

RESEARCH ARTICLE

# Uncovering Molecular Bases Underlying Bone Morphogenetic Protein Receptor Inhibitor Selectivity

Abdelaziz Alsamarah<sup>1</sup>, Alecander E. LaCuran<sup>1</sup>, Peter Oelschlaeger<sup>1</sup>, Jijun Hao<sup>2\*</sup>, Yun Luo<sup>1\*</sup>

**1** Department of Pharmaceutical Sciences, College of Pharmacy, Western University of Health Sciences, Pomona, California, United States of America, **2** College of Veterinary Medicine, Western University of Health Sciences, Pomona, California, United States of America

\* [luoy@westernu.edu](mailto:luoy@westernu.edu) (YL); [jhao@westernu.edu](mailto:jhao@westernu.edu) (JH)



OPEN ACCESS

**Citation:** Alsamarah A, LaCuran AE, Oelschlaeger P, Hao J, Luo Y (2015) Uncovering Molecular Bases Underlying Bone Morphogenetic Protein Receptor Inhibitor Selectivity. PLoS ONE 10(7): e0132221. doi:10.1371/journal.pone.0132221

**Editor:** Jie Zheng, University of Akron, UNITED STATES

**Received:** April 15, 2015

**Accepted:** June 11, 2015

**Published:** July 2, 2015

**Copyright:** © 2015 Alsamarah et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All data underlying the findings in this study are freely available in the paper and its Supporting Information files.

**Funding:** The computations were made possible by the Director's Discretionary award at Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. This work was supported by the seed funds of Western University of Health Sciences for Yun Luo and Jijun Hao. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Abnormal alteration of bone morphogenetic protein (BMP) signaling is implicated in many types of diseases including cancer and heterotopic ossifications. Hence, small molecules targeting BMP type I receptors (BMPRI) to interrupt BMP signaling are believed to be an effective approach to treat these diseases. However, lack of understanding of the molecular determinants responsible for the binding selectivity of current BMP inhibitors has been a big hindrance to the development of BMP inhibitors for clinical use. To address this issue, we carried out *in silico* experiments to test whether computational methods can reproduce and explain the high selectivity of a small molecule BMP inhibitor DMH1 on BMPRI kinase ALK2 vs. the closely related TGF- $\beta$  type I receptor kinase ALK5 and vascular endothelial growth factor receptor type 2 (VEGFR2) tyrosine kinase. We found that, while the rigid docking method used here gave nearly identical binding affinity scores among the three kinases; free energy perturbation coupled with Hamiltonian replica-exchange molecular dynamics (FEP/H-REMD) simulations reproduced the absolute binding free energies in excellent agreement with experimental data. Furthermore, the binding poses identified by FEP/H-REMD led to a quantitative analysis of physical/chemical determinants governing DMH1 selectivity. The current work illustrates that small changes in the binding site residue type (e.g. pre-hinge region in ALK2 vs. ALK5) or side chain orientation (e.g. Tyr219 in caALK2 vs. wtALK2), as well as a subtle structural modification on the ligand (e.g. DMH1 vs. LDN193189) will cause distinct binding profiles and selectivity among BMP inhibitors. Therefore, the current computational approach represents a new way of investigating BMP inhibitors. Our results provide critical information for designing exclusively selective BMP inhibitors for the development of effective pharmacotherapy for diseases caused by aberrant BMP signaling.

**Competing Interests:** The authors have declared that no competing interests exist.

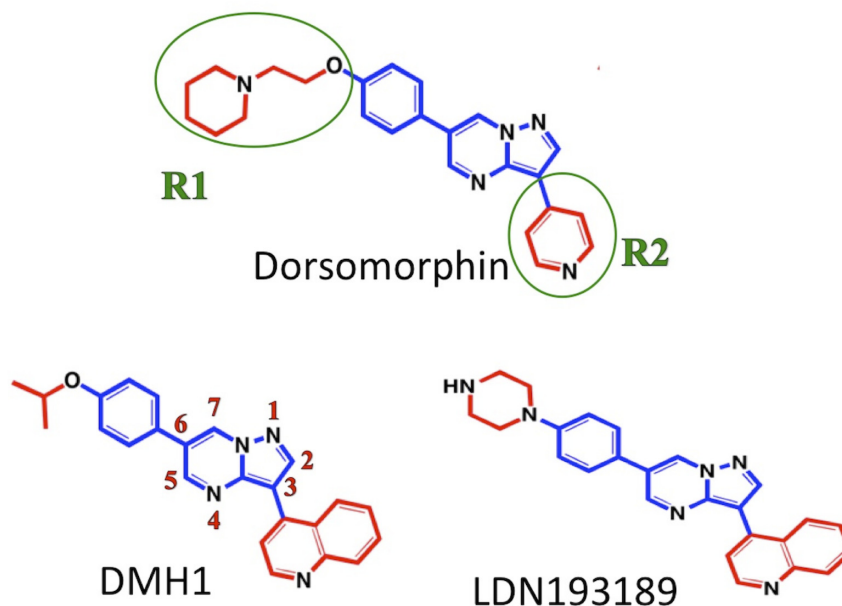
## Introduction

The bone morphogenetic proteins (BMPs), a subgroup of the transforming growth factor- $\beta$  (TGF- $\beta$ ) superfamily, play critical and diverse roles in cellular processes [1]. The biological activities of BMPs are mediated through formation of heteromeric BMP receptor complexes consisting of two type I receptors and two type II receptors. When BMPs bind to the extracellular part of the receptor complex, the type I BMP receptors are activated and their intracellular kinase domain then phosphorylates R-Smads protein family to trigger downstream gene transcription [2]. A series of serine/threonine kinases, termed activin receptor-like kinases (ALK), have been identified to constitute BMP type I receptors (BMPRI): ALK1, ALK2, ALK3, and ALK6.

Aberrant activation of BMP signaling is involved in numerous diseases and targeting BMPRI is believed to be an effective therapeutic approach for treating these diseases. For instance, mutation R206H in ALK2, which constitutively activates BMP signaling in the absence of BMP ligands, is responsible for ~97% of patients with fibrodysplasia ossificans progressiva (FOP) disorder, one of the most devastating and rare bone diseases [3, 4]. Thus small molecular ALK2 inhibitors, which may be effective therapeutic agents against FOP, have been highly sought after. In addition, abundant expression of ALK1 was found in the vasculature of many types of tumors, but weak or no expression of ALK1 was detected in tumor cells and normal tissues, suggesting that ALK1 inhibition may be a potential therapeutic approach complementary to the current anti-angiogenic modalities in the clinic [5]. Similarly, ALK3 and ALK6 are also implicated in other distinct diseases [6–8]. Therefore, development of selective small molecule inhibitors of each subtype of BMPRI to block BMP signaling may represent an effective therapeutic approach to treat these different types of disease.

Recently, significant efforts have been made to develop small molecule ALK2 inhibitors to interrupt abnormal activation of BMP signaling. Dorsomorphin (Fig 1), the first small molecule BMPRI inhibitor, was identified in a screen for compounds that perturb the zebrafish embryonic dorsoventral axis [9]. Although dorsomorphin inhibits ALK2 activity by binding to the ATP-binding pocket of the ALK2 Ser/Thr kinase domain [10], it displays significant “off-target” inhibition of the vascular endothelial growth factor receptor type 2 (VEGFR-2) tyrosine kinase and other BMP type I receptors [11–13]. Over the past several years, a series of dorsomorphin analogs with pyrazolo[1,5-*a*]pyrimidine or aminopyridine scaffold have been developed to improve compound selectivity towards ALK2 [13–20] (S1 Table). For instance, DMH1 [13] was developed with higher selectivity towards BMP type I receptors vs. TGF- $\beta$ /Activin pathway receptor ALK5 and VEGFR2 than dorsomorphin. Other derivatives such as LDN-193189 [16, 21], exhibited higher potency against BMP type I receptors but less selectivity against ALK5 and VEGFR2 than DMH1 (Fig 1 and S1 Table). Despite the continuous efforts in chemical synthesis in recent years, it remains unclear how these BMP inhibitors can discriminate one receptor over others. A consensus has emerged that understanding the selectivity mechanisms is critical for designing exclusively selective inhibitors for each subtype of BMPRI that are urgently needed today.

To address this question, we apply all-atom molecular dynamics-based free energy calculations to investigate the physicochemical contributions underlying BMP inhibitors' binding characteristics, which are often difficult to obtain from ligand-based structure-activity relationship (SAR) analysis or static crystal structures. The main computational approach applied here is free energy perturbation coupled with Hamiltonian replica-exchange molecular dynamics (FEP/H-REMD) simulations. The FEP/H-REMD approach has recently provided a wealth of molecular details on the energetic determinants of the binding affinity in tyrosine kinases [22–25]. We have chosen DMH1 as a model compound with the aim of capturing the origin of its



**Fig 1. Structures of dorsomorphin and its analogs.** The common scaffold is shown in blue, R1 and R2 groups are shown in red.

doi:10.1371/journal.pone.0132221.g001

excellent selectivity towards ALK2 vs. the structurally closely related ALK5 and VEGFR2 kinases.

For ALK2 kinase, two crystal structures were used for this study. One is from the wild-type ALK2-dorsomorphin complex (PDB ID: 3H9R [10]), denoted as wtALK2. The other is that of the Q207D mutant ALK2-LDN193189 complex (PDB ID: 3Q4U [20]). The mutation Q207D is located at the GS domain (glycine-serine rich region) on top of the kinase N-lobe (N-terminal lobe) motif (S1A Fig). Q207D ALK2 has been reported as constitutively active ALK2 (caALK2); it leads to ectopic endochondral bone formation in a mouse model [16]. ALK2 and ALK5 are highly similar (S1B Fig for structural alignment result), and their kinase domains contain a conserved sequence of three amino acids (Asp-Leu-Gly) known as DLG-motif at the beginning of the activation loop (A-loop), while VEGFR2 tyrosine kinase contains a DFG motif (Asp-Phe-Gly). In general, the crystal structures of ALK2 and ALK5 with inhibitors all represent DLG-in like conformation (S1C Fig) [10, 20, 26], while in VEGFR2, there are various inhibitors bound to either the DFG-in or DFG-out conformation (sometimes denoted as type-I and type-II inhibitors). Therefore, we have chosen crystal structures of both DFG-in and DFG-out conformations of VEGFR2 (indicated as VEGFR2-in and VEGFR2-out) as separate topologies for DMH1 binding free energy calculation.

