

Evaluation and Characterization of Trk Kinase Inhibitors for the Treatment of Pain: Reliable Binding Affinity Predictions from Theory and Computation

Shunzhou Wan,[†] Agastya P. Bhati,[†] Sarah Skerratt,[‡] Kiyoyuki Omoto,[§] Veerabahu Shanmugasundaram,^{||} Sharan K. Bagal,[‡] and Peter V. Coveney^{*,†}

[†]Centre for Computational Science, Department of Chemistry, University College London, London WC1H 0AJ, United Kingdom

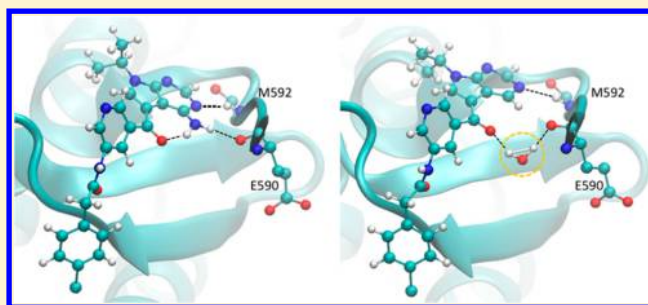
[‡]Pfizer Global R&D, The Portway Building, Granta Park, Cambridge CB21 6GS, United Kingdom

[§]Pfizer Worldwide Research and Development, 610 Main Street, Cambridge, Massachusetts 02139, United States

^{||}Pfizer Worldwide Research and Development, Groton Laboratories, Groton, Connecticut 06340, United States

Supporting Information

ABSTRACT: Optimization of ligand binding affinity to the target protein of interest is a primary objective in small-molecule drug discovery. Until now, the prediction of binding affinities by computational methods has not been widely applied in the drug discovery process, mainly because of its lack of accuracy and reproducibility as well as the long turnaround times required to obtain results. Herein we report on a collaborative study that compares tropomyosin receptor kinase A (TrkA) binding affinity predictions using two recently formulated fast computational approaches, namely, Enhanced Sampling of Molecular dynamics with Approximation of Continuum Solvent (ESMACS) and Thermodynamic Integration with Enhanced Sampling (TIES), to experimentally derived TrkA binding affinities for a set of Pfizer pan-Trk compounds. ESMACS gives precise and reproducible results and is applicable to highly diverse sets of compounds. It also provides detailed chemical insight into the nature of ligand–protein binding. TIES can predict and thus optimize more subtle changes in binding affinities between compounds of similar structure. Individual binding affinities were calculated in a few hours, exhibiting good correlations with the experimental data of 0.79 and 0.88 from the ESMACS and TIES approaches, respectively. The speed, level of accuracy, and precision of the calculations are such that the affinity predictions can be used to rapidly explain the effects of compound modifications on TrkA binding affinity. The methods could therefore be used as tools to guide lead optimization efforts across multiple prospective structurally enabled programs in the drug discovery setting for a wide range of compounds and targets.



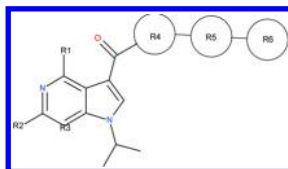
INTRODUCTION

The availability of computational methods that can reliably, rapidly, and accurately predict the binding affinities of ligands to a target protein of interest would greatly facilitate drug discovery programs by enabling project teams to more effectively triage design ideas and therefore synthesize only those compounds with a high probability of being pharmacologically active. The overall effect of this approach would be to reduce the number of “design–synthesis–test” cycles needed to generate compounds of sufficient bioactivity to progress to the clinic. Until recently, *in silico* methods of binding affinity prediction have not been regarded as reliable enough to produce such actionable results. The American Chemical Society’s Cross-Pharmaceutical Industry group has been discussing this particular challenge of conducting elaborative studies and has come out with an opinion piece recently, in which it is specifically noted that “coordinated, blinded prediction challenges offer the best opportunity to develop

broad understanding and broadly applicable methods”.¹ The Coveney group has now developed a suite of computational methods that deliver rapid, accurate, precise, and reliable binding affinity predictions. We report here a prospective computational study on a then-ongoing project at Pfizer² to assess the effectiveness of these methods based on the use of a Binding Affinity Calculator (BAC) software tool and associated services.³ The approach makes use of an automated workflow⁴ running in a high-performance computing environment that builds models, runs large numbers of calculations, and analyzes the output data in order to place reliable error bounds on predicted ligand binding affinities. In order to assess the reliability of these methods, the predictions were performed blind at UCL and subsequently compared with the experimentally determined TrkA binding affinity data.² For a

Received: December 21, 2016

Published: March 20, 2017

Table 1. Ligands Considered in This Study, Numbered as Per the Order in the Files Provided by Pfizer^a

Compound	R1	R2	R3	R4	R5	R6	IC ₅₀ (nM)	ΔG (kcal/mol)
1	NH ₂	H	N				6.97 4.25	-11.20 -11.49
3	NH ₂	H	N				24.53	-10.45
4	H	H	N				10.86	-10.93
6	NH ₂	H	N				8.67	-11.07
7	H	H	CH				44.25	-10.09
8	H	H	CH				735.79	-8.42
11	H	H	N				11.04 5.45	-10.92 -11.34
12	H	H	N				14.18	-10.77
13	H	H	N				224.74	-9.13
16	H	H	N				9800.00	-6.87
17	H	NH ₂	N				6.84 3.73	-11.21 -11.57
22	OH	H	N				1756.36	-7.90
23	H	NHCH ₃	N				9.55	-11.01
24	H	OH	N				249.42	-9.06
26	H	H	N				20.53	-10.55
30	H	H	N				201.70 47.87	-9.19 -10.05

^aAll of the ligands have the same net neutral charge. The experimental TrkA inhibitory values (IC₅₀) and the binding free energies derived from them are shown. Experimental IC₅₀ measurements were conducted independently in two separate laboratories using an identical protocol;² Pfizer, Sandwich (U.K.) IC₅₀ values are shown in black; TCG Lifescience (India) IC₅₀ values are shown in blue.

new method being pursued by an industrial computational chemist, these types of live blinded individual project team collaborations are the best kinds of pilot studies that can be performed.

Tropomyosin receptor kinase A (TrkA) is used as the target protein in this study. TrkA is the cognate receptor of the nerve growth factor (NGF) neuropeptide, a neurotrophic factor involved in the regulation of growth, maintenance, proliferation, and survival of certain target neurons.⁵ Preclinical and clinical studies have identified a crucial role for NGF in the pathogenesis of pain; the clinical efficacy of anti-NGF monoclonal antibody (mAb) therapies against several pain end points is well-documented, and preclinically the inhibition of the TrkA kinase domain by small-molecule kinase inhibitors has been shown to reverse the effects of NGF-mediated pain transduction.^{6,7} In view of the key role of NGF in modulating pain, there is significant interest in the clinical development of

small-molecule TrkA inhibitors to complement anti-NGF mAb treatment options.

The UCL group has recently introduced new approaches termed Enhanced Sampling of Molecular dynamics with Approximation of Continuum Solvent (ESMACS) and Thermodynamic Integration with Enhanced Sampling (TIES)⁸ for the reliable prediction of ligand–protein free energies. The ESMACS approach centers on the molecular mechanics Poisson–Boltzmann surface area (MMPBSA) method,⁹ which employs a continuum approximation for the aqueous solvent, while TIES is based on thermodynamic integration. The approaches emphasize the necessity of invoking ensemble-based sampling to reliably compute macroscopic quantities using microscopic modeling methods.¹⁰ The BAC tool³ is employed to automate much of the complexity of running and marshalling the required molecular dynamics simulations as well as collecting and analyzing data. The BAC tool utilized to perform the ESMACS and TIES studies

constitutes a computational pipeline built from an array of software tools and services and requires access to suitable computing resources. Integration and automation are central to the reliability of the method, ensuring that the results are reproducible and can be delivered rapidly.

