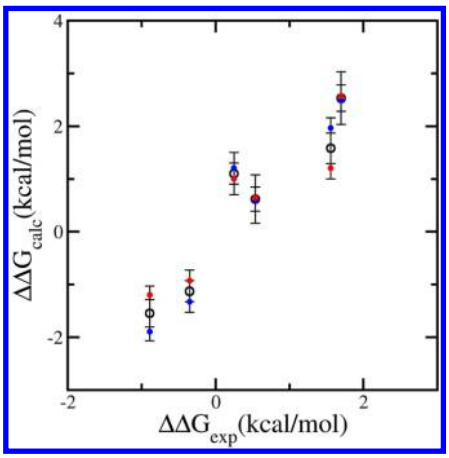


Figure 6. Reproducibility of TIES: The relative binding affinities of the CDK2 ligands (black circles) were calculated using 10 values resampled from an ensemble of 20 replica simulations. Error bars are represented as standard deviations of σ and 2σ. For each calculated $\Delta\Delta G$ value, results are also shown for two randomly chosen nonoverlapping 10-replica samples (blue and red dots). The data demonstrate that a 10-replica prediction will lie within $\pm\sigma$ and $\pm 2\sigma$ of the averaged relative binding affinities (open black circles) with confidence intervals of 0.68 and 0.95, respectively.

Thereafter, two random nonoverlapping subsamples of ten MD simulations at each λ-window were chosen in order to obtain two independent TIES predictions (denoted by red and blue filled circles in Figure 6) for the 6 transformations. Both the predictions (red and blue circles) for four of the transformations lie within the $\pm\sigma$ range, while for the remaining two they lie within the $\pm 2\sigma$ range.

4.2. Comparison of Results from Ensemble Based TIES with Those Using Extended Simulations. We compared the predictions from the use of the "standard" TIES protocol with the case when only one "long" simulation is performed at each λ-window. For this purpose, we chose the transformation from ligands L1Q to LI9 binding to CDK2. Five independent TI calculations were performed for this transformation based on a 20 ns molecular dynamics simulation at each λ window to check its reproducibility. The ΔG_{alch}^{bound} from these five TI calculations vary by as much as 0.8 kcal/mol. On the other hand, the ΔG_{alch}^{bound} of the same transformation from TIES has the 1σ uncertainty range as 0.38 kcal/mol. Therefore, at least in this case, the results from TIES employing the standard protocol are only slightly better than those from TI employing a single longer simulation at each λ-window for the same amount of computer time (equivalent to 20 ns per lambda window in each case). This observation may be compared with other studies,[17−19,26−28] where it has been found that the prediction from an ensemble of 50 short (4 ns) simulations is better than using a single 1 μs simulation. This may be related to our earlier observation (section 3.4) that the TIES/TI algorithm involves a number of internal cancellation of "errors", unlike in end point MD. However, for cases with larger and more flexible ligands and/or proteins and when the size of the alchemically changing part of the ligand is large, TIES yields much more precise predictions than can be realized from single "long" simulations.[27] Moreover, the use of multiple replicas within TIES has the important advantage of reducing the wall clock time required to get results since these can all be performed concurrently on a supercomputer.

5. DISCUSSION

Mobley and Klimovitch[63] quantified the impact of reliable binding free energy predictions on the drug discovery process. According to them, during lead optimization, a computational method screening 10−100 molecules per week with an error of 0.5 kcal/mol would reduce by a factor of 8 the number of compounds requiring synthesis in order to achieve a 10-fold improvement in binding free energy. A similar conclusion can be drawn from the analysis of biological assay variability.[61] In our study, TIES has been shown to achieve this level of precision with a short turnaround time and hence has considerable potential to influence drug design. Moreover, in this study we have employed TIES on a broad range of target proteins and for a large set of alchemical transformations including many types of chemical interactions and processes such as hydrophobic, hydrogen-bonding, and electrostatic interactions and solvent effects including displacement of water molecules from binding pockets. All these interactions have been accurately predicted with TIES within the limits of the accuracy of the classical force-field and ligand parameters employed. Below we draw attention to several cases in which specific interactions impact results.