The results show that, while molecular docking method used here gave nearly identical scores among the three kinases, FEP/H-REMD simulations successfully reproduced that DMH1 only binds to ALK2, but not to ALK5, VEGFR2-in or VEGFR2-out. The binding free energies are in good agreement with experimental measurements (Table 1). The FEP/H-REMD identified the potential binding poses of DMH1, which led to the quantitative analysis of the origin of DMH1 selectivity for these kinases. Our calculations indicate that DMH1 selectivity originates from a favorable electrostatic interaction between DMH1 and the ATP-binding pocket of ALK2. This interaction is absent in ALK5 and VEGFR2 because of subtle binding pose changes. Confirming our computational predictions, we further elucidate that the compound LDN193189 has more favorable interaction with ALK5 than DMH1, which is

**Table 1. Calculated binding free energy  $\Delta G_b^\circ$  of DHM1 to different kinases, docking scores and results from experimental kinase assays in kcal/mol.**

Protein	Expt <sup>a</sup>	Docking score <sup>b</sup>	$\Delta G_b^\circ$ <sup>c</sup> from FEP/HREMD	Decomposition <sup>d</sup> of $\Delta G_b^\circ$			
				repu	disp	elec	rst
wtALK2	-9.6	-8.9 ± 0.1	-8.5 ± 0.6	12.19	-17.65	-3.16	0.08
caALK2	-9.6	-9.6 ± 0.1	-6.2 ± 0.4	14.41	-18.93	-1.73	0.02
ALK5 pose3*	>-5.5	-9.1 ± 0.1	-0.4 ± 0.7	7.48	-15.86	7.42	0.55
ALK5 pose4	>-5.5	-8.7 ± 0.1	-0.9 ± 0.7	9.76	-16.16	4.78	0.68
VEGFR2-in	>-5.5	-8.6 ± 0.2	-3.4 ± 0.6	10.29	-18.91	4.63	0.60
VEGFR2-out	>-5.5	-9.7 ± 0.1	-1.7 ± 0.6	9.67	-15.54	3.66	0.49

<sup>a</sup>. The experimental binding free energies were estimated from IC<sub>50</sub> of kinase assay [14] using  $\Delta G = -RT \ln(C^\circ/K_d)$ , in which  $R$  is the gas constant  $1.987 \times 10^{-3}$  kcal/K/mol,  $C^\circ$  is the standard reference concentration 1 mol/L, and  $T$  is 300 K.

<sup>b</sup>. The docking score is a cluster average value. Errors in docking scores are standard deviations from all binding poses within a single cluster.

<sup>c</sup>. The average and the standard deviations in  $\Delta G_b^\circ$  are calculated from the last five simulations of 400 ps per replica with different initial velocities.

<sup>d</sup>. The free energy decomposition gives repulsive, dispersive, electrostatic, and distance restraint contributions in the total  $\Delta G_b^\circ$ .

\* The pose 3\* in Fig 2.

doi:10.1371/journal.pone.0132221.t001

consistent with previous experimental reports [14, 19]. Our computational study highlights the importance of structural dynamics and demonstrates that the FEP/H-REMD approach can serve as a robust method to explain and predict binding selectivities of BMP inhibitors among highly conserved ATP binding sites. The molecular mechanism illustrated here provides critical information for future rational design of exclusively selective and potent inhibitors for each subtype of BMPRI.

## Materials and Methods

### Atomic model preparation

The initial structures of ALK2, ALK5 and VEGFR2 for simulations were taken from the crystal structures of the protein-inhibitor complexes (PDB ID: 3H9R for wtALK2 [10], PDB ID: 3Q4U for caALK2 [20], PDB ID: 3TZM for ALK5 [26], PDB ID: 3VO3 for VEGFR2 DFG-out [27], PDB ID: 3CJG for VEGFR2 DFG-in [28]) of *Homo sapiens* species. In the wtALK2 complex, part of the A-loop (residues 362 to 374), and the  $\beta$ -turn between  $\beta$ 4 and  $\beta$ 5 (residues 273 to 275) were not present in the crystal structure. To address this issue, the missing A-loop portion in wtALK2 was transplanted from the crystal structure of the constitutively active Q207D mutant ALK2 (caALK2). The three missing residues in the  $\beta$ -turn were patched using the PATCH command in CHARMM program [29, 30]. Then these patched residues underwent energy minimization with the rest of the protein fixed to optimize the conformation. The pKa calculations using PROPKA GUI [31] plugin in VMD [32] indicate that the ionization states of protein residues remain the same as that of the individual residues at physiological pH. All the crystal water molecules were kept unchanged. CHARMM-GUI [33] was used to read in the PDB files and solvate each system in a rectangular water box (94 Å × 94 Å × 76 Å). Since potassium and chloride ions are the two major cytosolic ions, each system was neutralized with K<sup>+</sup> and Cl<sup>-</sup> ions at a physiological salt concentration of 150 mM. The solvated DMH1 complexes with wtALK2, caALK2, ALK5, VEGFR2 DFG-in and VEGFR2 DFG-out consist of 53747, 53706, 68303, 67950 and 53824 atoms, respectively. All simulations employed the all-atom CHARMM C36 force field [34–36] for proteins and ions, and the TIP3P force field [37] for water. In addition, the missing partial P-loop (residues 843 to 846) and the partial A-loop

(residues 1052 to 1065) in the crystal structure of VEGFR2 DFG-in were patched using the CHARMM PATCH command. Likewise, in ALK5, the A-loop residues 370 and 371 were patched using CHARMM. The patched residues were subjected to 500 steps of energy minimization using the steepest descent method [38], followed by 500 steps of minimization using the adopted-basis Newton-Raphson method [38], with the remaining parts of the protein held fixed using CHARMM.

Small ligands were first prepared and minimized using the ArgusLab program [39]. DMH1 is expected to be neutral in bulk solution. The unsubstituted N atom on the piperazine ring of LDN193189 is solvent-exposed in the binding site and is expected to be protonated in a physiological pH aqueous solution. Small ligand force field parameters were generated using the General Automated Atomic Model Parameterization (GAAMP) web server [40]. For DMH1 in bulk solution, the center-of-mass of the ligand was placed at the center of a cubic water box with a side length of 90 Å, resulting in a total of 49102 atoms. The solvated system was equilibrated for 1 nanosecond (ns) before being submitted to solvation free energy calculations.

## Docking and Solvation

In order to determine the potential binding poses of DMH1 in different kinases, the ligand was docked into the ATP binding site of each minimized crystal structure using the flexible ligand docking protocol in Autodock4.2 [41]. Using AutoGrid [42], the grid box was set to 70, 80, and 70 grid points, along the x, y and z-axis, with 0.375 Å grid spacing, centered on the ATP binding site. For conformational search, docking calculations were carried out using the Lamarckian genetic algorithm and default parameters. The docking protocol was first tested by docking dorsomorphin back into the ALK2 crystal structure and showed perfect alignment with the ALK2-dorsomorphin crystal structure (PDB ID: 3H9R). For DMH1 docking, the top 100 poses were sorted by their docking scores and clustered by root mean squared deviation (RMSD) of ligand heavy atoms that differ from each other within 2 Å.

The top ranked docking pose of each cluster was solvated in 150 mM KCl aqueous solution using CHARMM-GUI, and the molecular dynamics equilibrium (see simulation protocol below) was set to relax the atomic system by releasing the harmonic constraints (force constant 50 kcal/mol/Å<sup>2</sup>) stepwise (every 200 ps) on water and ion molecules, protein side chains, protein backbone, and eventually the ligand. At least 40 ns of equilibration were carried out for each system without constraint before submitting for free energy calculation. RMSD of protein backbone and ligand heavy atoms as well as center-of-mass distance between protein and ligand were monitored to obtain stable binding poses (S2 Fig).

## Simulations Protocol

All the simulations were performed with NAMD2.9b [43] using periodic boundary conditions at constant temperature and pressure (NPT ensemble) of 300 K and 1 atm using Langevin thermostat and Andersen-Hoover barostat. Long-range electrostatics interactions were treated using the particle-mesh Ewald (PME) method [44]. The non-bonded interaction list was updated on every integration step using a cutoff of 13.5 Å. A smoothing function is applied to both electrostatics and van der Waals forces between 10 Å and 12 Å. The dynamics were propagated using Langevin dynamics with Langevin damping coefficient of 1 ps<sup>-1</sup> and a time step of 2 fs. The SHAKE algorithm [45] was applied to all hydrogen atoms.

## Absolute Binding Free Energy Calculations

The binding of a ligand to a protein in aqueous solution can be described by the top part of the thermodynamic cycle in (S3 Fig). At equilibrium, the dissociation constant of the binding

reaction is defined by  $K_d = \frac{k_{off}}{k_{on}} = \frac{[Protein][Ligand]}{[Protein* Ligand]}$ , which has the unit of concentration. Under the constant pressure and constant temperature ensemble (NPT), the Gibbs free energy of binding  $\Delta G_{bind}$ , i.e., the binding affinity, between receptor and ligand in solution is related to the dissociation constant  $K_d$  by  $\Delta G = -RT \ln(C^\circ/K_d)$ , in which  $R$  is the gas constant  $1.987 \times 10^{-3}$  kcal/K/mol,  $C^\circ$  is the standard reference concentration 1 mol/L or 1 molecule/1660 Å<sup>3</sup>, and  $T$  is the absolute temperature in Kelvin. At room temperature, a nanomolar range dissociation constant corresponds to a binding free energy  $\Delta G_{bind}$  range from -8 to -12 kcal/mol.

The free energy of binding can be estimated, in principle, from a long molecular dynamics trajectory, as long as the binding and unbinding events have occurred many times so as to give an accurate thermodynamic average. In practice, this brute-force approach is often hindered by the current computational limitations. Since the free energy is a function of state, the Free Energy Perturbation (FEP) approach [46, 47] can be used instead. In FEP, the bound and unbound states are connected through an arbitrary path by perturbing the Hamiltonian of the system in a series of alchemical steps. To calculate the absolute binding free energy using FEP, the double decoupling protocol developed by Deng and Roux [48] is applied. As illustrated in (S3 Fig), the total binding free energy is calculated from two separate steps: (1) a bound ligand is decoupled from its environment, i.e., the electrostatics and Lennard-Jones interactions between ligand and its environment (protein and solvent) are turned off gradually ( $\Delta G_{decouple}$  step). This step corresponds to the free energy cost of moving a bound ligand from binding site into vacuum. (2) In order to close the thermodynamic cycle, the solvation free energy of ligand alone, i.e., the free energy cost of moving the ligand from vacuum to aqueous solution is calculated by decoupling the electrostatics and Lennard-Jones interactions between ligand and water molecules ( $\Delta G_{solvation}$  step). The sum of the free energies of step (1) and step (2) equals the negative standard free energy of binding ( $\Delta G_{bind} = -\Delta G_{decouple} - \Delta G_{solvation}$ ) (S3 Fig).