■ COMPUTATIONAL METHODS

The TrkA binding affinities of the compounds in Table 1 were calculated using ESMACS and TIES. These compounds were picked as representatives of the full series previously studied experimentally,² covering the dynamic range of the TrkA pharmacology assay and containing the structural features deemed as key determinants of TrkA activity. (A second series of Pfizer pan-Trk inhibitors, structurally dissimilar to those highlighted in Table 1, also showed good agreement between their ESMACS-predicted and experimentally determined TrkA binding affinities. Studies of those inhibitors will be published in due course.) These congeneric compounds all have the same net neutral charge. In ESMACS, the MMPBSA.py.MPI¹¹ module of the AMBER12 package¹² was employed for the free energy calculations of the complex (G_{complex}), the receptor (G_{receptor}), and the compound (G_{ligand}) (eq 1). The polar solvation free energy was calculated using the Poisson–Boltzmann equation with a grid spacing of 0.5 Å and dielectric constants of 1 and 80 for the protein and solvent, respectively. The energetic analyses, including the configurational entropy calculations, were conducted on 50 snapshots of complexes, protein, and compounds extracted evenly from each 4 ns production run in the ensemble MD simulations of complexes only (1-traj and 2-traj; see below) or from separate ensemble simulations of complexes and compounds (3-traj). In TIES, the energy derivatives $\partial V/\partial \lambda$ (eq 2) were recorded every 2 fs during the simulations, and a stochastic integration^{10,13} was performed using a trapezoidal method. Hydrogen-bond analyses were performed for the ensemble simulation trajectories. A hydrogen bond is considered to be formed when the distance between a hydrogen-bond acceptor and a hydrogen-bond donor is less than a defined distance cutoff and the acceptor–hydrogen–donor angle is greater than an angle cutoff. The cutoffs used in the current study are 3.0 Å for the distance and 135° for the angle. The protein and the compounds can have intramolecular and intermolecular hydrogen bonds, of which the intermolecular ones contribute significantly to the binding affinities. Some “bridging” water molecules, which are hydrogen-bonded to the protein and the compounds at the same time, also play an important role in binding of ligands to proteins, in addition to their direct interactions. In ESMACS predictions, the inter- and intramolecular interactions of the compounds and the protein are calculated explicitly, while the interactions with water, including such bridging water molecules, are taken into account implicitly.

In ESMACS, the free energy is evaluated approximately on the basis of the extended MMPBSA method,^{8,14} including configurational entropy and the free energy of association.¹⁵ Free energy changes ($\Delta G_{\text{binding}}$) are determined for the molecules in their solvated states. The binding free energy change is then calculated as

$$\Delta G_{\text{binding}} = G_{\text{complex}} - G_{\text{receptor}} - G_{\text{ligand}} \quad (1)$$

where G_{complex} , G_{receptor} , and G_{ligand} are the free energies of the complex, the receptor, and the ligand, respectively.

The three terms on the right-hand side of eq 1 can be generated from single simulations of the complexes or from separate simulations of complexes, receptor(s), and ligand(s). The former is the so-called one-trajectory (1-traj) method, in which the trajectories of the receptor(s) and ligand(s) are extracted from those of the complexes. The latter is the so-called three-trajectory (3-traj) method when the energies are derived from independent simulations of the three components or the two-trajectory (2-traj) method when the energy is derived from independent simulations for only the receptor or the ligand. In the drug development field, binding is usually investigated for a set of ligands bound to the same protein target. The free energy of the receptor, G_{receptor} , is then the same for all of ligands when it is derived from an independent simulation of the receptor and hence can be treated as a constant. In the current study, we employ three approaches to calculate $\Delta G_{\text{binding}}$: (a) employing only simulations of the complexes (1-traj); (b) the same as (a) but using the average of the receptor free energies $\langle G_{\text{receptor}} \rangle$ from the 1-traj method (denoted as 2-traj); and (c) the same as (b) but also invoking separate ligand simulations for the derivation of G_{ligand} (denoted as 3-traj).

In our more recent TIES method,¹³ an “alchemical transformation”¹⁶ for the mutated entity, either the ligand or the protein (in this study it is always the ligand), is used in both aqueous solution and within the ligand–protein complex. The relative free energy changes for the alchemical mutation processes, $\Delta G_{\text{aq}}^{\text{alch}}$ and $\Delta G_{\text{complex}}^{\text{alch}}$, are calculated as

$$\Delta G^{\text{alch}} = \int_0^1 \left\langle \frac{\partial V(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda \quad (2)$$

where λ ($0 \leq \lambda \leq 1$) is a coupling parameter such that $\lambda = 0$ and $\lambda = 1$ correspond to the initial and final thermodynamic states, $V(\lambda)$ is the potential energy of an intermediate state λ , and $\langle \dots \rangle_{\lambda}$ denotes an ensemble average over configurations representative of the state λ . The relative binding free energy difference is then calculated as

$$\Delta \Delta G^{\text{binding}} = \Delta G_{\text{aq}}^{\text{alch}} - \Delta G_{\text{complex}}^{\text{alch}} \quad (3)$$

Compared with ESMACS, the TIES approach, like all other alchemical-based free energy methods, is usually more accurate in comparing one congeneric ligand to another. However, the nature of these alchemical methods implies that their applications will be limited when diverse sets of compounds are of interest. A recent publication¹⁷ in which the authors claimed that their implementation of free energy perturbation (FEP) calculations successfully predicted a number of more active compounds (a correlation coefficient of 0.71 between the predicted and experimental binding affinities was achieved) has the same limitations. ESMACS, however, has no such limitations. Our previous studies⁸ have shown that ESMACS can be applied to sets of compounds that are highly diverse in terms of both the number of atoms and the net charge. In addition, ESMACS is able to make assessments of the effect that an individual ligand has on the protein to which it binds. In particular, the application of two- and three-trajectory versions of ESMACS allows one to probe the extent to which both the ligand and protein adjust their conformations on binding, broadly according to a “lock and key” or “induced fit” recognition mechanism.

In the current study, we apply the ESMACS approach for all of TrkA compounds highlighted in Table 1 and consider TIES

for a selected subset. In the evaluation and characterization of TrkA binding affinity predictions from the Pfizer compound set, the ESMACS and TIES protocols were conducted as recently described.^{8,13} TrkA compounds were optimized at the Hartree–Fock level with the 6-31G* basis (HF/6-31G*) in Gaussian 03¹⁸ and parametrized using Antechamber and RESP in AmberTools 12 with the general AMBER force field (GAFF).¹⁹ The Amber ff99SBildn force field²⁰ was used for the protein. In ESMACS, we used 25 replicas in an ensemble calculation for each ligand. In TIES, five replicas were used for each selected pair of ligands. The MD package NAMD2.9²¹ was used throughout the equilibration and production simulations with periodic boundary conditions. We used a protocol established in our previous publications^{13,14} in which 2 ns of equilibration and 4 ns of production were conducted for each replica. Energy analyses showed that a 4 ns production run was sufficient to have convergence results for the current molecular systems (see Figure S1 in the Supporting Information). In ESMACS, the MMPBSA.py.MPI¹¹ module of the Amber package¹² was used to extract the free energies of the complexes, the protein, and the compounds (eq 1), including configurational entropy from normal mode calculations.

ESMACS was performed on SuperMUC, two separate high-end supercomputers with a total of ca. 245 000 cores in a combination of various processor technologies (<https://www.lrz.de/services/compute/supermuc/systemdescription/>), at the Leibniz Rechenzentrum (Leibniz Supercomputing Centre) in Garching, Germany. We were able to run our calculations on them as if it were a single supercomputer.²² TIES calculations were performed on ARCHER, a Cray XC30 supercomputer (equipped with ca. 110 000 cores), the U.K.'s National High Performance Computing Service located in Edinburgh. A single ESMACS study of one ligand–protein interaction and a TIES study of the binding free energy difference for two congeneric ligands can both be completed in about 8 wall-clock hours on 1200 and 3120 CPU cores, respectively, on ARCHER (see details in our previous publication^{8,13}). (It is possible to further reduce the time required by use of GPU accelerators.) The ESMACS and TIES approaches are both scalable, allowing a large number of calculations to be performed concurrently, depending only on the computing resources available.

RESULTS

We report a collaborative study performed between academic computational chemists and a pharmaceutical company on a data set of 16 pan-Trk ligands (Table 1).² Of the TrkA cocrystal structures provided by Pfizer,² we elected to utilize the cocrystal structure of TrkA and **1** (Figure 1, PDB code 5JFV) as it contained the highest number of crystallographically defined TrkA residues. The missing residues were built by ModLoop,²³ while the initial structures of the complexes were constructed by a docking method employing UCSF DOCK.²⁴ Nine of the 16 compounds were successfully docked into the binding site of the TrkA protein with orientations agreeing with those from available X-ray structures. Compounds that failed to generate suitable cocomplexes via docking were manually positioned on the basis of X-ray and/or modeled structures of similar compounds.