Figure 7 shows the S1 pocket of the active site of the human thrombin protein which contains a water molecule in the case of the ligand containing an unsubstituted benzylamine ring. This water molecule is locked inside the pocket and mediates hydrogen bonding from the ligand amidino group to the protein. However, the water molecule is displaced by the *meta*-substituted benzylamide ring of the ligand. We studied the transformation from *meta*-substituted benzylamide to benzylamine (A) and the reverse transformation (B) using TIES; our results are in excellent agreement with the experimental values. $\Delta\Delta G_{TIES}$ for the transformations A and B are −0.6 ± 0.2 kcal/mol and 1.0 ± 0.2 kcal/mol, respectively, while the corresponding experimental values are −0.9 and 0.9 kcal/mol, respectively (no error bars reported). It is noteworthy here that the $\Delta\Delta G$ prediction of the transformation B from FEP/REST[31] is 1.5 kcal/mol. In order to quantify the contribution arising from the presence of the water molecule we calculated the probability of occurrence of a water molecule in a cubic box of volume 1000 Å^3 centered around the pocket over all conformations taken from the 5 replica simulations aligned to the initial structure. In the case of transformation A, the probability increases from 0 at $\lambda = 0$ to 0.7 at $\lambda = 1$, while in the case of transformation B, it decreases from 0.6 at $\lambda = 0$ to 0.3 at $\lambda = 1$. Figure 7 shows the two end points ($\lambda = 0,1$) of transformation A where the red frames denote the space with the number density of water oxygen centers greater than or equal to that in bulk water. Figure 7a shows that the presence of water is confined to the two distinct channels C1 and C2 in the presence of *m*-Cl, while Figure 7b shows that when the chlorine

DOI: 10.1021/acs.jctc.6b00979
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

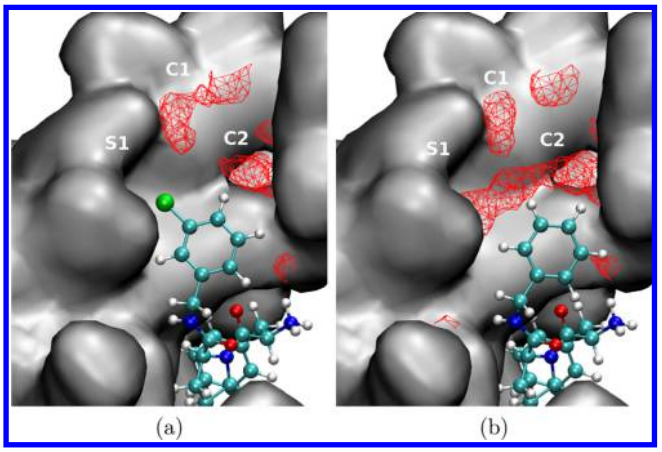


Figure 7. Cross sections of the S1 pocket of thrombin for the two end λ-windows of an alchemical transformation involving mutation of *m*-chlorobenzylamide to benzamidine. Experimentally, *meta*-substituents of benzylamides displace the water molecule from the S1 pocket which is present there in the case of an unsubstituted benzylamine ring. Red wireframes show the regions of water occupancy averaged over all the conformations from the 5 replica simulations aligned to the corresponding initial structures. (a) No water molecule bound in the S1 pocket in the presence of chlorine (green) at $\lambda = 0$. (b) The water molecule enters the S1 pocket through channels C1 and C2 on fully transforming Cl to H at $\lambda = 1$. The protein surface is shown in gray, and ligand atoms are colored by element: hydrogen in white, carbon in cyan, oxygen in red, and nitrogen in blue.

atom is transformed fully to a hydrogen atom the water enters into the S1 pocket as well as through C1 and C2.

In the intermediate λ-windows of the TIES calculation the electrostatic interactions are weakly scaled. Therefore, when the alchemically changing groups contain charges, it is sometimes possible to encounter larger fluctuations in $\partial V/\partial\lambda$ due to instabilities caused by the loss or increase of strong electrostatic interactions. This results in larger variation in the stochastic integral (eq 2), and hence, less precise prediction of $\Delta\Delta G$ using the standard TIES protocol as mentioned in section 3.2. In this study, 5 out of the 16 transformations of MCL1 ligands have a charged carboxylate group in the perturbing region, with the uncertainties of their $\Delta\Delta G$ predictions lying in the range of 0.7−0.8 kcal/mol, while the uncertainties of the remaining 11 predictions for MCL1 ligands lie within the range of 0.2−0.4 kcal/mol (see Table 1). In such cases, one can modify the standard protocol described here by, for example, increasing the ensemble size at various λ-windows and/or excluding the charged group from the alchemically mutating part of the ligands to further emphasize this point. A comparison of results with and without charged groups in PTP1B ligands is examined in detail in the SI.