The decoupling of a ligand with its environment (protein+solvent or solvent alone) was carried out using a scaling factor  $\lambda$  between 0 and 1, and a number of intermediate states.  $\lambda = 1$  represents a fully interacting (bound) state, and  $\lambda = 0$  represents a fully decoupled state (ligand in vacuum). The Weeks-Chandler-Andersen decoupling scheme [49] was utilized to separate the Lennard-Jones 6–12 potential into repulsive and dispersive parts. The decoupling of the ligand from the binding pockets was carried out as a stepwise reversible process, staged by the three thermodynamic coupling parameters representing repulsive, dispersive and electrostatic interactions of the ligand with its environment. Each coupling parameter has a different number of intermediate states between 0 and 1 using 128  $\lambda$  in total (72  $\lambda_{dispersive}$ , 24  $\lambda_{repulsive}$  and 32  $\lambda_{electrostatic}$ ) or 64  $\lambda$  in total (36  $\lambda_{dispersive}$ , 12  $\lambda_{repulsive}$  and 16  $\lambda_{electrostatic}$ ). In addition, a harmonic distance restraint between the center of mass of the ligand and center of mass of the protein is introduced using a force constant of 10 kcal/mol/Å<sup>2</sup> to confine the sampling volume of the decoupled ligand. The free energy cost of introducing the distance restraint to the total binding free energy is calculated using equation  $\Delta \Delta G_t = -k_B T \ln(F_t C^\circ) - \Delta G_t^{site}$ , where  $F_t$  corresponds to numerical integrals over the restraining quadratic potential and  $C^\circ$  is the standard reference concentration.  $\Delta G_t^{site}$  is the free energy cost of introducing the distance restraint when the ligand is fully coupled with the receptor.  $\Delta G_t^{site}$  is calculated by decreasing the force constant stepwise (13 windows and 200 ps per window) from its original value to zero using the thermodynamic integration framework within the *colvars* module in NAMD2.9b. The contributions of the distance restraint are quite small (see Table 1), which indicates that all the reference distances are taken from well-equilibrated binding poses.

To further improve Boltzmann sampling, the Hamiltonian replica-exchange algorithm was coupled to FEP (FEP/H-REMD) [50–52]. The FEP/H-REMD method has been recently used to predict binding affinities of small molecular ligands to tyrosine kinases [23], oligosaccharide

binding to glycoside hydrolase [53], as well as the binding of flexible decapeptide to SH3 domain of Abl kinase [25]. In this method, each  $\lambda$  value represents one replica of the system and frequent attempted  $\lambda$ -swapping ( $0.2 \text{ ps}^{-1}$ ) between neighboring replicas was performed during the simulation. The attempted exchange was either rejected or accepted according to the Metropolis Monte Carlo criterion to ensure the detailed balance condition [54]. The average acceptance ratio among 128 replicas was  $> 80\%$ . During the post-processing, the shuffled trajectories due to  $\lambda$ -swapping were un-shuffled for each  $\lambda$  value using the *sortreplicas* program in NAMD2.9b. The shuffled energy outputs were sorted for the final WHAM [55] post-processing. The insertion of the ligand into the bulk phases was calculated with the same protocol using 64 replicas in total. For each FEP/H-REMD job, all replicas were running in parallel mode on 512 or 1024 compute nodes on the IBM Blue Gene/Q supercomputer Mira at Argonne National Laboratory.

## Statistical Analysis

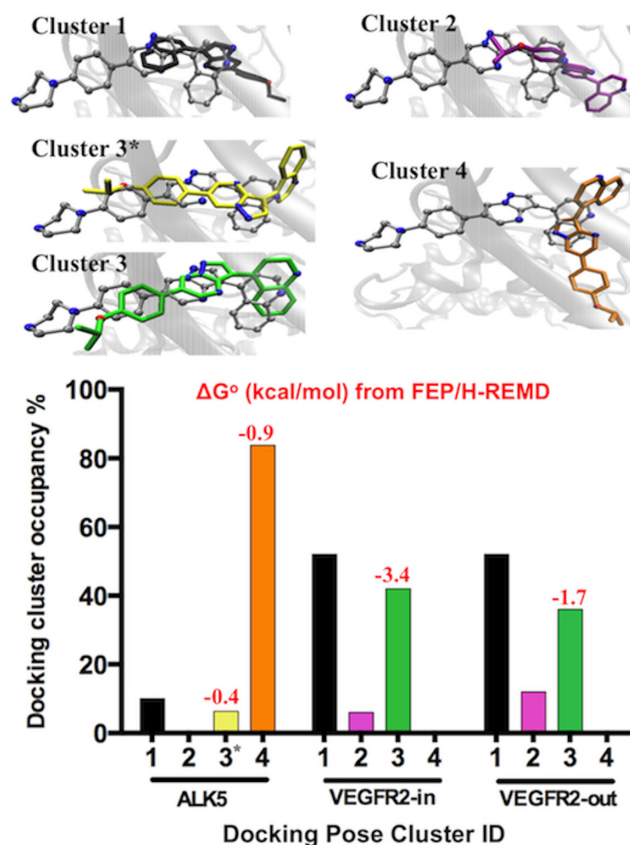
Paired T-tests for the root of mean standard fluctuation (RMSF) in the A-loop and P-loop were used to quantify the magnitude of the change between the apo (unbound) and DMH1 bound states utilizing Prism6 software (GraphPad Software Inc., La Jolla, CA). The individual amino acids in each loop were paired, for example: D354 in the apo A-loop is paired to D354 in the DMH1 bound A-loop. All statistically different results are denoted by one or more asterisks; \* =  $p < 0.05$ , \*\* =  $p < 0.01$ , and \*\*\* =  $p < 0.001$ .

## Results and Discussion

### Multiple docking poses

The top ranked binding pose of ALK2-DMH1 is in excellent agreement with crystal structures of the ALK2-dorsomorphin and ALK2-LDN193189 complexes. However, the docking scores failed to distinguish DMH1 as a binder vs. non-binder between ALK2, ALK5 and VEGFR2 (Table 1). The flexible ligand docking offered three clusters of binding poses in both ALK5 and VEGFR2-in/out with nearly identical docking scores as ALK2-DMH1 (Fig 2). These docking poses cannot be experimentally confirmed since there is no crystal structure of these DMH1 analogs binding to ALK5 or VEGFR2. Multiple binding poses are also more likely to occur in the case of weak binding. Therefore, the FEP/H-REMD simulations were carried out for the entire top ranked DMH1 binding poses from each cluster identified by AutoDock4.2 regardless of the binding scores. The free energy calculations revealed that some docking poses were false positives (positive binding free energy). Therefore we eliminated those poses. When there were multiple poses that gave negative binding free energies from FEP/H-REMD, we combined the binding free energies using equation (1)  $\Delta G^\circ = -k_B T \ln \left[ \sum_{i \neq j} \exp \left( -\frac{\Delta G_{i \neq j}^\circ}{k_B T} \right) \right]$ , similarly to the strategy applied in relative free energy calculations [56, 57].

**FEP/H-REMD results show only one out of three docking poses in VEGFR2 has negative binding free energy.** The docking binding poses of DMH1 in VEGFR2-out were clustered into three groups (Fig 2). The FEP/H-REMD simulations yielded positive binding free energies of +2.0 and +1.7 kcal/mol for poses (1) and (2), respectively. Only pose (3) gave a small negative binding free energy of -1.7 kcal/mol. This is consistent with the fact that only cluster (3) has binding features similar to the ALK2-LDN193189 crystal structure (Fig 2 top). The small negative binding free energy agrees with experimental kinase selectivity data [14]. Same as in VEGFR2-in, only the ALK2-LDN193189 like binding pose (3) gave a negative binding free energy of -3.4 kcal/mol. Therefore, for both systems, we used pose (3) for further analysis.



**Fig 2. Clustered DMH1 docking poses in ALK5, VEGFR2-in, and VEGFR2-out proteins.** Each docking cluster represents a group of ligand poses with ligand RMSD < 2 Å. **Top:** Representative docking poses in each cluster superimposed on LDN193189 (gray) in the ALK2 x-ray structure. **Bottom:** The occupancy of each docking cluster, with FEP/H-REMD binding free energy shown in red on top of each histogram (positive values are not shown here). The color of the ligand-binding pose corresponds to those in the bar graphs in the bottom.

doi:10.1371/journal.pone.0132221.g002

**FEP/H-REMD results show two docking poses in ALK5 have negative binding free energy.** The dominant binding pose cluster (4) is unique in ALK5, characterized by the quinoline ring overlapping with the pyrazolo-pyrimidine moiety of LDN193189 in the crystal structure, and the rest of the molecule projects away from the ATP site (Fig 2). The DMH1 binding pose (3\*) in ALK5 is similar to the pose (3) in VEGFR2-in/out, but shifted into a hydrophobic region adjacent to the ATP binding site with a different rotation of the pyrazolo[1,5-a]pyrimidine and quinoline rings. For FEP/H-REMD calculations, we excluded pose (1) since it was shown to be a false positive binding pose in VEGFR2. Both poses (3\*) and (4) turned out to be weak binding poses with binding free energies of -0.4 and -0.9 kcal/mol. Combining the two negative binding free energies using equation (1) yielded the binding free energy of -1.1 kcal/mol, which agrees with the experimental result that DMH1 is a weak binder for ALK5 [14].

### Absolute binding free energies and decomposed contributions

The absolute binding free energy of wtALK2-DMH1 complex was computed from a total of 384 ns FEP/H-REMD sampling (6 ns per replica, 64 replicas). caALK2-DMH1 and VEGFR2-out-DMH1 complexes were sampled for a total of 358.4 ns each. The evolution of the calculated free energy over the sampling length indicates the progress towards convergence



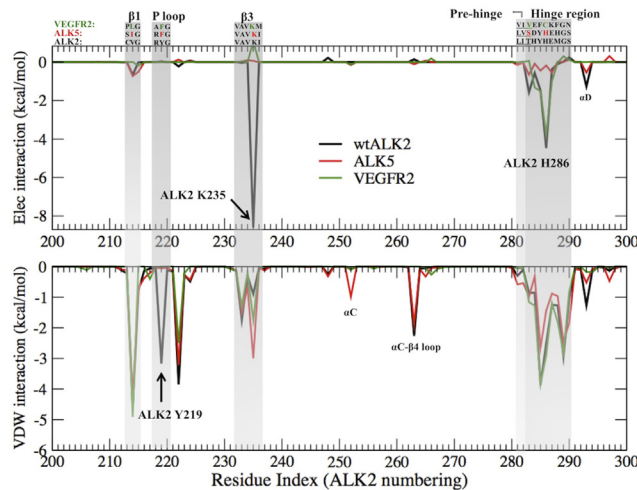
(S4A Fig). The progression of the free energy components (repulsive, dispersive and electrostatic) with respect to the coupling parameters ( $\lambda_{\text{rep}}$ ,  $\lambda_{\text{dis}}$  and  $\lambda_{\text{elec}}$ ) for DMH1 in ALK2 *versus* DMH1 in water is shown in (S4 Fig). The ALK5-DMH1 and VEGFR2-in-DMH1 systems were sampled only for 256 ns, as the evolution of the calculated free energies clearly shows that the binding affinities were already less negative than -5 kcal/mol. The data collected from the last 128 ns were used to compute the block average using a block size of 400 ps and standard error of the free energies of binding for DMH1 to ALK2s, yielding  $-8.5 \pm 0.6$  kcal/mol in wtALK2 and  $-6.2 \pm 0.4$  kcal/mol in caALK2. The data collected from the last 76.8 ns were used to compute free energies for the other systems, yielding  $-0.4 \pm 0.7$  kcal/mol and  $-0.9 \pm 0.7$  kcal/mol in ALK5 (see multiple docking poses section),  $-3.4 \pm 0.6$  kcal/mol in VEGFR-in, and  $-1.7 \pm 0.6$  kcal/mol in VEGFR-out (Table 1).