Comparison of Experimental TrkA Binding Free Energies with ESMACS Predictions. The comparison of the TrkA binding free energies obtained from ESMACS predictions and those determined experimentally is shown in Figure 2. The predicted binding free energies from the 1-traj

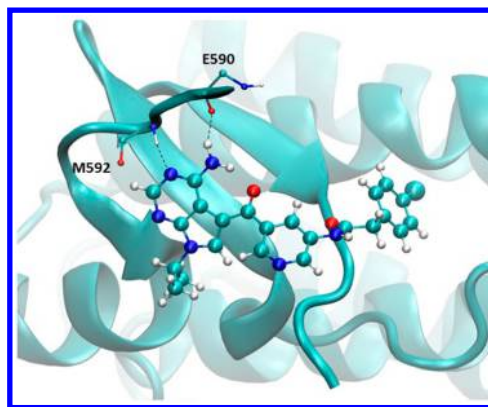


Figure 1. Crystal structure of **1** bound to TrkA, viewed from the N-lobe to the C-lobe of the kinase. Hydrogen bonds are displayed by dashed lines. The protein is shown in cyan cartoon, and ligand atoms are colored by element: hydrogen in white, carbon in cyan, oxygen in red, and nitrogen in blue. For clarity, the N-lobe is not shown.

approach exhibit a moderate Pearson correlation with the experimental data (with a coefficient of 0.42; Figure 2a). The 2-traj approach, in which the same free energy of the receptor averaged from the 1-traj method is used, significantly improves the correlation between the predictions and experimental data (with a coefficient of 0.76; Figure 2b). Incorporating the free energies of the ligands from their individual simulations in water (the 3-traj approach) generates a correlation similar to that from the 2-traj approach (with a coefficient of 0.79; Figure 2c). The improvements are achieved by including the adaptation energies⁸ of the receptor in the current study, and possibly of the ligands in general cases, which quantify the free energy changes of a molecule between its bound and free states. The 3-traj ESMACS results are in good agreement with the experimental measurements. It should be noted that the ESMACS method is capable of comparing the binding affinities of the entire set of ligands, notwithstanding the relatively significant structural differences. It should also be noted that the experimental TrkA binding free energies were approximated from half-maximal inhibitory concentration (IC₅₀) values, which provide only semiquantitative estimates. In addition, experimental IC₅₀ values for individual compounds vary between intralab measurements, and their average IC₅₀ values vary between 1.6- and 4.2-fold when measured independently across the two laboratories (Table 1 and Table S1 in the Supporting Information). A root-mean-square (RMS) error of ~24% has recently been inferred from interlaboratory variations of reported binding affinities by Chodera and Mobley.²⁵

Improvement of Docking Predictions by the ESMACS Approach. For the nine compounds that were successfully docked into the TrkA protein, the binding free energies were also calculated using the structures after docking and energy minimization. The minimization process optimizes the geometry of a collection of atoms so that the net interatomic force on each atom is acceptably close to zero. The remaining seven compounds were not included in these calculations because the manually constructed structures exhibit some close contacts between the protein and the compounds, to which the binding free energy estimations are very sensitive. When the compounds could be docked into a single protein structure, the ranking of their binding affinities was reasonably predicted from the docked structures (Figure 3a). However, the results from

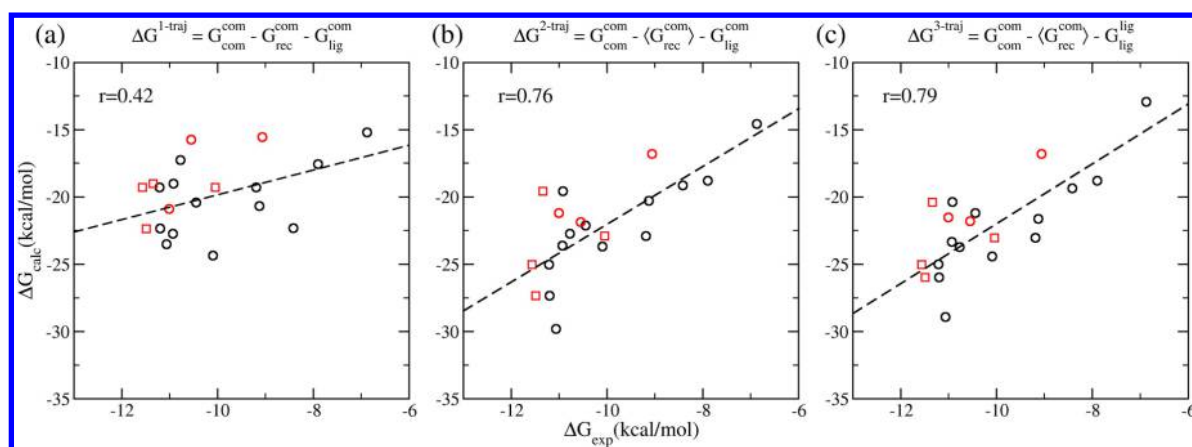


Figure 2. Comparisons of the TrkA experimental data and the calculated binding free energies from (a) 1-traj, (b) 2-traj, and (c) 3-traj ESMACS approaches. The equation in each panel indicates the calculations used in each case. The term $G_{\text{com/rec/lig}}^{\text{com/lig}}$ represents the free energy of the component (subscript: complex, receptor, or ligand) obtained from simulation of the same or different component (superscript: complex or free ligand in aqueous solution). The calculated binding free energies are associated with standard errors of ca. 0.6 kcal/mol in the 1-traj approach and ca. 3 kcal/mol in the 2- and 3-traj approaches (see Table S2), which are not shown in the figures for reasons of clarity. The experimental data from two sites (Pfizer, Sandwich and TCG Lifescience) are displayed in black and red, respectively. The correlation coefficients shown in the figure were calculated using the averages of the calculated binding free energies and the experimental data from Pfizer, Sandwich (black circles) and TCG Lifescience (red circles) where the former are not available. Large uncertainties are associated with the correlation coefficients because of the large error bars of the calculated and experimental binding affinities. Further analyses with bootstrapping resampling (see details in the Supporting Information) generate correlation coefficients of 0.39 ± 0.26 , 0.60 ± 0.26 , and 0.62 ± 0.27 for the 1-, 2-, and 3-traj approaches, respectively. It is evident that the 2- and 3-traj methods improve the ranking provided by the 1-traj method.

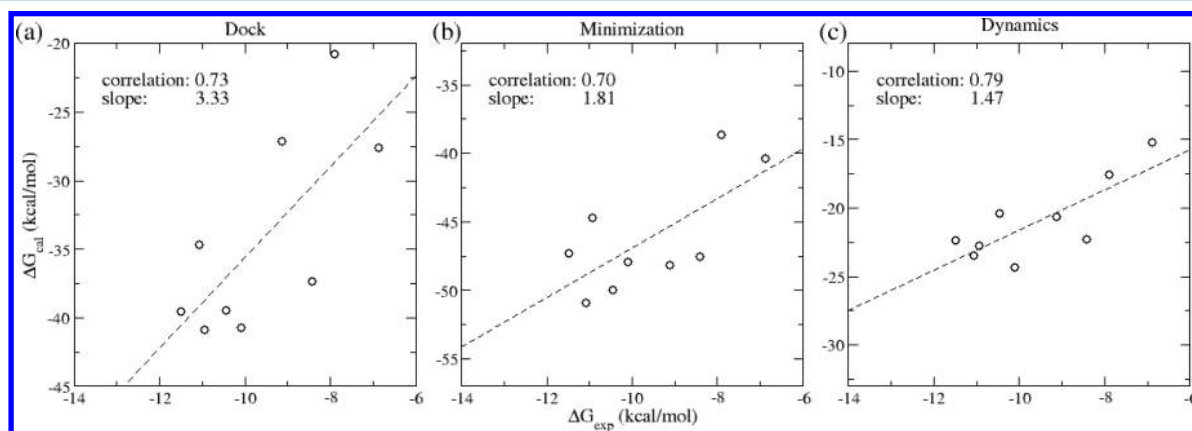


Figure 3. Calculated binding free energies for 1, 3, 4, 6, 7, 8, 13, 16, and 22 (a) from single structures after docking, (b) from 25 structures for each compound after 11 000-step minimization with a sophisticated conjugate-gradient method, and (c) from ensemble averages from three-trajectory ESMACS studies. Compounds 1, 3, 4, 6, 7, 8, 13, 16, and 22 are highly structurally similar (see Table 1).

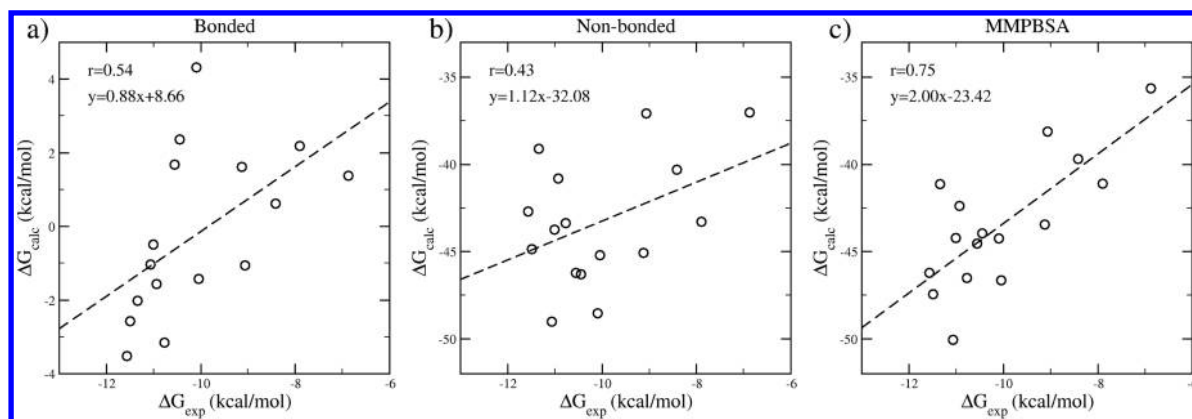


Figure 4. Correlations of free energy components and the experimental data from the 3-traj approach. Both (a) bonded and (b) nonbonded energy terms contribute to the ranking of binding affinities, with similar correlation coefficients between the calculations and experimental data. Their combination, the MMPBSA energy (c), exhibits better correlations with the experimental data than the components themselves.