It is evident from Table 1 and Figure 5 that the TIES predictions for MCL1 have larger deviations from the experimental results, with the largest RMSE and MAE. In this case interactions with R263 become too weak to hold it in a stable position at intermediate λ-windows due to the scaling down of the electrostatic interactions. Such behavior can be attributed to the highly flexible nature of the ligand (detail shown in the SI).

The free energy methods based on alchemical transformation found in the literature have been applied to cases where the size of the change in the chemical group is small, usually limited to one or two heavy atoms. However, the recently reported FEP/REST method handles perturbations up to 10 heavy atoms.[31] In this study, we have achieved the same level by successfully applying TIES to a diverse range of chemical modifications with "perturbations" including up to 10 heavy atoms. For example, the range of chemical transformations with corresponding absolute errors (AE) include benzamide to *p*-fluorobenzenesulfonamide (AE = 0.6 kcal/mol), benzyloxy to ethanamide (AE = 0.5 kcal/mol), hydrogen to methyl cyclohexane (AE = 0.1 kcal/mol), cyclohexyl to phenyl (AE = 0.5 kcal/mol), cyclohexyl methyl to 2,2,4,4-tetramethylcyclohexyl (AE = 0.2 kcal/mol), cyclopentyl to cycloheptyl (AE = 0.0 kcal/mol), and indole to indane (AE = 2.1 kcal/mol) to mention a few. TIES correctly predicts that small changes in binding affinity are associated with these large perturbations (up to 10 heavy atoms).[30]

6. CONCLUSIONS

In this article, a new approach based on thermodynamic integration is described to predict relative binding free energies which has a high level of accuracy and precision. The methodology described here is robust and reliable. Automation of the entire workflow yields results within a few hours. It provides users with the freedom to set desired statistical uncertainties for purposes of accuracy and precision, based on both the length of the MD simulations and the number of replicas per λ-window, which can all be adjusted separately. The method gives predictions in very good agreement with the experimental values for a large set of ligands bound to five different target proteins. Out of the total 55 predictions made here, 30 deviate from experimental values by less than 0.6 kcal/mol, with an overall mean absolute error of 0.7 kcal/mol and root-mean-square error of 0.9 kcal/mol. A rich diversity of chemical modifications has been successfully captured with this approach, ranging from single atom perturbation to large functional group modifications (up to 10 heavy atoms). Accurate predictions for flexible ligands with large numbers of rotatable bonds were made. Most importantly, the calculated relative binding affinities are shown to be reproducible within carefully controlled statistical uncertainties. Such a high level of accuracy and precision make the predictions reliable, and, hence, this approach may prove of substantial value in the drug discovery and design process.

A direct comparison of our results with those from FEP/REST[31] demonstrates enhanced accuracy of predictions for the five biomolecular systems studied in terms of reduced mean absolute error and root mean squared error and improved correlation coefficients (section 4). Unlike TIES where it is a built-in feature, there is no mention of the reproducibility of the FEP/REST calculations. Compounding the issue, the FEP/REST methodology is proprietary and not accessible to open evaluation.

In the present study, the potential parametrizations for both proteins and ligands are all known to be reliable. However, this cannot be guaranteed to be the case in new situations. Care must always be taken in selecting such parametrizations, particularly for new ligands. Problems pertaining to alchemical transformations between congeners involving a change in the net charge using free energy methods are well-known, and TIES is no exception to them.[64,65] The current version of the TIES protocol cannot reliably handle such situations. Moreover, when alchemically mutating parts of the ligands contain charges it may be necessary to modify the standard TIES

DOI: 10.1021/acs.jctc.6b00979
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

protocol to ensure adequate sampling in the calculation of ensemble averages.

TIES represents one of a new class of ensemble-based free energy methods which is rapid, accurate, precise, and reproducible. The approach has the potential to make an impact in drug design and personalized medicine.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.6b00979.

Tables of all the experimental and predicted $\Delta\Delta G$ values for all five protein targets bound to pairs of ligands, along with the structures of all the ligands studied and discussion of some notable ligand−protein interactions studied with TIES; hysteresis values for selected ligands forming closed cycles and plots exhibiting the convergence of $\langle \partial V/\partial \lambda \rangle$ (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +44 (0)20 7679 4560. E-mail: p.v.coveney@ucl.ac.uk.