**Electrostatic contribution is a key molecular determinant responsible for the binding specificity.** The staged FEP/H-REMD strategy separates the interaction free energy of ligand with its surrounding into repulsive, dispersive and electrostatic components (Methods section). Although the absolute value of each decomposed free energy is path dependent, comparing the relative values between studied kinases offers useful insights into the binding mechanism. The positive repulsive contribution of the binding free energy in all proteins versus in bulk solution (Table 1) suggests that, in order to accommodate the bulky ligand DMH1, the binding pocket of all three kinases must undergo a certain amount of structural rearrangements, including certain numbers of water molecules expelled from the binding pocket and rearrangements of binding site residues. These rearrangements are associated with an unfavorable free energy penalty. The major favorable contribution of the binding affinity is the dispersive component. The negative dispersion contribution in protein relative to bulk solvent suggests that the protein binding site provides an environment with a higher density of van der Waals centers to stabilize DMH1 in the binding pocket. However, the dispersive component itself does not reflect the trend of the binding affinity among ALK2, ALK5 and VEGFR2, which indicates that the difference in van der Waals dispersive contribution is not sufficient to determine the binding specificity of DMH1 among the three kinases.

In contrast to the van der Waals dispersive contributions, which consistently favor the binding process, the contribution of electrostatic interactions is only favorable in ALK2 (Table 1). This indicates that the sum of favorable electrostatic interactions associated with hydrogen bonding and charge-charge (salt bridge) interactions established between DMH1 and ALK2 surpass the loss of water-DMH1 interactions in bulk solution. However, in ALK5 and VEGFR2, the sum of electrostatic interactions between ligand and binding site is not enough to compensate for the desolvation penalty of DMH1. Therefore, although the total binding free energy is dominated by the dispersive contribution, the electrostatic contribution is the key determinant responsible for the binding specificity of DMH1 to ALK2 over ALK5 and VEGFR2. In the following computational analysis section, we discuss what those key interactions are.

## Computational Analysis

**Per-residue interaction energy reveals key amino acids dominating ALK2 specificity.** The per-residue electrostatic interaction between DMH1 with ALK2, ALK5, and VEGFR2 shows clearly that the most favorable electrostatic interaction between DMH1 and ALK2 Lys235 is missing in ALK5 and VEGFR2 (Fig 3 top). It is consistent with the fact that a direct hydrogen bond between the Lys235 side chain and the N atom on the quinoline moiety of DMH1 was observed in 20% of the whole simulation time in ALK2, but in 0% with the corresponding Lys232 in ALK5 and Lys868 in VEGFR2. In addition, the favorable electrostatic



**Fig 3. Per-residue interaction with DMH1 in wtALK2 (black), ALK5 (red) and VEGFR2 (green), averaged from the last 2 ns of trajectory.** Top: electrostatic interaction. Bottom: van der Waals interaction. The three protein sequences are aligned by homology and the numbering is that of ALK2. The rest of the residues that have no interaction with DMH1 are not shown.

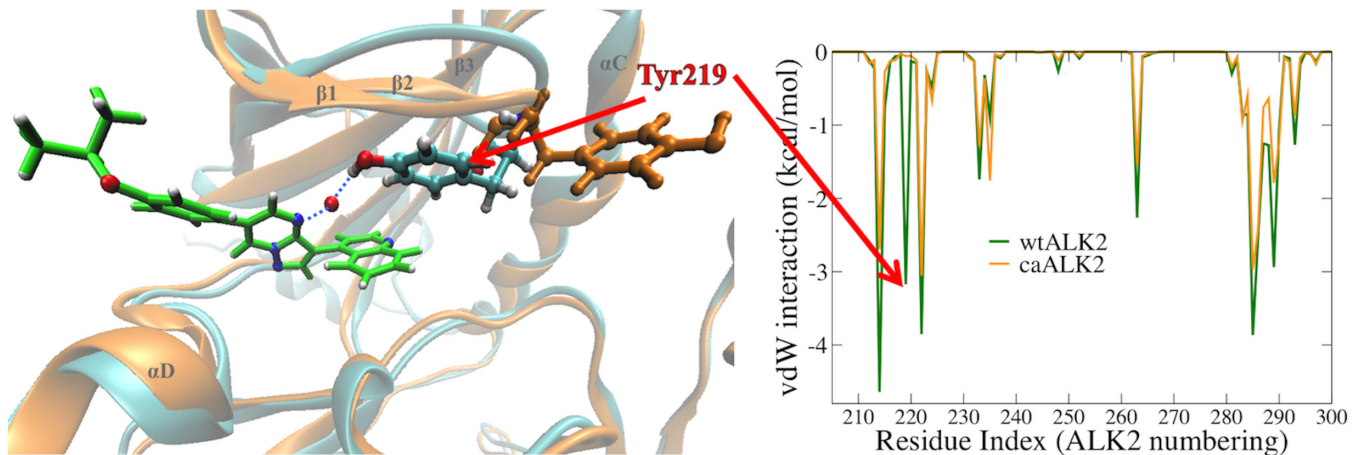
doi:10.1371/journal.pone.0132221.g003

interactions at the hinge region are much more significant in ALK2 than in ALK5. This is because at the ALK2 hinge region His286 forms a stable hydrogen bond with the N1 atom of the DMH1 pyrazole ring (Fig 1). This conserved hydrogen bond has been observed in the crystal structures of ALK2-dorsomorphin and ALK2-LDN193189 complexes [10, 20]. The van der Waals interactions are similar among ALK2, ALK5 and VEGFR2, except that the favorable van der Waals interaction between DMH1 and Tyr219 in ALK2 is absent in ALK5 and VEGFR2 (see discussion on Tyr219 below).

**P-loop Tyr219 in ALK2 significantly contributes to binding affinity and selectivity.**

During our investigation of ALK2, we first used the caALK2 crystal structure (PDB ID 3Q4U). Although the docking score (estimated binding free energy) of DMH1 agrees well with the experimental value of -9.6 kcal/mol, the FEP/H-REMD calculated binding free energy is -6.2 kcal/mol (Table 1). When we used the wtALK2 crystal structure (PDB ID 3H9R), the binding free energy is -8.5 kcal/mol. Interestingly, the binding poses of DMH1 in wtALK2 and caALK2 are nearly identical. The averaged van der Waals interactions between the binding site and DMH1 also show a similar pattern, except that a favorable P-loop (also termed phosphate-binding loop) Tyr219 peak in wtALK2 is completely missing in caALK2 (Fig 4). Dynamic trajectories showed Tyr219 in wtALK2 always pointing inwards (Tyr219-in), forming favorable hydrophobic contacts with the quinoline moiety of DMH1 and a water-mediated hydrogen bond with N4 on the pyrimidine moiety 25% of simulation time (Fig 4). In contrast, Tyr219 in caALK2 is pointing outwards away from the ligand throughout the simulation (Fig 5 left). In addition, the root mean square fluctuation (RMSF) of the P-loop indicated that its fluctuation did not change upon DMH1 binding to wtALK2; whereas, DMH1 binding to caALK2 increased RMSF (Fig 5 right). Thus, the differences in both the interaction energy and the P-loop fluctuation confirm that the favorable interaction between DMH1 and Tyr219 can only be established when Tyr219 is pointing inward as in the wtALK2.

Our structural analysis and PDB database search suggest that this conformational difference in the P-loop is unlikely to be caused by a single mutation in Q207D. In fact, another Q207D ALK2 crystal structure (PDB ID: 4BGG) has the same Tyr219-in conformation as in wtALK2. Different inhibitors in the binding site may also contribute to this difference in the Tyr219

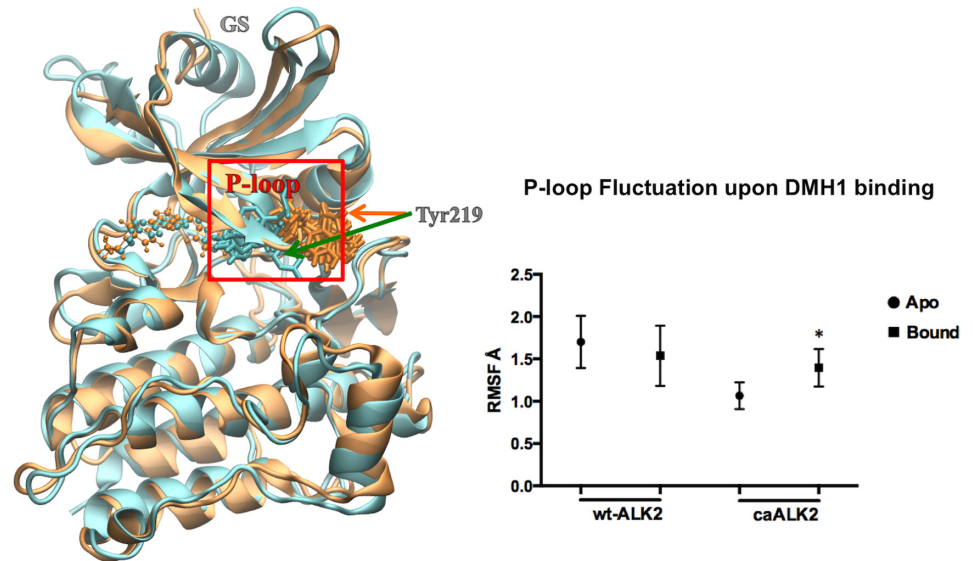


**Fig 4. Left:** Superposition of wtALK2 (cyan) and caALK2 (orange) backbone in cartoon. P-loop Tyr219 is shown in sticks. The ligand DMH1 is shown in light green. Tyr219 in caALK2 (orange) is pointing away from DMH1. Tyr219 in wtALK2 (cyan) is pointing towards the ligand and forms a water-mediated hydrogen bond with N4 on the pyrimidine moiety of DMH1. **Right:** van der Waals interaction energies in kcal/mol between each ALK2 residues and DMH1: wtALK2 in green, caALK2 in orange.