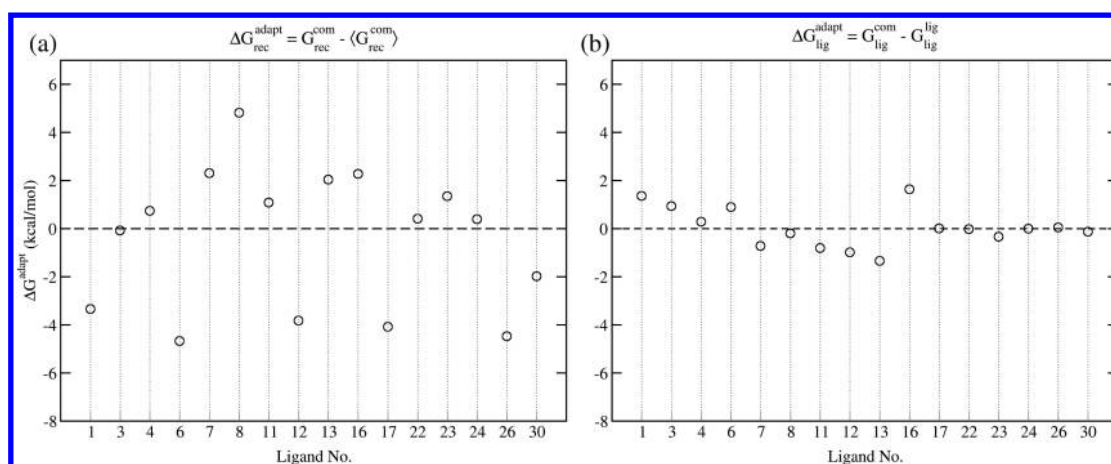


Figure 5. Adaptation free energies of (a) the receptor and (b) the compounds showing the binding free energy changes between the 1- and 3-traj approaches. The terms $G_{\text{rec/lig}}^{\text{com/lig}}$ are the free energies of receptor or ligands (subscript) calculated from simulations performed for the complex or the ligands (superscript).

the docking structures overestimate the free energy differences among the compounds, as shown in Figure 3a, in which the large slope of the regression line (3.33) indicates that the predicted energy differences are on average more than 3 times larger than the experimentally measured ones. The energy minimization significantly decreases the overestimations (Figure 3b, with a regression slope of 1.81). The ESMACS study further decreases the overestimation and makes the points on the scatter plot more concentrated around the regression line (Figure 3c). The docking method ranks the binding affinities reasonably well only for the compounds with closely related structures. Its results are also strongly dependent on the quality and choice of the structure(s) of the target protein.²⁶ The ESMACS approach, however, is not sensitive to the initial structure and achieves a similar correlation coefficient between the calculated and experimental binding affinities for all of the compounds irrespective of whether they can be successfully docked (Figure 2c).

ESMACS Binding Affinities: Contributions from Energy Components. The components of the free energy calculations (see the Supporting Information for more details) may provide insight into the mechanism of compound binding.²⁷ In this study, no apparent correlations could be found for the individual energy components of the internal van der Waals, electrostatic, or electrostatic solvation energies. The total bonded energies, including bond, angle, and dihedral interactions, manifest a moderate correlation with the experimental binding free energies in the 3-traj approach, with a Pearson coefficient of 0.54 (Figure 4a). The nonbonded interactions, including the internal electrostatic, solvation electrostatic, and van der Waals interactions, have a slightly weaker correlation (Pearson coefficient of 0.43) than the bonded energies. The combination of bonded and nonbonded interactions significantly improves the correlation, with a Pearson coefficient of 0.75 (Figure 4c). The bonded and nonbonded energies are the two components of the total adaptation energy (see below), which is associated with the conformational changes upon compound binding. It is therefore not surprising that both of them show correlations with the binding affinities. The configurational entropy components do not exhibit a significant correlation with the experimental binding affinity (Figure S2). They vary in a range of 4.6 kcal/mol; the MMPBSA energies vary in a range of 14.41

kcal/mol by comparison. Inclusion of the contribution from configurational entropy into the MMPBSA energy does not improve the Pearson correlation coefficients significantly (see Figure 2c, which includes configurational entropy, and Figure 4c, which does not).

Adaptation Energy: A Measure of Conformational Change upon Binding. The improvement obtained with the 3-traj approach compared with the 1-traj approach is achieved by relaxing the assumption that the receptor and the compounds sample similar conformations in both the free and bound states. Unfavorable adaptation energies⁸ are usually induced when the conformations within the free state are significantly perturbed upon compound binding. The adaptation energies from the 3-traj approach indeed provide more insights into the mechanism of compound binding. They are related to the structural modifications made to the compounds and the protein and shed important light on the optimization. As the ranking of the ligand binding is the main concern of the study and the relative adaptation energies of the protein are energetically as informative as the absolute ones in the binding affinity comparison, no attempt has been made here to compute accurate absolute adaptation energies for the protein. All of the adaptation energies for the protein are relative ones in the current paper. Non-negligible adaptation energies are associated with conformational changes upon compound binding, meaning that the binding involves “induced fit” recognition. Compound 8, for example, introduces the largest adaptation energy within the protein, while 16 has the largest adaptation energy within the compound (Figure 5). The binding of 8 induces a conformational change within the protein because there is a reluctance to accommodate the methoxy group in the linker region of 8 into the binding site; the arrangement of the amide group in 16 forces the compound to adopt a high-energy conformation in the protein’s binding site (see details below).

Comparison of Experimental TrkA Binding Free Energies with TIES Predictions. The TIES method provides a more accurate approach to estimate relative binding affinities, but the scope for its use is rather tightly circumscribed, as it is applicable only for pairs of compounds that do not present significant structural differences. Our previous studies have demonstrated that TIES offers more quantitative accuracy in its predictions than ESMACS.^{13,28} Here we apply TIES to the

TrkA data set to highlight the accuracy of the approach. In TIES, studies are performed for 14 pairs of TrkA compounds (Figure 6). The calculated binding free energy differences

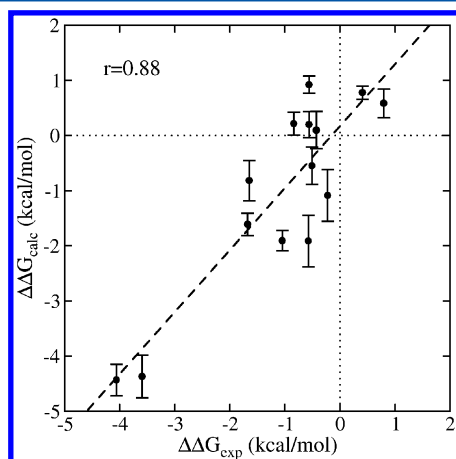


Figure 6. Correlation between TIES-predicted relative binding affinities and experimental data. The black line is the correlation line, while the dotted lines ($x = 0$ and $y = 0$) create four quadrants. Ten out of the 14 data points are in quadrants I ($x > 0$ and $y > 0$) and III ($x < 0$ and $y < 0$), meaning that the calculated binding free energy differences have the same sign as those from the experimental data.

correlate well with the experimental data, with a Pearson correlation of 0.88 (Figure 6), compared with 0.79 from the 3-traj ESMACS study (Figure 2c). A directional agreement is achieved if a prediction has the same sign as the experimental observation; otherwise, the result is deemed to be directional disagreement. For 10 out of the 14 pairs studied, TIES successfully achieves directional agreements, meaning that the results predict the direction of the change in the binding affinity correctly (Figure 6). As shown in the figure, all but one of the TIES predictions that directionally disagree lie on the border of quadrants II and III within their calculated error bars. When the error bars of the experimental $\Delta\Delta G$ are also taken into account, a better agreement might be achieved than that reported here.

DISCUSSION

The calculations, especially the ESMACS ones, can provide further insights for binding affinity variation between structurally similar ligands. In order to analyze the results, it is necessary to consider the differences between the one-, two-, and three-trajectory methods. The 1-traj simulation calculates $\Delta G_{\text{binding}}$ assuming there is no energetic penalty for the adaptation of the protein or ligand to the binding conformation; the 2-traj method takes the change in protein energy into consideration, and the 3-traj method accounts for the changes in both protein and ligand energies. The simulations enable the exploration of key differences in binding modes for structurally related ligands, as we now discuss.