ORCID

Agastya P. Bhati: 0000-0003-4539-4819
Shunzhou Wan: 0000-0001-7192-1999
David W. Wright: 0000-0002-5124-8044
Peter V. Coveney: 0000-0002-8787-7256

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Frank Lovering of Pfizer for helpful discussions pertaining to the comparison of experimental and theoretical free energy predictions. The authors would like to acknowledge the support of EPSRC via the 2020 Science programme (http://www.2020science.net/, EP/I017909/1), the Qatar National Research Fund (7-1083-1-191), the MRC Medical Bioinformatics project (MR/L016311/1), the EU H2020 projects ComPat (http://www.compat-project.eu/, Grant No. 671564), and CompBioMed (http://www.compbiomed.eu/, Grant No. 675451) and funding from the UCL Provost. We acknowledge the Leibniz Supercomputing Centre for providing access to SuperMUC (https://www.lrz.de/services/compute/) and the very able assistance of its scientific support staff. We also made use of ARCHER, the UK's national High Performance Computing Service, funded by the Office of Science and Technology through EPSRC's High-End Computing Programme. Access to ARCHER was provided through the 2020 Science programme. This research was supported in part by PLGrid Infrastructure (specifically the Prometheus supercomputer run by ACK Cyfronet AGH in Krakow). A.P.B. is supported by an Overseas Research Scholarship from UCL and an Inlaks Scholarship from the Inlaks Shivdasani Foundation.

■ REFERENCES

(1) Kong, X.; Brooks, C. L. λ-dynamics: A new approach to free energy calculations. *J. Chem. Phys.* **1996**, *105*, 2414.
(2) Knight, J. L.; Brooks, C. L. λ-dynamics free energy simulation methods. *J. Comput. Chem.* **2009**, *30*, 1692.
(3) Zwanzing, R. W. High-temperature equation of state by a perturbation method in non polar gases. *J. Chem. Phys.* **1954**, *22*, 1420.
(4) Straatsma, T. P.; Berendsen, H. J. C.; Postma, J. P. M. Free energy of hydrophobic hydration: A molecular dynamics study of nobel gases in water. *J. Chem. Phys.* **1986**, *85*, 6720.
(5) Straatsma, T. P.; Berendsen, H. J. C. Free energy of ionic hydration: Analysis of a thermodynamic integration technique to evaluate free energy differences by molecular dynamics simulation. *J. Chem. Phys.* **1988**, *89*, 5876.
(6) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245.
(7) Shirts, M. R.; Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* **2008**, *129*, 124105.
(8) Hamelberg, D.; McCammon, J. A. Standard free energy of releasing a localized water molecule from the binding pockets of proteins: Double-decoupling method. *J. Am. Chem. Soc.* **2004**, *126*, 7683.
(9) Gilson, M.; Given, J.; Bush, B.; McCammon, J. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* **1997**, *72*, 1047.
(10) Hansson, T.; Åqvist, J. Estimation of binding free energies for HIV proprotein inhibitors by molecular dynamics simulations. *Protein Eng., Des. Sel.* **1995**, *8*, 1137.
(11) Åqvist, J.; Medina, C.; Samuelsson, J. E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng., Des. Sel.* **1994**, *7*, 385.
(12) Åqvist, J. Calculation of absolute binding free energies for charged ligands and eeffect of long-range electrostatic interactions. *J. Comput. Chem.* **1996**, *17*, 1587.
(13) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889. PMID: 11123888.
(14) Postma, J. P. M.; Berendsen, H. J. C.; Haak, J. R. Thermodynamics of cavity formation in water. *Faraday Symp. Chem. Soc.* **1982**, *17*, 55.
(15) Tembre, B. L.; McCammon, J. A. Ligand-receptor interactions. *Comput. Chem.* **1984**, *8*, 281.
(16) Jorgensen, W. L.; Ravimohan, C. Monte Carlo simulation of differences in free energies of hydration. *J. Chem. Phys.* **1985**, *83*, 3050.
(17) Sadiq, S. K.; Wright, D. W.; Kenway, O. A.; Coveney, P. V. Accurate ensemble molecular dynamics binding free energy ranking of multidrug-resistance HIV-1 protease. *J. Chem. Inf. Model.* **2010**, *50*, 890.
(18) Wright, D. W.; Hall, B. A.; Kenway, O. A.; Jha, S.; Coveney, P. V. Computing clinically relevant binding free energies of HIV-1 protease inhibitors. *J. Chem. Theory Comput.* **2014**, *10*, 1228.
(19) Wan, S.; Knapp, B.; Wright, D. W.; Deane, C. M.; Coveney, P. V. Rapid, precise and reproducible prediction of peptide-MHC binding affinities from molecular dynamics that correlate well with experiment. *J. Chem. Theory Comput.* **2015**, *11*, 3346.
(20) Coveney, P. V.; Wan, S. On the calculation of equilibrium thermodynamic properties from molecular dynamics. *Phys. Chem. Chem. Phys.* **2016**, *18*, 30236.
(21) Genheden, S.; Ryde, U. How to obtain statistically converged MM/GBSA results. *J. Comput. Chem.* **2010**, *31*, 837.
(22) Auffinger, P.; Louise-May, S.; Westhof, E. Multiple molecular dynamics simulations of the anticodon loop of tRNAAsp in aqueous solution with counterions. *J. Am. Chem. Soc.* **1995**, *117*, 6720.
(23) Zagrovic, B.; van Gunsteren, W. F. Computational analysis of the mechanism and thermodynamics of inhibition of phosphodiesterase 5A by synthetic ligands. *J. Chem. Theory Comput.* **2007**, *3*, 301. PMID: 26627173.
(24) Genheden, S.; Ryde, U. A comparison of different initialization protocols to obtain statistically independent molecular dynamics simulations. *J. Comput. Chem.* **2011**, *32*, 187.
(25) Lawrenz, M.; Baron, R.; McCammon, J. A. Independent-trajectories thermodynamic-integration free-energy changes for