doi:10.1371/journal.pone.0132221.g004

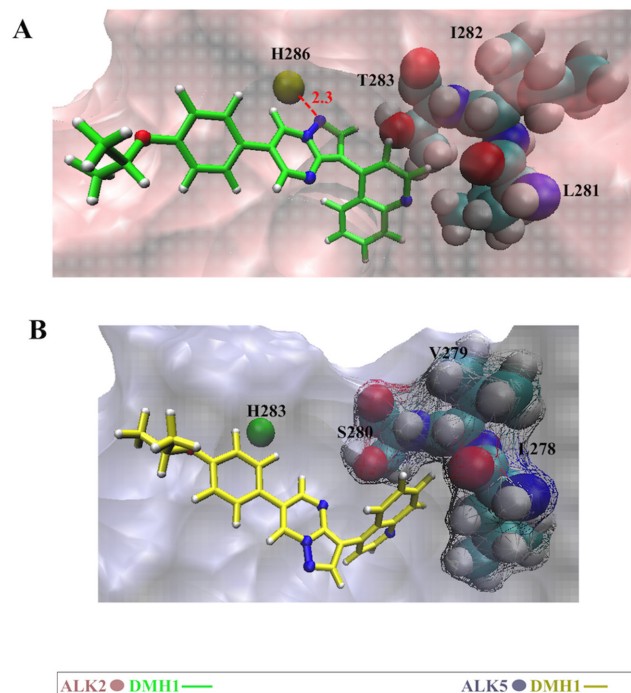
orientation among crystal structures. Recently there was a study about the effect of different inhibitors on the orientation of P-loop Phe632 in PRK1 kinase crystal structure (PDB ID: 4OTI) [58]. Other factors such as the presence of the GS domain on top of kinase N-lobe motif may also affect the P-loop conformation. Noticeably, four currently available crystal structures of Q207D caALK2 (PDB ID: 3Q4U, 4BGG [18], 3OOM and 3MTF [20]) all have the truncated GS domain, which may remove potential inhibitory interactions in the crystal structure of the protein, thus change the P-loop conformation. The caALK2 structure used here may represent a metastable state of ALK2 inactive conformation. In conclusion, the current analysis reveals quantitatively the crucial impact of the P-loop Tyr219 residue on the inhibitor binding affinity and selectivity, which indeed deserves special attention during structure-based ALK2 inhibitor design.

**The difference in the pre-hinge region affects the DMH1 binding pose in ALK2 and ALK5.** As previously described, the docking pose of DMH1 in ALK2 shares similar features with the ALK2-LDN193189 complex crystal structure, in which the pyrazolo[1,5-a]pyrimidine moiety of DMH1 faces the hinge region and forms a direct hydrogen bond with His286 in ALK2 (Fig 6A). This hydrogen bond is also seen in other crystal structures of BMPRI kinases, ALK6 (PDB ID: 3MDY) and ALK1 (PDB ID: 3MY0), in complex with LDN193189. However this favorable electrostatic interaction is absent between DMH1 and ALK5 (Fig 3). In search of a possible explanation, we compared the sequence conservation among three families of kinases (Table 2). We found that BMPRI kinases (ALK1, 2, 3 and 6) share a conserved amino acid triad Leu281, Ile282 and Thr283 (ALK2 numbering) adjacent to the ATP binding site, which we denote the “pre-hinge” region. Thr283 is known as the “gatekeeper” as it blocks access of ligands to a hydrophobic pocket next to the site of ATP binding [26]. In TGF- $\beta$  kinases (ALK4, ALK5, ALK7), this pre-hinge triad consists instead of a conserved sequence Leu278, Val279 and Ser280 (ALK5 numbering) [14, 19, 59]. In our simulation, this less bulky pre-hinge triad in ALK5 allows the DMH1 quinoline ring to reach deeper within this hydrophobic region (Fig 6B). This binding pose shift results in the loss of two major favorable electrostatic interactions between DMH1 and His283 and Lys232 in ALK5. Noticeably, ALK4, which has the same pre-hinge triad as ALK5, also does not bind DMH1. On the other hand, crystal structures of ALK5 with the potent ALK5 inhibitors SB431542 (PDB ID 3T2M),



**Fig 5.** Left: Superposition of wtALK2 (cyan) and caALK2 (orange) backbone in cartoon. P-loop Tyr219 is shown in sticks with overlapped snapshots from the last 2 ns of simulation. DMH1 atoms are shown as small spheres. Right: RMSF of the P-loop in the unbound (apo) and DMH1 bound states. Statistical difference is represented by an \* above the bound conformation; \* =  $P < 0.05$ .

doi:10.1371/journal.pone.0132221.g005



**Fig 6.** Binding conformations of DMH1 in ALK2 (top) and ALK5 (bottom) from molecular dynamics simulations. The conserved triad of amino acids consists of the gatekeeper and two pre-hinge residues shown in VDW mode. The rest of the protein is shown in surface mode.

doi:10.1371/journal.pone.0132221.g006

**Table 2. Hinge region sequences of ALK1-7 and the IC<sub>50</sub> values of LDN193189 and DMH1.**

Family	Kinase <sup>a</sup>	Pre-hinge	Gatekeeper	Hinge	LDN 193189 IC <sub>50</sub> (nM)	DMH1 IC <sub>50</sub> (nM)
BMPR-I	ALK1	LI	T	HYHEHGS	13.3	27
	ALK2	LI	T	HYHEMGS	40.7	107.9
	ALK3	LI	T	DYHENGGS	<5	<5
	ALK6	LI	T	DYHENGGS	60	47.6
TGF-βR	ALK4	LV	S	DYHEHGS	1825	9622
	ALK5	LV	S	DYHEHGS	565	>100,000
	ALK7	LV	S	EYHEQGS	N/A	N/A

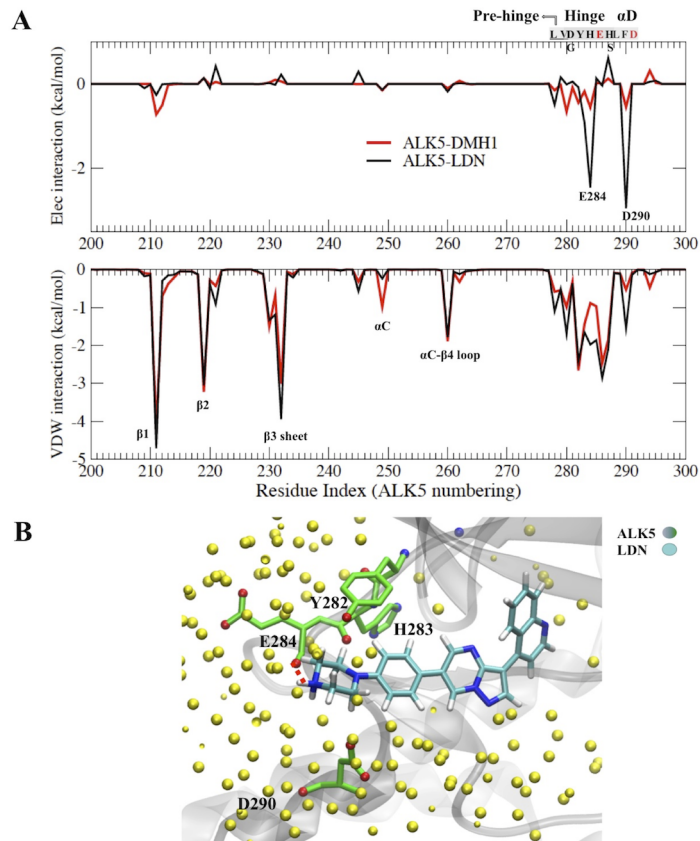
<sup>a</sup> IC<sub>50</sub> values are taken from reference [14] for each kinase.

doi:10.1371/journal.pone.0132221.t002

GW855857 (PDB ID 3HMM), compound 19 (PDB ID 2WOU), and indolinone (PDB ID 2X7O) [26, 60–62] indicate that their binding is associated with two hydrogen bonds, one of which must be with the hinge region His283. Therefore, our model can explain why DMH1 is not a potent inhibitor of ALK5.

**Solvent exposed piperazine ring makes LDN193189 a more potent inhibitor of ALK5 than DMH1.** Compound LDN193189 is a close analog of DMH1 (Fig 1). The piperazine ring in LDN193189 [16] was designed to replace the solvent exposed moiety of dorsomorphin in order to improve the solubility and metabolic stability by avoiding the phase I O-dealkylation metabolic pathway. However, LDN193189 turned out to be a more potent inhibitor of ALK2 and also ALK5 compared with DMH1 (Table 2). In order to explain the difference between DMH1 and LDN193189 in their interaction with ALK5, we used the fully equilibrated ALK5-DMH1 conformation, and replaced DMH1 with LDN193189 by substituting the isopropoxy moiety of DMH1 with a piperazine ring using the Molecular Operating Environment (MOE) program [63]. The new ALK5-LDN193189 complex was then solvated in explicit solvent and submitted for further minimization and molecular dynamics simulation. The RMSD and the distance of center of mass between ligand and receptor show that LDN193189 quickly reaches equilibrium in ALK5 within 18 ns of simulation (S5 Fig). The equilibrated binding poses of LDN193189 and DMH1 are essentially the same, since the two molecules are highly similar. The average per-residue electrostatic interaction from the last 6 ns reveals clearly a more favorable electrostatic interaction between LDN193189 and ALK5 Glu284 (hinge region) and Asp290 (αD helix) residues (Fig 7A). Hydrogen bonding analysis indicates that the protonated piperazine ring of LDN193189 forms a hydrogen bond with Glu284 40% of the simulation time (Fig 7B). The van der Waals interaction between ALK5 and LDN193189 is also stronger than with DMH1 at the hinge region (Fig 7A). Therefore, our model illustrates that the solvent exposed R2 group in dorsomorphin analogs (Fig 1) also plays an important role in binding selectivity. This group can be modified to manipulate the binding selectivity between ALK isoforms.