Hinge Binding Group Modifications: Compounds 1, 4, and 22. Initial compounds in the pyrrolopyrimidine series, such as **1**, contain an amino group at the 4-position of the hinge binding group (Table 2). Compounds such these, while TrkA-active, required optimization of their kinase selectivity profile.²⁹ It is known within the scientific literature that minimizing the number of hydrogen-bonding interactions made between a kinase inhibitor and the kinase hinge binding region can lead to an enhanced kinome selectivity profile.^{30,31} In order to design compounds with an optimal kinase selectivity profile, an analysis of the crystal structure of **1** bound to TrkA and the nonliganded TrkA apo structure was undertaken.²⁹ The aminopyrrolopyrimidine motif of **1** makes a two-point hydrogen-bonding interaction with hinge residues E590 and M592. When **1** is superpositioned onto the TrkA apo crystal structure, the $-\text{NH}_2$ motif of the aminopyrrolopyrimidine group overlays directly with a conserved water molecule observed in the apo structure. The $-\text{NH}_2$ group was therefore removed to generate compounds such as **4**, in which a bridging water molecule forms hydrogen bonds with the backbone carbonyl oxygen at E590 and the ketone oxygen of the TrkA ligand. The hydrogen bond between the ligand and the backbone N–H of hinge residue M592 is maintained. The deletion of the $-\text{NH}_2$ group had only a minimal effect on the TrkA activity, as **1** and **4** had broadly similar IC_{50} values in the TrkA pharmacology assay.

ESMACS calculations show that compounds without the $-\text{NH}_2$ group at the 4-position (all compounds except **1**, **3**, **6**, and **22**; Table 2) have a bridging water molecule between the

Table 2. Structures and Properties of **1**, **4**, and **22**^a

Compound	1	4	22
IC_{50} (nM)	4.25/6.97	10.86	1756.36
ΔG_{exp} (kcal/mol)	-11.49/-11.20	-10.93	-7.91
$\Delta\Delta G_{\text{exp}}$ (kcal/mol)	0.00	0.56/0.27	3.58/3.29
$\Delta\Delta G_{\text{ESMACS}}$ (kcal/mol)	0.00	2.63	7.18
$\Delta\Delta G_{\text{TIES}}$ (kcal/mol)	0.00	-0.20	4.37

^aTrkA free energy activities derived from experimental IC_{50} values are denoted as ΔG_{exp} . The relative binding free energies from experiment and ESMACS and TIES calculations are denoted as $\Delta\Delta G_{\text{exp}}$, $\Delta\Delta G_{\text{ESMACS}}$, and $\Delta\Delta G_{\text{TIES}}$, respectively. Experimental data from Pfizer, Sandwich (U.K.) are shown in black, and those from TCG Lifescience (India) are shown in blue.

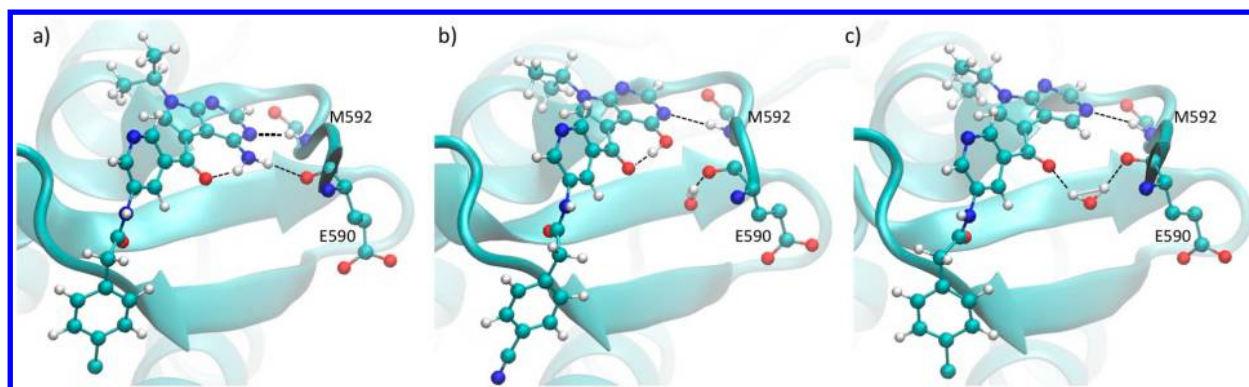


Figure 7. ESMACS simulations of (a) **1**, (b) **22**, and (c) **4** bound to TrkA. Final conformations of 4 ns production runs from all 25 replicas are overlapped and smoothed by averaging over 10 frames. Key residues E590 and M592 (for clarity, side-chain atoms are not shown) at the TrkA hinge region are highlighted. Hydrogen bonds are shown as dashed lines. The protein is shown in cyan cartoon, and ligand atoms are colored by element: hydrogen in white, carbon in cyan, oxygen in red, and nitrogen in blue. For reasons of clarity, the N-lobe of the protein has been removed.

Table 3. Structures and Properties of 12, 17, 23, and 24^a

Compound	12	17	23	24
IC ₅₀ (nM)	14.18	3.73/6.84	9.55	249.42
ΔG _{exp} (kcal/mol)	-10.77	-11.57/-11.21	-11.01	-9.06
ΔΔG _{exp} (kcal/mol)	0.80/0.44	0.00	0.56/0.20	2.51/2.15
ΔΔG _{ESMACS} (kcal/mol)	1.29	0.00	3.49	8.21
ΔΔG _{TIES} (kcal/mol)	0.58	0.00	-0.92	-

^aTrkA free energy activities derived from experimental IC₅₀ values are denoted as ΔG_{exp}. The relative binding free energies from experiment and ESMACS and TIES calculations are denoted as ΔΔG_{exp}, ΔΔG_{ESMACS}, and ΔΔG_{TIES}, respectively. Experimental data from Pfizer, Sandwich (U.K.) are shown in black, and those from TCG Lifescience (India) are shown in blue.

ketone oxygen of the ligand and the backbone carbonyl oxygen of hinge residue E590, with an average occupancy of 41 ± 9%. The occupancy indeed underscores the presence of a water molecule at the location as the molecule dynamically moves into and out of the cutoff distance and angle that designate the occurrence of a bridging water molecule. Simulations of compounds containing the -NH₂ group at the 4-position (**1**, **3**, and **6**; Table 2) show the same bridging water to be present but at a far lower frequency (occupancy values of just 2 ± 0% for **1**, **3**, and **6**) (Figure 7).

ESMACS confirms that the bridging water molecule highlighted in Figure 7 helps mediate and thus satisfy key protein–compound interactions, rendering the overall binding free energies of compounds such as **4** less affected by the absence of the H-bond when -NH₂ is not present. This capability is reflected in the close similarity between the calculated TrkA free energy values for **1** and **4** and those derived from TrkA pharmacology experiments (Table 2). Compound **22** was synthesized as part of an effort to characterize putative compound metabolites. As highlighted in Table 2, **22** is only weakly active in the TrkA pharmacology (IC₅₀ = 1756 nM). ESMACS simulations show a low average occupancy of the aforementioned bridging water molecule (average occupancy 10 ± 6%) due to the presence of the -OH

group at the 4-position (Figure 7). The calculations also show that the -OH group of **22** is more likely to form an intramolecular H-bond with the carbonyl group of the ligand than to form an intermolecular H-bond with hinge residue E590 of the protein. Within the 25-member ensemble, only one ESMACS replica has an intermolecular H-bond forming with an occupancy of 84%, compared with <5% for all of the others. The -OH group occupies the space of the bridging water molecule and does not have a direct hydrogen-bonding interaction with E590. The intramolecular H-bond of the -OH group places the -OH oxygen in a position where an electrostatic repulsion arises with the backbone carbonyl oxygen of E590 (Figure 7b). The result of having a nonsatisfied hydrogen-bonding group (-OH) in **22** is a reduction in overall ligand binding affinity. This is reflected not only in the measured activity data (ΔG_{exp} = -7.1 kcal/mol) but also in the relative binding affinities calculated by ESMACS and TIES (Table 2).

Hinge Binding Group Modifications: Compounds 12, 17, 23, and 24. The TrkA cocrystal structure of compounds such as **12** indicated that a polar atom capable of making an additional hydrogen-bonding interaction with M592 might be accommodated at the 2-position of the hinge binding group (Table 3). Compound **17** was synthesized initially to test this

theory and proved to be more active in the TrkA pharmacology model than the parent molecule **12**. Compound **23** was then synthesized to see whether adding a methyl group to the amine would further boost the potency through the introduction of hydrophobic interactions with the lipophilic side chain of Y591. Compound **23** was potent in the TrkA assay, albeit slightly less so than **12**. Compound **24** was synthesized as a putative metabolite of **12** and was found to be >15-fold less active at TrkA.