DOI: 10.1021/acs.jctc.6b00979
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

biomolecular systems: Determinants of H5N1 avian influenza virus neuraminidase inhibition by peramivir. *J. Chem. Theory Comput.* **2009**, *5*, 1106. PMID: 19461872.
(26) Wan, S.; Bhati, A. P.; Zasada, S.; Coveney, P. V. *Rapid, accurate, precise and reproducible ligand-protein binding free energy prediction* **2016**, preprint.
(27) Bunney, T. D.; Wan, S.; Thiyagarajan, N.; Sutto, L.; Williams, S. V.; Ashford, P.; Koss, H.; Knowles, M. A.; Gervasio, M. A.; Coveney, P. V.; Katan, M. The effect of mutations on drug sensitivity and kinase activity of fibroblast growth factor receptors: A combined experimental and theoretical study. *EBioMedicine* **2015**, *2*, 194.
(28) Wan, S.; Coveney, P. V. Rapid and accurate ranking of binding affinities of epidermal growth factor receptor sequences with selected lung cancer drugs. *J. R. Soc., Interface* **2011**, *8*, 1114.
(29) Wan, S.; Bhati, A. P.; Zasada, S. J.; Wall, I.; Green, D.; Bamborough, P.; Coveney, P. V. *Rapid and reliable binding affinity prediction for analysis of bromodomain inhibitors: a computational study* **2016**, preprint.
(30) Wan, S.; Bhati, A.; Coveney, P.; Skerratt, S. E.; Gore, K.; Bagal, S. K.; Shanmugasundaram, V.; Omoto, K. Rapid, accurate and reproducible binding affinity calculation for drug discovery: A retrospective analysis of the Pfizer Pan-Trk program. 2016; 252nd ACS National Meeting. Philadelphia, USA.
(31) 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.
(32) Aldeghi, M.; Heifetz, A.; Bodkin, M. J.; Knapp, S.; Biggin, P. C. Accurate calculation of the absolute free energy of binding for drug molecules. *Chem. Sci.* **2016**, *7*, 207.
(33) Lawrenz, M.; Baron, R.; Wang, Y.; McCammon, J. A. In *Computational Drug Discovery and Design*; Baron, R., Ed.; Springer: New York, 2012; Vol. *819*, Chapter Independent-Trajectory Thermodynamic Integration: A Practical Guide to Protein-Drug Binding Free Energy Calculations Using Distributed Computing, pp 469−486.
(34) Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. Systematic validation of protein force fields against experimental data. *PLoS One* **2012**, *7*, e32131.
(35) Sadiq, S. K.; Mazzeo, M. D.; Zasada, S.; Manos, S.; Stoica, I.; Gale, C. V.; Watson, S. J.; Kellam, P.; Brew, S.; Coveney, P. V. Patient-specific simulation as a basis for clinical decision-making. *Philos. Trans. R. Soc., A* **2008**, *366*, 3199.
(36) Straatsma, T. P.; McCammon, J. A. Multiconfigurational thermodynamic integration. *J. Chem. Phys.* **1991**, *95*, 1175.
(37) Straatsma, T. P.; McCammon, J. A. Computational alchemy. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407.
(38) Caves, L. S. D.; Evanseck, J. D.; Karplus, M. Locally accessible conformations of proteins: Multiple molecular dynamics simulations of crambin. *Protein Sci.* **1998**, *7*, 649.
(39) Elofsson, A.; Nilsson, L. How consistent are molecular dynamics simulations? *J. Mol. Biol.* **1993**, *233*, 766.
(40) Genheden, S.; Ryde, U. Will molecular dynamics simulations of proteins ever reach equilibrium? *Phys. Chem. Chem. Phys.* **2012**, *14*, 8662.
(41) Genheden, S.; Mikulskis, P.; Hu, L.; Kongsted, J.; Söderhjelm, P.; Ryde, U. Accurate prediction of nonpolar solvation free energies require explicit consideration of binding-site hydration. *J. Am. Chem. Soc.* **2011**, *133*, 13081.
(42) Fujitani, H.; Tanida, Y.; Ito, M.; Jayachandran, G.; Snow, C. D.; Shirts, M. R.; Sorin, E. J.; Pande, V. S. Direct calculation of the binding free energies of FKBP ligands. *J. Chem. Phys.* **2005**, *123*, 084108.