**Electrostatic interaction between DMH1 and VEGFR2 is not sufficient to establish a strong binding.** Our FEP/H-REMD calculations demonstrate that DMH1 has very low binding affinity toward both VEGFR2 DFG-in and DFG-out conformations (binding free energy of -3.4 and -1.7 kcal/mol, respectively). The DMH1 pose (3) in VEGFR2-in/out from docking is similar to LDN193189 in the ALK2 crystal structure (Fig 2). However, the molecular dynamics simulation in the fully solvated system brought to light the deviation of DMH1 from its original docked pose in VEGFR2 within 10 ns. The most relevant ligand motion occurs by the rotation of the quinoline ring when it binds to VEGFR2 (S6A Fig). Due to this deviation, DMH1 in VEGFR2, compared to that in ALK2, misses a major electrostatic interaction and hydrogen

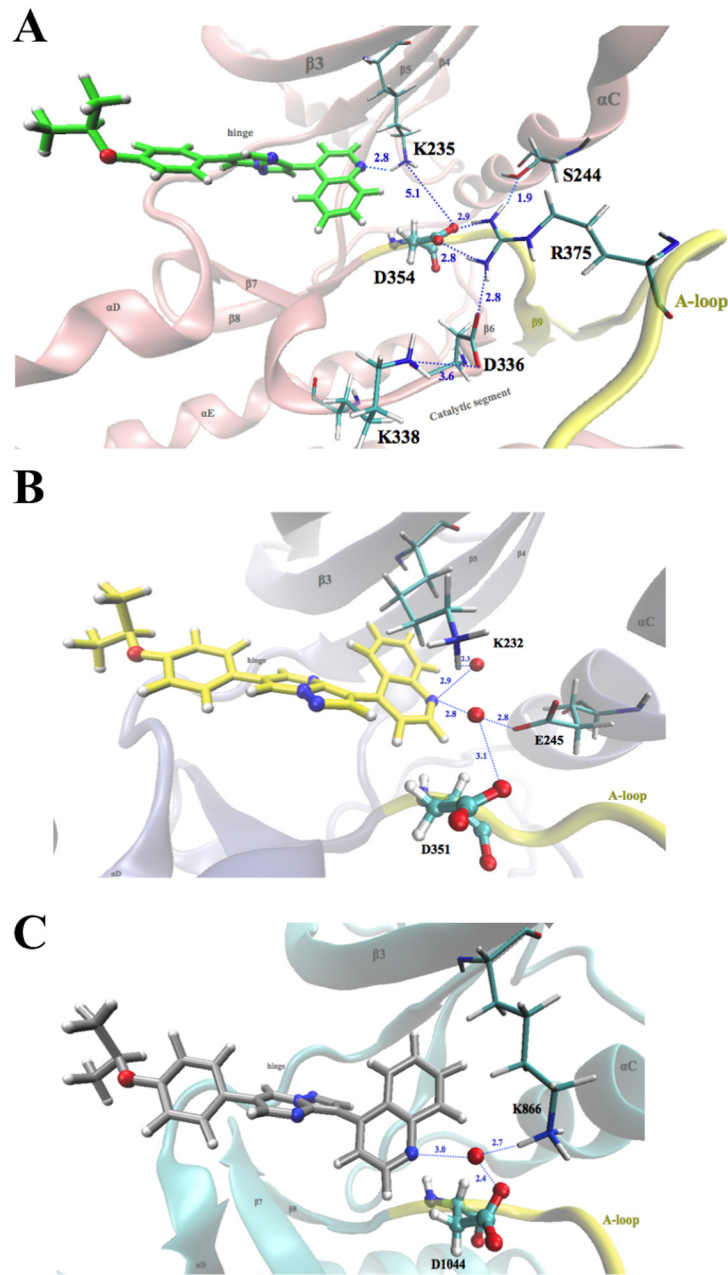


**Fig 7. A.** Residue decomposed electrostatic interactions (top) and van der Waals interaction (bottom) between LDN193189 and ALK5 (black line), compared with DMH1 (red line). **B.** Residues in ALK5 that contribute favorable electrostatic or van der Waals interactions with the LDN193189 piperazine moiety are highlighted in green. A hydrogen bond between the LDN193189 protonated N and Glu284 backbone is shown as a red dotted line. Water molecules within 8 Å of LDN193189 are shown as yellow spheres.

doi:10.1371/journal.pone.0132221.g007

bond with Lys868 of the  $\beta 3$  strand. A survey of 28 x-ray crystal structures of VEGFR2-inhibitor complexes (see all PDB IDs [64]) also indicates that potent VEGFR2 inhibitors typically form two to three direct hydrogen bonds with Cys919 (hinge region) and/or Asp1046 (DFG) and occasionally Glu885 ( $\alpha C$  helix). Compared to all the potent VEGFR2 inhibitors, the molecular dynamics-equilibrated DMH1 only forms one direct hydrogen bond with Cys919 (S6B and S6C Fig). In conclusion, both the positive electrostatic free energy component (Table 1) and the PDB database survey reveal that DMH1 does not establish the necessary favorable electrostatic interactions with VEGFR2.

**Effect of DMH1 binding on the salt-bridge network and the A-loop.** Previous ALK2 crystal structures show that there are hydrogen-bonding and salt-bridge networks between Lys235 ( $\beta 3$  strand) and Asp354 (DLG motif), and between Arg375 (A-loop) and Ser244 ( $\alpha C$ -helix), Asp336 (catalytic segment), and Asp354 [10]. Here, we analyze the dynamics of hydrogen bonding and salt bridges by calculating the probability of their occurrence (i.e., occupancy) in presence of DMH1 during the simulation. In ALK2, a direct hydrogen bond between DMH1 quinoline nitrogen and Lys235 side chain weakens the salt bridge between Lys235 in the  $\beta 3$  strand and Asp354 in the A-loop DLG-motif (Fig 8). As a result, the A-loop fluctuation in wtALK2 and caALK2 increases upon DMH1 binding ( $p = 0.0069$  and  $p = 0.0174$ , respectively) (S7 Fig). In the binding models of ALK5 and VEGFR2 in the in and out conformations, DMH1



**Fig 8. A.** Hydrogen-bonding and salt-bridge network of the wtALK2-DMH1 complex. Protein backbone is partially shown. DMH1 and residues involved in the network are shown in stick with atoms colored by type: red, oxygen; blue, nitrogen; green, carbon; and white, hydrogen. A-loop backbone is shown in yellow. The average distance of each salt bridge and hydrogen bond is indicated. **B.** ALK5-DMH1 complex showing water molecule as red ball bridging two hydrogen bonds between Lys232 and Asp351 and the nitrogen atom on the quinolone ring of DMH1. **C.** VEGFR2-in conformation and numbering showing water molecule bridging two hydrogen bonds between Lys866 and Asp1044 and the nitrogen atom on the quinolone ring of DMH1.

doi:10.1371/journal.pone.0132221.g008

only forms water-mediated hydrogen bonds with Lys232 ( $\beta 3$  strand) and Asp351 (DLG motif) in ALK5, or Lys866 ( $\beta 3$  strand) and Asp1044 (DFG motif) in VEGFR2 in both in and out conformations (Fig 8B and 8C). The network of Lys232-DMH1-Asp351 water-mediated hydrogen bonding plus Lys232-Asp351 salt bridge (ALK5 numbering) stabilizes the A-loop compared to

the unbound structure ( $p < 0.001$  for ALK5 and VEGFR2-in, and  $p = 0.0468$  for VEGFR2-out). Therefore, the dynamic analysis here emphasizes that the hydrogen bonding and salt bridge network near the ATP binding site are tightly tied. Weakening or strengthening a single hydrogen bond by a small molecule such as DMH1 is likely to have long distance effects on the strength of the hydrogen-bonding and salt-bridge network as well as the dynamics of the A-loop. This finding is consistent with the dynamically coupled allosteric network in Src kinase previously reported [65].

## Conclusions

Given the important roles of BMP signaling in embryogenesis and homeostasis, small molecules that specifically target BMPRI are highly sought after. In recent years, BMP inhibitors including dorsomorphin, DMH1, LDN193189 and other analogs, have been developed to inhibit BMPRI subtype ALK2. However, the molecular mechanism underlying their binding selectivity between ALK2 and other structurally closely related kinases has remained unknown. In the present study, we used computational tools such as docking, molecular dynamics simulation and free energy calculations to address this issue. While our docking scores from AutoDock did not differentiate the binding selectivity of DMH1 among ALK2, ALK5 and VEGFR2, our FEP/H-REMD simulations successfully reproduced the fact that DMH1 only binds to ALK2, but not ALK5 and VEGFR2, in excellent agreement with experimental measurements. The free energy decomposition analysis showed that van der Waals dispersive interactions dominate the total binding affinity, but electrostatic interactions are largely responsible for DMH1 discrimination between ALK2/5 and VEGFR2. The per-residue interactions between the ligand and the kinases clearly revealed that the favorable electrostatic interaction with catalytic Lys235 and van der Waals interaction with the P-loop Tyr219 play critical roles in ALK2 binding specificity. A shift in the DMH1 binding pose in ALK5, mainly caused by the pre-hinge triad including gatekeeper Ser280 residue, results in the loss of several favorable interactions between the ligand and receptor. To understand the tighter binding of LDN193189 to ALK5, we performed molecular dynamics simulation of LDN193189 in ALK5 with explicit solvent. The simulation showed that the protonated piperazine ring on LDN193189 forms stable hydrogen bonds with Glu284 in ALK5. Our analysis provides the rationale for improving ALK2/ALK5 selectivity of LDN193189 analogs through modifying the solvent exposed group.

In summary, the current study reveals how small changes in the binding site residue type (e.g. pre-hinge region in ALK2 vs. ALK5) or residue conformation (e.g. Tyr219 in caALK2 vs. wtALK2), as well as small ligand modification (e.g. DMH1 vs. LDN193189) will cause distinct binding profiles and selectivity. It is, therefore, difficult to predict the binding specificity of small molecules in BMPI receptors solely based on the ligand-based structure-activity relationship or static binding information from rigid protein docking and crystal structures. In contrast, the computational methodology applied in this study takes into consideration local conformational changes as well as the effect of explicit solvent, representing a new way in understanding binding specificity of small molecule BMP inhibitors to their receptor kinases, which is critical for developing exclusively selective inhibitors for each subtype of BMPRI.

In terms of computational cost, each 1ns FEP/H-REMD simulation took approximately 5 hours real time (128 replicas running on 512 compute nodes in parallel on the Mira Blue Gene/Q supercomputer). The time-evolution of the absolute binding free energy plot suggests that even though it took 2–4 times longer to get a converged absolute binding free energy, the rank of the binding free energies among the three kinases is correct within 10 hours real time for each system. Therefore, the current study demonstrates that the FEP/H-REMD approach can serve as a robust method to validate the binding poses from virtual docking when the crystal



structure of a ligand-receptor complex is not available. Furthermore, we show that molecular dynamics-based free energy simulation can explain and predict binding selectivities of BMP inhibitors among highly conserved ATP binding sites. Our computational approach presented here would play a significant role in the rational design of exclusively selective and potent BMP inhibitors.

## Supporting Information

**S1 Fig.** Structure of ALK2 with labeled regions (Fig A). Structural alignment of ALK2 and ALK5 (Fig B). Structural alignment of ALK2 with VEGFR2-in and VEGFR2-out (Fig C). (PDF)

**S2 Fig.** The time evolutions of the RMSD of protein and ligand, as well as the center-of-mass distance between protein and ligand in each simulation system. (DOCX)

**S3 Fig.** Thermodynamic cycle used to calculate the absolute binding free energy. (DOCX)

**S4 Fig.** Progression of the total binding free energies of DMH1 in different systems with respect to the FEP/H-REMD simulation time (Fig A). Progression of the free energy components with respect to the coupling parameters (Fig B). (DOCX)

**S5 Fig.** The time evolutions of the RMSD of ALK5 and the distance between the center-of-mass of the LDN193189 and ALK5 during the equilibrium simulation. (DOCX)

**S6 Fig.** The motion of DMH1 deviating from its original docked pose (Figure A). Alignment of VEGFR2-out DMH1 complex with compound-19 in PDB ID 3VO3 (Figure B). Alignment of VEGFR2-in DMH1 complex with compound 11-b in PDB ID 3CJG (Figure C). (DOCX)

**S7 Fig.** Fluctuation of the A-loop backbone upon DMH1 binding. (DOCX)

**S1 Table.** Structures of BMP inhibitors and fold selectivity against ALK2 kinase. (DOCX)

## Acknowledgments

The computations were made possible by the Director's Discretionary award at Argonne Leadership Computing Facility (ALCF), which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. This work was supported by the seed funds of Western University of Health Sciences for Yun Luo and Jijun Hao. The authors are grateful to Dr. Alex N. Bullock at University of Oxford for helpful discussions regarding ALK2 crystal structures and Dr. Benoît Roux at University of Chicago and Dr. Wei Jiang at ALCF for helpful discussions regarding FEP/H-REMD implementation.

## Author Contributions

Conceived and designed the experiments: AA JH YL. Performed the experiments: AA AEL. Analyzed the data: AA YL. Wrote the paper: AA PO JH YL.

## References

1. Miyazono K, Kamiya Y, Morikawa M. Bone morphogenetic protein receptors and signal transduction. *J Biochem.* 2010; 147(1):35–51. doi: [10.1093/jb/mvp148](https://doi.org/10.1093/jb/mvp148) PMID: [19762341](https://pubmed.ncbi.nlm.nih.gov/19762341/)
2. Huse M, Muir TW, Xu L, Chen YG, Kuriyan J, Massague J. The TGF beta receptor activation process: an inhibitor- to substrate-binding switch. *Mol Cell.* 2001; 8(3):671–682. PMID: [11583628](https://pubmed.ncbi.nlm.nih.gov/11583628/)
3. Shore EM, Xu M, Feldman GJ, Fenstermacher DA, Cho TJ, Choi IH, et al. A recurrent mutation in the BMP type I receptor ACVR1 causes inherited and sporadic fibrodysplasia ossificans progressiva. *Nature genetics.* 2006; 38(5):525–527. PMID: [16642017](https://pubmed.ncbi.nlm.nih.gov/16642017/)
4. Kaplan FS, Xu M, Seemann P, Connor JM, Glaser DL, Carroll L, et al. Classic and atypical fibrodysplasia ossificans progressiva (FOP) phenotypes are caused by mutations in the bone morphogenetic protein (BMP) type I receptor ACVR1. *Hum Mutat.* 2009; 30(3):379–390. doi: [10.1002/humu.20868](https://doi.org/10.1002/humu.20868) PMID: [19085907](https://pubmed.ncbi.nlm.nih.gov/19085907/)
5. Hu-Lowe DD, Chen E, Zhang L, Watson KD, Mancuso P, Lappin P, et al. Targeting activin receptor-like kinase 1 inhibits angiogenesis and tumorigenesis through a mechanism of action complementary to anti-VEGF therapies. *Cancer Res.* 2011; 71(4):1362–1373. doi: [10.1158/0008-5472.CAN-10-1451](https://doi.org/10.1158/0008-5472.CAN-10-1451) PMID: [21212415](https://pubmed.ncbi.nlm.nih.gov/21212415/)
6. Zhou XP, Woodford-Richens K, Lehtonen R, Kurose K, Aldred M, Hampel H, et al. Germline mutations in BMPR1A/ALK3 cause a subset of cases of juvenile polyposis syndrome and of Cowden and Bannayan-Riley-Ruvalcaba syndromes. *Am J Hum Genet.* 2001; 69(4):704–711. PMID: [11536076](https://pubmed.ncbi.nlm.nih.gov/11536076/)
7. Cheah PY, Wong YH, Chau YP, Loi C, Lim KH, Lim JF, et al. Germline bone morphogenesis protein receptor 1A mutation causes colorectal tumorigenesis in hereditary mixed polyposis syndrome. *Am J Gastroenterol.* 2009; 104(12):3027–3033. doi: [10.1038/ajg.2009.542](https://doi.org/10.1038/ajg.2009.542) PMID: [19773747](https://pubmed.ncbi.nlm.nih.gov/19773747/)
8. Saetrom P, Biesinger J, Li SM, Smith D, Thomas LF, Majzoub K, et al. A risk variant in an miR-125b binding site in BMPR1B is associated with breast cancer pathogenesis. *Cancer Res.* 2009; 69(18):7459–7465. doi: [10.1158/0008-5472.CAN-09-1201](https://doi.org/10.1158/0008-5472.CAN-09-1201) PMID: [19738052](https://pubmed.ncbi.nlm.nih.gov/19738052/)
9. Yu PB, Hong CC, Sachidanandan C, Babbitt JL, Deng DY, Hoyng SA, et al. Dorsomorphin inhibits BMP signals required for embryogenesis and iron metabolism. *Nature Chemical Biology.* 2008; 4(1):33–41. PMID: [18026094](https://pubmed.ncbi.nlm.nih.gov/18026094/)
10. Chaikuad A, Alfano I, Kerr G, Sanvitale CE, Boergemann JH, Triffitt JT, et al. Structure of the bone morphogenetic protein receptor ALK2 and implications for fibrodysplasia ossificans progressiva. *J Biol Chem.* 2012; 287(44):36990–36998. doi: [10.1074/jbc.M112.365932](https://doi.org/10.1074/jbc.M112.365932) PMID: [22977237](https://pubmed.ncbi.nlm.nih.gov/22977237/)
11. Cai J, Pardali E, Sanchez-Duffhues G, ten Dijke P. BMP signaling in vascular diseases. *FEBS Lett.* 2012; 586(14):1993–2002. doi: [10.1016/j.febslet.2012.04.030](https://doi.org/10.1016/j.febslet.2012.04.030) PMID: [22710160](https://pubmed.ncbi.nlm.nih.gov/22710160/)
12. Mueller TD, Nickel J. Promiscuity and specificity in BMP receptor activation. *FEBS Lett.* 2012; 586(14):1846–1859. doi: [10.1016/j.febslet.2012.02.043](https://doi.org/10.1016/j.febslet.2012.02.043) PMID: [22710174](https://pubmed.ncbi.nlm.nih.gov/22710174/)
13. Hao J, Ho JN, Lewis JA, Karim KA, Daniels RN, Gentry PR, et al. In vivo structure-activity relationship study of dorsomorphin analogues identifies selective VEGF and BMP inhibitors. *ACS Chem Biol.* 2010; 5(2):245–253. doi: [10.1021/cb9002865](https://doi.org/10.1021/cb9002865) PMID: [20020776](https://pubmed.ncbi.nlm.nih.gov/20020776/)
14. Engers DW, Frist AY, Lindsley CW, Hong CC, Hopkins CR. Synthesis and structure-activity relationships of a novel and selective bone morphogenetic protein receptor (BMP) inhibitor derived from the pyrazolo[1,5-a]pyrimidine scaffold of dorsomorphin: the discovery of ML347 as an ALK2 versus ALK3 selective MLPCN probe. *Bioorg Med Chem Lett.* 2013; 23(11):3248–3252. doi: [10.1016/j.bmcl.2013.03.113](https://doi.org/10.1016/j.bmcl.2013.03.113) PMID: [23639540](https://pubmed.ncbi.nlm.nih.gov/23639540/)
15. Hong CC, Yu PB. Applications of small molecule BMP inhibitors in physiology and disease. *Cytokine Growth Factor Rev.* 2009; 20(5–6):409–418. doi: [10.1016/j.cytogfr.2009.10.021](https://doi.org/10.1016/j.cytogfr.2009.10.021) PMID: [19914855](https://pubmed.ncbi.nlm.nih.gov/19914855/)
16. Yu PB, Deng DY, Lai CS, Hong CC, Cuny GD, Bouxsein ML, et al. BMP type I receptor inhibition reduces heterotopic ossification. *Nat Med.* 2008; 14(12):1363–1369. doi: [10.1038/nm.1888](https://doi.org/10.1038/nm.1888) PMID: [19029982](https://pubmed.ncbi.nlm.nih.gov/19029982/)
17. Yu PB, Hong CC, Sachidanandan C, Babbitt JL, Deng DY, Hoyng SA, et al. Dorsomorphin inhibits BMP signals required for embryogenesis and iron metabolism. *Nat Chem Biol.* 2008; 4(1):33–41. PMID: [18026094](https://pubmed.ncbi.nlm.nih.gov/18026094/)
18. Mohedas AH, Wang Y, Sanvitale CE, Canning P, Choi S, Xing X, et al. Structure-activity relationship of 3,5-diaryl-2-aminopyridine ALK2 inhibitors reveals unaltered binding affinity for fibrodysplasia ossificans progressiva causing mutants. *J Med Chem.* 2014; 57(19):7900–7915. doi: [10.1021/jm501177w](https://doi.org/10.1021/jm501177w) PMID: [25101911](https://pubmed.ncbi.nlm.nih.gov/25101911/)
19. Mohedas AH, Xing X, Armstrong KA, Bullock AN, Cuny GD, Yu PB. Development of an ALK2-biased BMP type I receptor kinase inhibitor. *ACS Chem Biol.* 2013; 8(6):1291–1302. doi: [10.1021/cb300655w](https://doi.org/10.1021/cb300655w) PMID: [23547776](https://pubmed.ncbi.nlm.nih.gov/23547776/)