ESMACS shows that addition of polar groups at the 2-position increases the total occupancy of hydrogen bonds with M592 from $23 \pm 4\%$ (for **12**) to $85 \pm 4\%$, $80 \pm 4\%$, and $91 \pm 19\%$ (for **17**, **23**, and **24**, respectively). This includes the occupancy of the hydrogen bond made between the pyrimidyl nitrogen and the $-\text{NH}$ group of M592, which is common to all of the compounds. The increased occupancies of **17**, **23**, and **24** result in the additional hydrogen bonds between the polar groups at the 2-position of the compounds and the $\text{C}=\text{O}$ group of M592. The simulations also show that addition of an $-\text{NH}_2$ group (**17**) does not introduce any unfavorable steric interaction within the protein or the ligand (see Figure 8). The

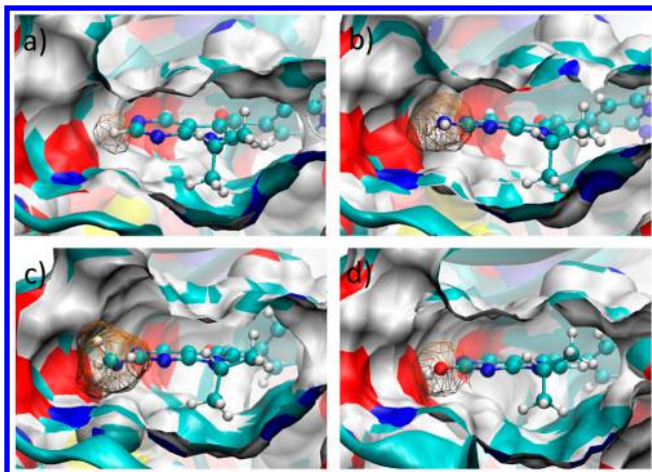


Figure 8. ESMACS simulations of (a) **12**, (b) **17**, (c) **23**, and (d) **24** bound to TrkA. Representative conformations are displayed using the final conformations of 4 ns production runs of one replica from each 25-member ensemble. The protein is shown by the surface representation and the ligand by the ball-and-stick representation with hydrogen in white, carbon in cyan, oxygen in red, and nitrogen in blue. The surfaces of the R2-position groups are shown in the wireframe representation.

slightly improved binding affinity of **17** versus **12** stems from the favorable electrostatic interaction (including the solvation electrostatic component from the Poisson–Boltzmann calculation) between the compound and the protein. The addition of the methyl group on the amine (**23**), however, induces an unfavorable adaptation energy within the protein (see the comparison of relative adaptation energies for **12/17** and **23** in Figure 5a) due to steric hindrance between the protein and **23** (Figure 8c). The presence of an $-\text{NH}_2$ group at the 2-position (**17**) does not induce an unfavorable adaptation energy. The introduction of an $-\text{OH}$ group (**24**), however, introduces a significant adaptation energy in the protein because the $-\text{OH}$ group forms the strongest hydrogen bond with M592 among the subgroup of compounds **12**, **17**, **23**, and **24**, which induces steric hindrance within the complex and reduces the overall binding affinity (Figure 8).

Linker Group Modifications: Compounds **1**, **4**, and **22**.

The TrkA cocrystal structure of compounds such as **1** highlighted a conserved water molecule that formed hydrogen-bonding interactions with the $\text{N}-\text{H}$ group of the amide linker and protein residues K544 and E560 (Table 4 and Figure 9). A methyl group was added to the amide nitrogen to assess the effect of displacing the conserved water molecule (and the potentially altered amide conformation relative to the pyridyl ring) on the TrkA activity. The effect of transposing the amide $\text{N}-\text{H}$ to the other side of the carbonyl group was also investigated, as this, if the activity were retained, could open up new parallel chemistry opportunities within the chemotype. Compound **3**, an example of the *N*-methylamide, remained active at TrkA, although with reduced affinity versus its desmethyl congener **1**. Compound **6**, an example of the transposed $\text{N}-\text{H}$ amide, also remained active at TrkA with almost identical affinity as **1**.

ESMACS shows that the modifications highlighted in Table 4 do not alter the key hydrogen-bonding interaction between the carbonyl oxygen of the amide group of **1**, **3**, and **6** and the backbone $\text{N}-\text{H}$ of D668, with occupancies of $56 \pm 7\%$, $54 \pm 3\%$, and $51 \pm 6\%$, respectively (Figure 9). The altered position of the amide group in **1** and **6** slightly affects the appearance of the bridging water molecules between the ligands and residue E560, with occupancies of $15 \pm 6\%$ and $20 \pm 5\%$ for **1** and **6**, respectively. Minimal direct hydrogen-bonding interactions are predicted between the amide $-\text{NH}$ group of **1** and **6** and residue K544 or E560. The replacement of the amide $-\text{NH}$ group of **1** by the more hydrophobic *N*-methyl group is predicted to displace the adjacent bridging water molecule (Figure 9a,b) and overall make binding to TrkA energetically less favorable. There is a good correlation between the ESMACS-predicted and experimentally derived free energies of TrkA binding for **1**, **3**, and **6** (Table 4).

Linker Group Modifications: Compounds **7** and **8**.

Although the linker group of the pyrrolopyrimidine/pyrrolopyridine series binds to a relatively narrow region of the ligand binding site, between the two adjacent ATP and DFG pockets, the tolerance of the TrkA protein toward bulkier linker groups was briefly investigated as part of the TrkA program. An example from this work is **8**, in which a methoxy group has been added to the linker-group carbonyl. As highlighted in Table 5, a significant reduction in TrkA activity is observed for **8** versus **7**.

ESMACS calculations indicate that the oxygen atom of the methoxy group of **8** is located proximal ($3.47 \pm 0.31 \text{ \AA}$) to the carbonyl group of residue D668. The occupancy of a bridging water molecule between the methoxy group and the carbonyl is only $5 \pm 2\%$, indicating that the orientation and the space between these groups preclude a bridging water molecule being available to quench any electrostatic repulsion (Table 5 and Figure 10). The chirality of the methoxy group also prevents the polar oxygen from having any direct or favorable water-bridged interactions with E560 or D668. The binding of **8** induces the largest adaptation energy (Figure 5a) in the TrkA protein of all compounds assessed. Energetic analyses show that although the bonded energy and the van der Waals interactions are favorable for **8** compared with **7** (by 3.70 and 2.29 kcal/mol, respectively), a significantly unfavorable electrostatic energy (10.59 kcal/mol), mainly from the interactions between the polar oxygen of the methoxy group and the polar/charged groups of the protein (carbonyl group of D668 and carboxylate

Table 4. Structures and Properties of 1, 3, and 6^a

Compound	1	3	6
IC ₅₀ (nM)	4.25/6.97	24.53	8.67
ΔG _{exp} (kcal/mol)	-11.49/-11.20	-10.45	-11.07
ΔΔG _{exp} (kcal/mol)	0.00	1.04/0.75	0.42/0.13
ΔΔG _{ESMACS} (kcal/mol)	0.00	4.78	-2.94
ΔΔG _{TIES} (kcal/mol)	0.00	1.91	-0.10

^aTrkA free energy activities derived from experimental IC₅₀ values are denoted as ΔG_{exp}. The relative binding free energies from experiment and ESMACS and TIES calculations are denoted as ΔΔG_{exp}, ΔΔG_{ESMACS}, and ΔΔG_{TIES}, respectively. Experimental data from Pfizer, Sandwich (U.K.) are shown in black, and those from TCG Lifescience (India) are shown in blue.

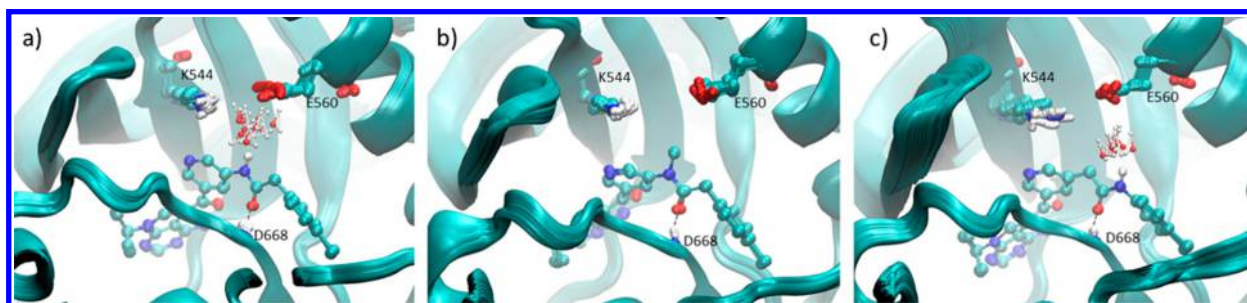


Figure 9. ESMACS simulations of (a) 1, (b) 3, and (c) 6 bound to TrkA. Final conformations of 4 ns production runs from all 25 replicas are overlapped and smoothed by averaging over 10 frames. The key TrkA protein residues D668, K544, and E560 are highlighted. The protein is shown in cyan cartoon, and ligand atoms are colored by element: hydrogen in white, carbon in cyan, oxygen in red, and nitrogen in blue. Hydrogen bonds between the linker group carbonyl oxygen of 1, 3, and 6 and the backbone –NH group of D668 are shown as dashed lines.