(43) Friberg, A.; Vigil, D.; Zhao, B.; Daniels, R. N.; Burke, J. P.; Garcia-Barrantes, P. M.; Camper, D.; Chauder, B. A.; Lee, T.; Olejniczak, E. T.; Fesik, S. W. Discovery of potent myeloid cell leukemia 1 (Mcl-1) inhibitors using fragment-based methods and structure-based design. *J. Med. Chem.* **2013**, *56*, 15. PMID: 23244564.
(44) Wilson, D. P.; Wan, Z.-K.; Xu, W.-X.; Kirincich, S. J.; Follows, B. C.; Joseph-McCarthy, D.; Foreman, K.; Moretto, A.; Wu, J.; Zhu, M.; Binnun, E.; Zhang, Y.-L.; Tam, M.; Erbe, D. V.; Tobin, J.; Xu, X.; Leung, L.; Shilling, A.; Tam, S. Y.; Mansour, T. S.; Lee, J. Structure-based optimization of protein tyrosine phosphatase 1B inhibitors: From the active site to the second phosphotyrosine binding site. *J. Med. Chem.* **2007**, *50*, 4681. PMID: 17705360.
(45) Baum, B.; Mohamed, M.; Zayed, M.; Gerlach, C.; Heine, A.; Hangauer, D.; Klebe, G. More than a simple lipophilic contact: A detailed thermodynamic analysis of nonbasic residues in the S1 pocket of thrombin. *J. Mol. Biol.* **2009**, *390*, 56.
(46) Liang, J.; Tsui, V.; Abbema, A. V.; Bao, L.; Barrett, K.; Beresini, M.; Berezhkovskiy, L.; Blair, W. S.; Chang, C.; Driscoll, J.; Eigenbrot, C.; Ghilardi, N.; Gibbons, P.; Halladay, J.; Johnson, A.; Kohli, P. B.; Lai, Y.; Liimatta, M.; Mantik, P.; Menghrajani, K.; Murray, J.; Sambrone, A.; Xiao, Y.; Shia, S.; Shin, Y.; Smith, J.; Sohn, S.; Stanley, M.; Ultsch, M.; Zhang, B.; Wu, L. C.; Magnuson, S. Lead identification of novel and selective TYK2 inhibitors. *Eur. J. Med. Chem.* **2013**, *67*, 175.
(47) Liang, J.; van Abbema, A.; Balazs, M.; Barrett, K.; Berezhkovsky, L.; Blair, W.; Chang, C.; Delarosa, D.; DeVoss, J.; Driscoll, J.; Eigenbrot, C.; Ghilardi, N.; Gibbons, P.; Halladay, J.; Johnson, A.; Kohli, P. B.; Lai, Y.; Liu, Y.; Lyssikatos, J.; Mantik, P.; Menghrajani, K.; Murray, J.; Peng, I.; Sambrone, A.; Shia, S.; Shin, Y.; Smith, J.; Sohn, S.; Tsui, V.; Ultsch, M.; Wu, L. C.; Xiao, Y.; Yang, W.; Young, J.; Zhang, B.; Zhu, B.-y.; Magnuson, S. Lead optimization of a 4-aminopyridine benzamide scaffold to identify potent, selective, and orally bioavailable TYK2 Inhibitors. *J. Med. Chem.* **2013**, *56*, 4521. PMID: 23668484.
(48) Hardcastle, I. R.; Arris, C. E.; Bentley, J.; Boyle, F. T.; Chen, Y.; Curtin, N. J.; Endicott, J. A.; Gibson, A. E.; Golding, B. T.; Griffin, R. J.; Jewsbury, P.; Menyerol, J.; Mesguiche, V.; Newell, D. R.; Noble, M. E. M.; Pratt, D. J.; Wang, L.-Z.; Whitfield, H. J. N2-substituted O6-cyclohexylmethylguanine derivatives: Potent inhibitors of cyclin-dependent kinases 1 and 2. *J. Med. Chem.* **2004**, *47*, 3710. PMID: 15239650.
(49) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235.
(50) Pearlman, D. A. A comparison of alternative approaches to free energy calculation. *J. Phys. Chem.* **1994**, *98*, 1487.
(51) Wang, B.; Li, L.; Hurley, T. D.; Meroueh, S. O. Molecular Recognition in a Diverse Set of Protein-Ligand Interactions Studied with Molecular Dynamics Simulations and End-Point Free Energy Calculations. *J. Chem. Inf. Model.* **2013**, *53*, 2659. PMID: 24032517.
(52) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157.
(53) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781.
(54) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration. *J. Chem. Phys.* **1994**, *100*, 9025.
(55) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem. Phys. Lett.* **1994**, *222*, 529.
(56) Sadiq, S. K.; Wright, D. W.; Watson, S. J.; Zasada, S.; Stoica, I.; Coveney, P. V. Automated molecular simulation based binding affinity calculator for ligand-bound HIV-1 protease. *J. Chem. Inf. Model.* **2008**, *48*, 1909.
(57) Highfield, R. Science Museum Blog. https://blog.sciencemuseum.org.uk/supercomputer-bid-to-create-the-first-truly-personalised-medicine/ (accessed November 2, 2016).