20. Sanvitale CE, Kerr G, Chaikuad A, Ramel MC, Mohedas AH, Reichert S, et al. A new class of small molecule inhibitor of BMP signaling. *PLOS One*. 2013; 8(4):e62721. doi: [10.1371/journal.pone.0062721](https://doi.org/10.1371/journal.pone.0062721) PMID: [23646137](https://pubmed.ncbi.nlm.nih.gov/23646137/)
21. Cuny GD, Yu PB, Laha JK, Xing X, Liu JF, Lai CS, et al. Structure-activity relationship study of bone morphogenetic protein (BMP) signaling inhibitors. *Bioorg Med Chem Lett*. 2008; 18(15):4388–4392. doi: [10.1016/j.bmcl.2008.06.052](https://doi.org/10.1016/j.bmcl.2008.06.052) PMID: [18621530](https://pubmed.ncbi.nlm.nih.gov/18621530/)
22. Lin YL, Meng Y, Huang L, Roux B. Computational Study of Gleevec and G6G Reveals Molecular Determinants of Kinase Inhibitor Selectivity. *J Am Chem Soc*. 2014; 136(42):14753–14762. doi: [10.1021/ja504146x](https://doi.org/10.1021/ja504146x) PMID: [25243930](https://pubmed.ncbi.nlm.nih.gov/25243930/)
23. Lin Y-L, Roux B. Computational Analysis of the Binding Specificity of Gleevec to Abl, c-Kit, Lck, and c-Src Tyrosine Kinases. *J Am Chem Soc*. 2013; 135(39):14741–14753. doi: [10.1021/ja405939x](https://doi.org/10.1021/ja405939x) PMID: [24001034](https://pubmed.ncbi.nlm.nih.gov/24001034/)
24. Lin Y-L, Meng Y, Jiang W, Roux B. Explaining why Gleevec is a specific and potent inhibitor of Abl kinase. *Proc Natl Acad Sci U S A*. 2013; 110(5):1664–1669. doi: [10.1073/pnas.1214330110](https://doi.org/10.1073/pnas.1214330110) PMID: [23319661](https://pubmed.ncbi.nlm.nih.gov/23319661/)
25. Gumbart JC, Roux B, Chipot C. Standard binding free energies from computer simulations: What is the best strategy? *J Chem Theory Comput*. 2013; 9(1):794–802. PMID: [23794960](https://pubmed.ncbi.nlm.nih.gov/23794960/)
26. Ogunjimi AA, Zeqiraj E, Ceccarelli DF, Sicheri F, Wrana JL, David L. Structural basis for specificity of TGFbeta family receptor small molecule inhibitors. *Cellular signalling*. 2012; 24(2):476–483. doi: [10.1016/j.cellsig.2011.09.027](https://doi.org/10.1016/j.cellsig.2011.09.027) PMID: [21983015](https://pubmed.ncbi.nlm.nih.gov/21983015/)
27. Miyamoto N, Sakai N, Hirayama T, Miwa K, Oguro Y, Oki H, et al. Discovery of N-[5-((2-((cyclopropylcarbonyl)amino)imidazo[1,2-b]pyridazin-6-yl)oxy)-2-methylphenyl]-1,3-dimethyl-1H-pyrazole-5-carboxamide (TAK-593), a highly potent VEGFR2 kinase inhibitor. *Bioorganic & medicinal chemistry*. 2013; 21(8):2333–2345.
28. Harris PA, Boloor A, Cheung M, Kumar R, Crosby RM, Davis-Ward RG, et al. Discovery of 5-[4-[(2,3-dimethyl-2H-indazol-6-yl)methylamino]-2-pyrimidinyl]amino]-2-methylbenzenesulfonamide (Pazopanib), a novel and potent vascular endothelial growth factor receptor inhibitor. *J Med Chem*. 2008; 51(15):4632–4640. doi: [10.1021/jm800566m](https://doi.org/10.1021/jm800566m) PMID: [18620382](https://pubmed.ncbi.nlm.nih.gov/18620382/)
29. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem*. 1983; 4(2):187–217.
30. Brooks BR, Brooks CL III., MacKerell AD Jr., Nilsson L, Petrella RJ, Roux B, et al. CHARMM: The Biomolecular Simulation Program. *J Comput Chem*. 2009; 30(10):1545–1614. doi: [10.1002/jcc.21287](https://doi.org/10.1002/jcc.21287) PMID: [19444816](https://pubmed.ncbi.nlm.nih.gov/19444816/)
31. Rostkowski M, Olsson MH, Sondergaard CR, Jensen JH. Graphical analysis of pH-dependent properties of proteins predicted using PROPKA. *BMC Struct Biol*. 2011; 11:6. doi: [10.1186/1472-6807-11-6](https://doi.org/10.1186/1472-6807-11-6) PMID: [21269479](https://pubmed.ncbi.nlm.nih.gov/21269479/)
32. Humphrey W, Schulten K. VMD: visual molecular dynamics. *J Mol Graphics*. 1996; 14:27–38.
33. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem*. 2008; 29(11):1859–1865. doi: [10.1002/jcc.20945](https://doi.org/10.1002/jcc.20945) PMID: [18351591](https://pubmed.ncbi.nlm.nih.gov/18351591/)
34. Best RB, Zhu X, Shim J, Lopes PEM, Mittal J, Feig M, et al. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone  $\phi$ ,  $\psi$  and Side-Chain  $\chi_1$  and  $\chi_2$  Dihedral Angles. *J Chem Theory Comput*. 2012; 8(9):3257–3273. PMID: [23341755](https://pubmed.ncbi.nlm.nih.gov/23341755/)
35. MacKerell AD Jr., Feig M, Brooks CL 3rd. Improved treatment of the protein backbone in empirical force fields. *J Am Chem Soc*. 2004; 126(3):698–699. PMID: [14733527](https://pubmed.ncbi.nlm.nih.gov/14733527/)
36. MacKerell AD Jr., Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B*. 1998; 102(18):3586–3616. doi: [10.1021/jp973084f](https://doi.org/10.1021/jp973084f) PMID: [24889800](https://pubmed.ncbi.nlm.nih.gov/24889800/)
37. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein MLK. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 1983; 79(2):926–935.
38. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. *J Comput Chem*. 1983; 4:187–217.
39. Planaria Software LLC, Seattle, WA.
40. Huang L, Roux B. Automated Force Field Parameterization for Non-Polarizable and Polarizable Atomic Models Based on Target Data. *J Chem Theory Comput*. 2013; 9(8):3543–3556.
41. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*. 2009; 30(16):2785–2791. doi: [10.1002/jcc.21256](https://doi.org/10.1002/jcc.21256) PMID: [19399780](https://pubmed.ncbi.nlm.nih.gov/19399780/)

42. Muegge I, Enyedy IJ. Virtual screening for kinase targets. *Curr Med Chem*. 2004; 11(6):693–707. PMID: [15032724](#)
43. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. *J Comput Chem*. 2005; 26(16):1781–1802. PMID: [16222654](#)
44. Darden T D, York LP. Particle-mesh Ewald—an N-log(N) method for Ewald sums in large systems. *J Chem Phys*. 1993; 98:10089–10092.
45. Ryckaert J-P, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys*. 1977; 23(3):327–341.
46. Straatsma TP, Mccammon JA. Computational Alchemy. *Annu Rev Phys Chem*. 1992; 43:407–435.
47. Kollman P. Free-Energy Calculations—Applications to Chemical and Biochemical Phenomena. *Chem Rev*. 1993; 93(7):2395–2417.
48. Deng Y, Roux B. Calculation of standard binding free energies: aromatic molecules in the T4 lysozyme L99A mutant. *J Chem Theory Comput*. 2006; 2:1255–1273.
49. Weeks JD, Chandler D, Andersen HC. Role of Repulsive Forces in Determining Equilibrium Structure of Simple Liquids. *J Chem Phys*. 1971; 54(12):5237–5247.
50. Jiang W, Hodoscek M, Roux B. Computation of Absolute Hydration and Binding Free Energy with Free Energy Perturbation Distributed Replica-Exchange Molecular Dynamics (FEP/REMD). *J Chem Theory Comput*. 2009; 5(10):2583–2588. PMID: [21857812](#)
51. Jiang W, Phillips JC, Huang L, Fajer M, Meng Y, Gumbart JC, et al. Generalized Scalable Multiple Copy Algorithms for Molecular Dynamics Simulations in NAMD. *Comput Phys Commun*. 2014; 185(3):908–916. PMID: [24944348](#)
52. Jiang W, Roux B. Free Energy Perturbation Hamiltonian Replica-Exchange Molecular Dynamics (FEP/H-REMD) for Absolute Ligand Binding Free Energy Calculations. *J Chem Theory Comput*. 2010; 6(9):2559–2565. PMID: [21857813](#)
53. Payne CM, Jiang W, Shirts MR, Himmel ME, Crowley MF, Beckham GT. Glycoside Hydrolase Processivity Is Directly Related to Oligosaccharide Binding Free Energy. *J Am Chem Soc*. 2013; 135(50):18831–18839. doi: [10.1021/ja407287f](#) PMID: [24279927](#)
54. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett*. 1999; 314(1–2):141–151.
55. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM. *J Comput Chem*. 1992; 13:1011–1021.
56. Mobley DL, Klimovich PV. Perspective: Alchemical free energy calculations for drug discovery. *J Chem Phys*. 2012; 137(23):230901. doi: [10.1063/1.4769292](#) PMID: [23267463](#)
57. Rocklin GJ, Mobley DL, Dill KA. Separated topologies—a method for relative binding free energy calculations using orientational restraints. *J Chem Phys*. 2013; 138(8):085104. doi: [10.1063/1.4792251](#) PMID: [23464180](#)
58. Chamberlain P, Delker S, Pagarigan B, Mahmoudi A, Jackson P, Abbasian M, et al. Crystal structures of PRK1 in complex with the clinical compounds lestaurinib and tofacitinib reveal ligand induced conformational changes. *PLOS One*. 2014; 9(8):e103638. doi: [10.1371/journal.pone.0103638](#) PMID: [25111382](#)
59. Pruitt KD, Brown GR, Hiatt SM, Thibaud-Nissen F, Astashyn A, Ermolaeva O, et al. RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res*. 2014; 42(Database issue):D756–763. doi: [10.1093/nar/gkt1114](#) PMID: [24259432](#)
60. Goldberg FW, Ward RA, Powell SJ, Debreczeni JE, Norman RA, Roberts NJ, et al. Rapid generation of a high quality lead for transforming growth factor-beta (TGF-beta) type I receptor (ALK5). *J. Med. Chem*. 2009; 52(23):7901–7905. doi: [10.1021/jm900807w](#) PMID: [19736928](#)
61. Gellibert F, Fouchet MH, Nguyen VL, Wang R, Krysa G, de Gouville AC, et al. Design of novel quinazoline derivatives and related analogues as potent and selective ALK5 inhibitors. *Bioorg Med Chem Lett*. 2009; 19(8):2277–2281. doi: [10.1016/j.bmcl.2009.02.087](#) PMID: [19285388](#)
62. Roth GJ, Heckel A, Brandt T, Grauert M, Hoerer S, Kley JT, et al. Design, synthesis, and evaluation of indolinones as inhibitors of the transforming growth factor beta receptor I (TGFbetaRI). *J Med Chem*. 2010; 53(20):7287–7295. doi: [10.1021/jm100812a](#) PMID: [20919678](#)
63. Molecular Operating Environment (MOE), 2013.08; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2015. 2013.08.
64. 1YWN, 1Y6A, 1Y6B, 3VHE, 3VNT, 3VO3, 2OH4, 3EFL, 3C7Q, 4ASE, 4ASD, 3VHK, 2QU5, 2QU6, 2P2H, 2P2I, 3B8Q, 3B8R, 3BE2, 3CP9, 3DTW, 3EWH, 3U6J, 3CJG, 3CJF, 2R15, 4AG8, 2XIR.
65. Foda ZH, Shan Y, Kim ET, Shaw DE, Seeliger MA. A dynamically coupled allosteric network underlies binding cooperativity in Src kinase. *Nat Commun*. 2015; 6:5939. doi: [10.1038/ncomms6939](#) PMID: [25600932](#)