Table 5. Structures and Properties of 7 and 8^a

Compound	7	8
IC ₅₀ (nM)	44.25	735.79
ΔG _{exp} (kcal/mol)	-10.09	-8.42
ΔΔG _{exp} (kcal/mol)	0.00	1.67
ΔΔG _{ESMACS} (kcal/mol)	0.00	5.06
ΔΔG _{TIES} (kcal/mol)	0.00	1.61

^aTrkA free energy activities derived from experimental IC₅₀ values are denoted as ΔG_{exp}. The relative binding free energies from experiment and ESMACS and TIES calculations are denoted as ΔΔG_{exp}, ΔΔG_{ESMACS}, and ΔΔG_{TIES}, respectively. Experimental data were obtained from Pfizer, Sandwich (U.K.) for these two compounds.

group of E560; Figure 10), is introduced by the inclusion of the (S)-OMe group.

The calculated free energies of TrkA binding for 7 and 8 correlate well with those obtained experimentally. Hence,

utilization of ESMACS to predict the TrkA binding affinity of 8 during the Pfizer TrkA project might well have influenced the medicinal chemistry team to deprioritize this compound and redirect synthetic efforts in other directions.

Linker Group Modifications: Compounds 4 and 16.

The hydrogen bond between the linker-group carbonyl and the backbone N–H of D668 is a key interaction across the pyrrolopyrimidine/pyrrolopyridine series. During the Pfizer TrkA program, reversing the amide group was modeled to assess whether the usual ligand binding mode could be adapted to maintain this interaction in the new amide arrangement. Docking of 16 into TrkA using a Pfizer docking protocol (data not shown) suggested that the reverse amide would likely not be able to adapt a conformation in which this interaction would be retained. However, 16 was synthesized as an example of a reverse amide system to challenge and/or verify the docking result. As can be seen in Table 6, reversing the amide group led to a significant reduction in TrkA activity.

ESMACS-based component binding analysis shows that 16 has the least favorable bonding and nonbonding interactions of the compounds in Table 1. The reversed positions of the –NH and C=O groups prevent formation of the hydrogen bond between the ligand and residue D668 (Figure 11), with an H-bond occupancy of <1% compared with occupancies of 30–56% for the rest of the compound set (except for 26, whose

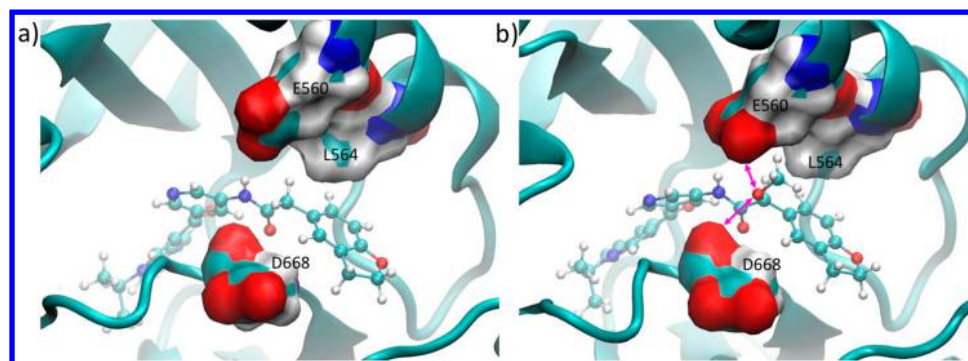


Figure 10. ESMACS simulations of (a) 7 and (b) 8 bound to TrkA. Final conformations of 4 ns production runs from all 25 replicas are overlapped and smoothed by averaging over 10 frames. The key TrkA protein residues D668, E560, and L564 are highlighted. The protein is shown in cyan cartoon, and ligand atoms are colored by element: hydrogen in white, carbon in cyan, oxygen in red, and nitrogen in blue. The unfavorable electrostatic interactions between the methoxy group of 8 and the carbonyl group of D668 and between the methoxy group and the carboxylate group of E560 are highlighted by pink arrows.

Table 6. Structures and Properties of 4 and 16^a

Compound	4	16
IC ₅₀ (nM)	10.86	9800.00
ΔG_{exp} (kcal/mol)	-10.93	-6.87
$\Delta\Delta G_{\text{exp}}$ (kcal/mol)	0.00	4.06
$\Delta\Delta G_{\text{ESMACS}}$ (kcal/mol)	0.00	10.42
$\Delta\Delta G_{\text{TIES}}$ (kcal/mol)	0.00	4.43

^aTrkA free energy activities derived from experimental IC₅₀ values are denoted as ΔG_{exp} . The relative binding free energies from experiment and ESMACS and TIES calculations are denoted as $\Delta\Delta G_{\text{exp}}$, $\Delta\Delta G_{\text{ESMACS}}$, and $\Delta\Delta G_{\text{TIES}}$, respectively. Experimental data were obtained from Pfizer, Sandwich (U.K.) for these two compounds.

occupancy is 14% as a result of the competition from an extra hydrogen bond between the polar oxygen at R4 and residue K544 (Table 1). Indeed, ESMACS calculations reveal that the -NH and C=O groups of 16 align in a noncomplementary fashion with the -NH and C=O groups of residue D668, introducing a large unfavorable electrostatic interaction (Figure

11). Compound 16 has the largest ligand adaptation energy, indicating that the switching of N-H and C=O does not enable an alternate binding mode to be adopted, and therefore, 16 must adopt a high-energy conformation to complex with the protein. As is clear in Table 6, ESMACS correctly predicts the reduced TrkA potency of 16 compared with 4.

CONCLUSIONS

Methodologies that can predict the binding affinities of molecules ahead of synthesis represent a key area of interest in the pharmaceutical industry. The ability to effectively prioritize prospective compounds on the basis of their likely activity in a key pharmacology assay offers the potential to rapidly improve lead optimization timelines in drug discovery programs by reducing the number of “design–synthesis–screen” cycles needed to identify candidate-quality molecules suitable for clinical testing.³²

Despite the relative diversity of the compounds in this study, the 3-traj version of the ESMACS approach provides good agreement between the theoretically predicted binding free energies and those derived from experimentally measured activities for the entire data set. TIES generates even better agreement between the calculated and experimental binding free energy differences for a selected subgroup of ligands. This is manifest in the better correlation coefficient obtained in TIES (0.88) than in ESMACS (0.79) and accurate binding free energy differences from the former. This is the case for all of

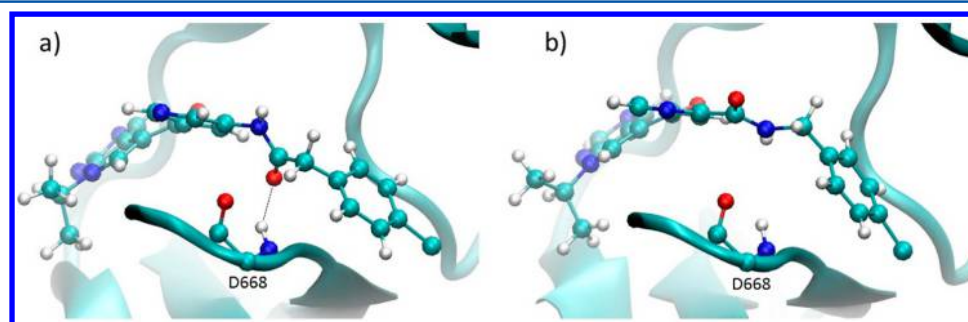


Figure 11. ESMACS simulations of (a) 4 and (b) 16 bound to TrkA. Final conformations of 4 ns production runs from all 25 replicas are overlapped and smoothed by averaging over 10 conformations. The key TrkA protein residue D668 is highlighted. The protein is shown in cyan cartoon, and ligand atoms are colored by element: hydrogen in white, carbon in cyan, oxygen in red, and nitrogen in blue. The hydrogen bond between the linker-group carbonyl oxygen of 4 and the backbone -NH group of D668 is shown as a dashed line. As can be seen in (b), switching of the N-H and C=O groups prevents the formation of a hydrogen bond between the carbonyl oxygen of 16 and the backbone -NH group of D668.

molecular systems we have studied.^{13,28,33} ESMACS also provides structural and energetic insight into the binding of individual ligand–protein systems: some compounds bind by a “lock and key” recognition mechanism, for which no significant conformational adjustment is required by the protein and the ligands; others bind according to an “induced fit” recognition mechanism involving significant conformational changes associated with non-negligible adaptation energies.

The Binding Affinity Calculator (BAC) used in this study automates the workflows for ESMACS and TIES calculations, making the use of these approaches easy and user-friendly. With powerful computing resources now widely available, these robust approaches should become more routine for industrial groups in the field of structure-based drug design. The results described herein, based on an analysis of TrkA ligands synthesized as part of the TrkA program at Pfizer, suggest that the binding affinity calculations have the potential to be successfully applied in real-time prospective ligand design.

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.6b00780.

Detailed description of the energy decomposition, the energy convergence, and the error analyses, experimental measurements of TrkA inhibitory activity, and the predicted binding free energies (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: p.v.coveney@ucl.ac.uk.

ORCID

Peter V. Coveney: 0000-0002-8787-7256

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge the support from the EPSRC via the 2020 Science Programme (<http://www.2020science.net/>), EP/I017909/1), the EU H2020 Projects ComPat (<http://www.compat-project.eu/>, 671564) and CompBioMed (<http://www.compbioimed.eu/>, 675451), the Qatar National Research Fund (7-1083-1-191), the MRC Medical Bioinformatics Project (MR/L016311/1), and the UCL Provost. We are grateful to the Hartree Centre for access to the BlueWonder2 computer and for the help of its scientific support staff. We acknowledge the Leibniz Supercomputing Centre for providing access to SuperMUC and the assistance of its scientific support staff. We also made use of ARCHER, the U.K.'s National High Performance Computing Service, funded by the Office of Science and Technology through the EPSRC's High-End Computing Programme. Access to ARCHER was provided through the 2020 Science Programme. This research was also partially supported by the PLGrid Infrastructure, through which access to Prometheus, the Polish supercomputer run by ACK Cyfronet AGH in Krakow, was provided. A.P.B. is supported by an Overseas Research Scholarship from UCL and an Inlaks Scholarship from the Inlaks Shivdasani Foundation.

■ REFERENCES

(1) Sherborne, B.; Shanmugasundaram, V.; Cheng, A. C.; Christ, C. D.; Desjarlais, R. L.; Duca, J. S.; Lewis, R. A.; Loughney, D. A.; Manas,

E. S.; McGaughey, G. B.; Peishoff, C. E.; van Vlijmen, H. Collaborating to Improve the Use of Free-Energy and Other Quantitative Methods in Drug Discovery. *J. Comput.-Aided Mol. Des.* **2016**, *30*, 1139–1141.

(2) Skerratt, S. E.; Andrews, M.; Bagal, S. K.; Bilsland, J.; Brown, D.; Bungay, P. J.; Cole, S.; Gibson, K. R.; Jones, R.; Morao, I.; Nedderman, A.; Omoto, K.; Robinson, C.; Ryckmans, T.; Skinner, K.; Stuppel, P.; Waldron, G. The Discovery of a Potent, Selective, and Peripherally Restricted Pan-Trk Inhibitor (PF-06273340) for the Treatment of Pain. *J. Med. Chem.* **2016**, *59*, 10084–10099.

(3) Sadiq, S. K.; Wright, D.; Watson, S. J.; Zasada, S. J.; Stoica, I.; Coveney, P. V. Automated Molecular Simulation Based Binding Affinity Calculator for Ligand-Bound HIV-1 Proteases. *J. Chem. Inf. Model.* **2008**, *48*, 1909–1919.

(4) 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*, 375–385.

(5) Norman, B. H.; McDermott, J. S. Targeting the Nerve Growth Factor (NGF) Pathway in Drug Discovery. Potential Applications to New Therapies for Chronic Pain. *J. Med. Chem.* **2017**, *60*, 66–88.

(6) Hefti, F. F.; Rosenthal, A.; Walicic, P. A.; Wyatt, S.; Vergara, G.; Shelton, D. L.; Davies, A. M. Novel Class of Pain Drugs Based on Antagonism of NGF. *Trends Pharmacol. Sci.* **2006**, *27*, 85–91.

(7) Mantyh, P. W.; Koltzenburg, M.; Mendell, L. M.; Tive, L.; Shelton, D. L. Antagonism of Nerve Growth Factor-TrkA Signaling and the Relief of Pain. *Anesthesiology* **2011**, *115*, 189–204.

(8) 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*, 3346–3356.

(9) 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., III Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000**, *33*, 889–897.

(10) Coveney, P. V.; Wan, S. On the Calculation of Equilibrium Thermodynamic Properties from Molecular Dynamics. *Phys. Chem. Chem. Phys.* **2016**, *18*, 30236–30240.

(11) Miller, B. R., 3rd; McGee, T. D., Jr.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. Mmpbsa.py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321.

(12) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.

(13) Bhati, A. P.; Wan, S.; Wright, D. W.; Coveney, P. V. Rapid, Accurate, Precise, and Reliable Relative Free Energy Prediction Using Ensemble Based Thermodynamic Integration. *J. Chem. Theory Comput.* **2017**, *13*, 210–222.

(14) 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*, 1228–1241.

(15) Swanson, J. M.; Henchman, R. H.; McCammon, J. A. Revisiting Free Energy Calculations: A Theoretical Connection to MM/PBSA and Direct Calculation of the Association Free Energy. *Biophys. J.* **2004**, *86*, 67–74.

(16) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. Alchemical Free Energy Methods for Drug Discovery: Progress and Challenges. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150–160.

(17) 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.

(18) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. J.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; 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.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*; Gaussian, Inc.: Wallingford, CT, 2004.

(19) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(20) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber FF99SB Protein Force Field. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1950–1958.

(21) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(22) Highfield, R. Supercomputer bid to create the first truly personalised medicine. <https://blog.sciencemuseum.org.uk/supercomputer-bid-to-create-the-first-truly-personalised-medicine/> (accessed Feb 22, 2017).

(23) Fiser, A.; Sali, A. Modloop: Automated Modeling of Loops in Protein Structures. *Bioinformatics* **2003**, *19*, 2500–2501.

(24) Allen, W. J.; Balias, T. E.; Mukherjee, S.; Brozell, S. R.; Moustakas, D. T.; Lang, P. T.; Case, D. A.; Kuntz, I. D.; Rizzo, R. C. Dock 6: Impact of New Features and Current Docking Performance. *J. Comput. Chem.* **2015**, *36*, 1132–1156.

(25) Chodera, J. D.; Mobley, D. L. Entropy-Enthalpy Compensation: Role and Ramifications in Biomolecular Ligand Recognition and Design. *Annu. Rev. Biophys.* **2013**, *42*, 121–142.

(26) Bissantz, C.; Folkers, G.; Rognan, D. Protein-Based Virtual Screening of Chemical Databases. 1. Evaluation of Different Docking/Scoring Combinations. *J. Med. Chem.* **2000**, *43*, 4759–4767.

(27) Gohlke, H.; Kiel, C.; Case, D. A. Insights into Protein-Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras-Raf and Ras-RaiGDS Complexes. *J. Mol. Biol.* **2003**, *330*, 891–913.

(28) Bunney, T. D.; Wan, S.; Thiyagarajan, N.; Sutto, L.; Williams, S. V.; Ashford, P.; Koss, H.; Knowles, M. A.; Gervasio, F. L.; 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*, 194–204.

(29) Andrews, M. D.; Bagal, S. K.; Gibson, K. R.; Omoto, K.; Ryckmans, T.; Skerratt, S. E.; Stuppel, P. A. Pyrrolo[2,3-*d*]pyrimidine Derivatives as Inhibitors of Tropomyosin-Related Kinases and Their Preparation and Use in the Treatment of Pain. WO2012137089A1, 2012.

(30) Xing, L.; Klug-Mcleod, J.; Rai, B.; Lunney, E. A. Kinase Hinge Binding Scaffolds and Their Hydrogen Bond Patterns. *Bioorg. Med. Chem.* **2015**, *23*, 6520–6527.

(31) Morphy, R. Selectively Nonselective Kinase Inhibition: Striking the Right Balance. *J. Med. Chem.* **2010**, *53*, 1413–1437.

(32) Ou-Yang, S. S.; Lu, J. Y.; Kong, X. Q.; Liang, Z. J.; Luo, C.; Jiang, H. Computational Drug Discovery. *Acta Pharmacol. Sin.* **2012**, *33*, 1131–1140.

(33) Wan, S.; Bhati, A. P.; Zasada, S. J.; Wall, I.; Green, D.; Bamborough, P.; Coveney, P. V. Rapid and Reliable Binding Affinity Prediction of Bromodomain Inhibitors: A Computational Study. *J. Chem. Theory Comput.* **2017**, *13*, 784–795.