DOI: 10.1021/acs.jctc.6b00979
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

(58) Øksendal, B. K. *Stochastic differential equations: an introduction with applications*; Universitext; Springer: Berlin, Heidelberg, NY, 1998; Corrected second printing 2000.
(59) Wan, S.; Coveney, P. V.; Flower, D. R. Molecular Basis of Peptide Recognition by the TCR: Affinity Differences Calculated Using Large Scale Computing. *J. Immunol.* **2005**, *175*, 1715−1723.
(60) Groen, D.; Bhati, A. P.; Suter, J.; Hetherington, J.; Zasada, S. J.; Coveney, P. V. FabSim: Facilitating computational research through automation on large-scale and distributed e-infrastructures. *Comput. Phys. Commun.* **2016**, *207*, 375.
(61) Kramer, C.; Dahl, G.; Tyrchan, C.; Ulander, J. A comprehensive company database analysis of biological assay variability. *Drug Discovery Today* **2016**, *21*, 1213.
(62) Chodera, J. D.; Mobley, D. L. Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. *Annu. Rev. Biophys.* **2013**, *42*, 121.
(63) Mobley, D. L.; Klimovich, P. V. Perspective: Alchemical free energy calculations for drug discovery. *J. Chem. Phys.* **2012**, *137*, 230901.
(64) Lin, Y.-L.; Aleksandrov, A.; Simonson, T.; Roux, B. An overview of electrostatic free energy computations for solutions and proteins. *J. Chem. Theory Comput.* **2014**, *10*, 2690. PMID: 26586504.
(65) Rocklin, G. J.; Mobley, D. L.; Dill, K. A.; Hünenberger, P. H. Calculating the binding free energies of charged species based on explicit-solvent simulations employing lattice-sum methods: An accurate correction scheme for electrostatic finite-size effects. *J. Chem. Phys.* **2013**, *139*, 184103.

DOI: 10.1021/acs.jctc.6b00979
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX