# Identification and Validation of Novel PERK Inhibitors

Qiantao Wang,[†,§] Jihyun Park,[†] Ashwini K. Devkota,[‡] Eun Jeong Cho,[‡] Kevin N. Dalby,[*,†,‡] and Pengyu Ren[*,§]

[†]Division of Medicinal Chemistry, College of Pharmacy, The University of Texas at Austin, Austin, Texas 78712, United States
[‡]Texas Screening Alliance for Cancer Therapeutics, The University of Texas at Austin, Austin, Texas 78712, United States
[§]Department of Biomedical Engineering, The University of Texas at Austin, Austin, Texas 78712, United States

**ABSTRACT:** PERK, as one of the principle unfolded protein response signal transducers, is believed to be associated with many human diseases, such as cancer and type-II diabetes. There has been increasing effort to discover potent PERK inhibitors due to its potential therapeutic interest. In this study, a computer-based virtual screening approach is employed to discover novel PERK inhibitors, followed by experimental validation. Using a focused library, we show that a consensus approach, combining pharmacophore modeling and docking, can be more cost-effective than using either approach alone. It is also demonstrated that the conformational flexibility near the active site is an important consideration in structure-based docking and can be addressed by using molecular dynamics. The consensus approach has further been applied to screen the ZINC lead-like database, resulting in the identification of 10 active compounds, two of which show $IC_{50}$ values that are less than 10 $\mu$M in a dose−response assay.
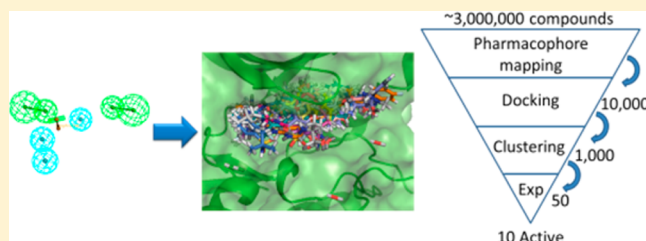
## ■ INTRODUCTION

Virtual library screening and molecular modeling have been used widely in the drug discovery process and have yielded experimentally confirmed hits for various protein targets.[1−6] Different virtual screening (VS) approaches have been used, including structure-based docking and ligand-based mapping. Not surprisingly, there are limitations in both approaches. For example, reliable and relevant structures of the target proteins are necessary for docking. In contrast ligand-based mapping only requires knowledge of known ligands of the target. Often, a novel target of therapeutic interest does not have a crystal structure. For instance, a recent survey[7] showed that there were crystal structures available for only 155 individual kinases among the total 518 human kinases. The time needed to obtain such crystal structures varies considerably, and the outcome is not guaranteed. In addition, crystal structures without bound ligands may not be relevant, especially for proteins that undergo large conformational changes upon ligand-binding. The solution in such situations would be either to generate a model structure (either entirely or partially) via homology modeling and/or molecular dynamics (MD) simulation[8−10] or to apply a ligand-based mapping approach, such as pharmacophore mapping and shape-based screening of the ligand so the protein structures are not used.[6,11−15]

PKR-like endoplasmic reticulum kinase (PERK), along with two other proteins IRE1 (inositol requiring enzyme 1) and ATF6 (activating transcription factor 6), are the three principle transducers of the unfolded protein response (UPR).[16−18] The UPR is activated in response to the accumulation of unfolded or misfolded proteins in the endoplasmic reticulum (ER), due to ER stress arising from a number of conditions including glucose deprivation, hypoxia, oxidative stress, viral infection, high cholesterol, and protein mutations. An active UPR can restore homeostasis by increasing the capacity of the ER for protein folding and degradation while reducing protein synthesis; however, prolonged UPR activity, implying an unresolved ER stress, may lead to cell apoptosis, thus protecting the organism from the potential harmful consequences. The PERK arm of the UPR regulates protein levels entering the ER by phosphorylating the translation initiation factor eIF2$\alpha$, thereby reducing protein synthesis. PERK is activated by autophosphorylation through a poorly understood mechanism, which may involve oligomerization.

Recent studies have implicated the UPR in several human diseases, for example, protein-misfolding diseases, like retinitis pigmentosa[19] and type II diabetes,[20] where apoptosis signals from the UPR triggered by misfolded proteins cause the death of normal cells. Certain types of cancer[21,22] and viruses[23] exploit the UPR signal to increase the ER capacity in order to sustain the rapid growth of cancer cells or viral replication. Given the integral roles of PERK in the UPR, an understanding of its interactions with other proteins in the signaling pathways may inspire the development of potential therapeutic strategies. Recently, GlaxoSmithKline reported their first-in-class PERK inhibitor (GSK2606414).[24] Here we discuss the discovery of novel inhibitors of PERK utilizing virtual library screening approaches in hopes of providing new scaffolds for the development of PERK inhibitors.
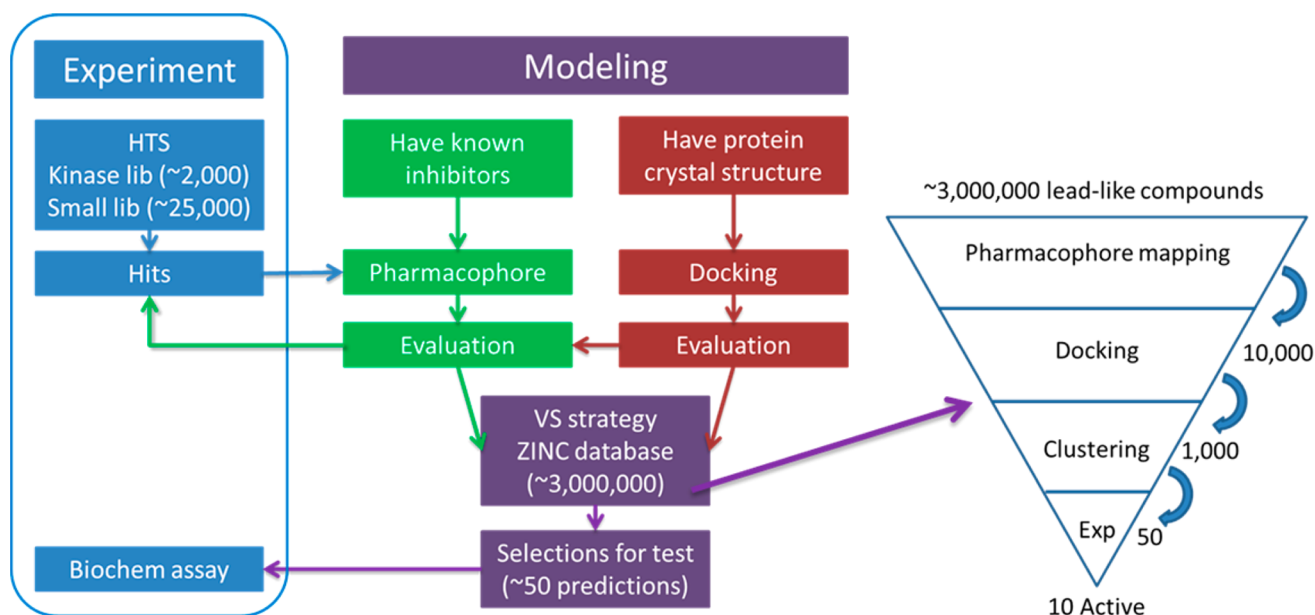
**Figure 1.** The schematic plot of the workflow of the screening process.

In this paper, we apply both structure-based docking and ligand-based screening approaches to identify potential novel inhibitors of PERK. We first discuss how MD simulations are necessary to refine a PERK crystal structure for docking-based virtual screening. Then we present a ligand-based pharmacophore model generated from four hits derived from high throughput screening (HTS). Both approaches are first validated against the HTS results of a screen against a library of about 27 000 compounds. The initial VS results suggest that a consensus approach by combining both pharmacophore modeling and docking are more effective than either one alone, which is in accordance with previous retrospective studies[25,26] on VEGFR-2 inhibitors using a number of combinations of VS methods. Our VS protocol is then applied to screen the ZINC lead-like database containing more than 3 million compounds. Finally, about 50 commercially available compounds from virtual screening were tested in biochemical kinase assays, confirming activities of 10.

## ■ METHOD

**Screening Work-Flow.** Two virtual screening approaches, ligand pharmacophore and docking, were used jointly. We first trained our protocol against previous high-throughput screening data[27] (the green and brown blocks in Figure 1). From the known active compounds obtained in the HTS, a ligand-based pharmacophore was generated and used to screen other potential compounds. Alternatively, we also performed protein structure-based docking to screen the compounds. The performance of both pharmacophore and docking were evaluated by comparing with the HTS result. On the basis of this, a protocol was proposed and applied to a VS of the ZINC database, which is the lower portion of the triangle shown in Figure 1. Finally, the selected compounds from the VS were tested *in vitro*.

**Structure Preparation.** The available apo mouse PERK structure (PDB code: 3QD2)[28] shows a closed G-loop when it is superimposed with a structurally similar kinase PKR (PDB code: 2A19).[29] It can be seen that the G-loop region in 3QD2 clashes with the ATP in 2A19 (Figure 2). With such clashing,
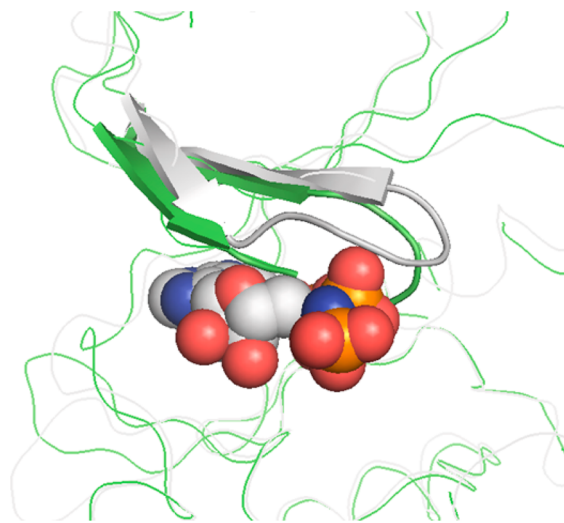


**Figure 2.** PERK (green) superimposed with PKR-ANP complex (silver).

the 3QD2 structure is not meaningful for docking. To obtain a PERK structure with an "open" active site, we first raised the G-loop region artificially by modeling after 2A19, then manually docked the ATP and two $Mg^{2+}$ ions into the ATP-binding site of 3QD2 (using 2A19 as the template). The mouse PERK-ATP complex was then solvated in an octahedral box of TIP3P water,[30] with a minimum buffer distance of 14 Å from the protein surface to the box edge. There are 12 095 water molecules in the box in total. Counter ions were also added to neutralize the box. Structural minimization was applied before running molecular dynamics simulation in order to remove any bad contacts between atoms. During the MD simulation, the system was heated up from 0 to 300 K in 200 ps with NVT ensemble and then switched to NPT ensemble for 10 ns. Positional restraints were applied to the two $Mg^{2+}$ ions, each with a weight of 5.0 kcal/mol/Å$^2$ in minimization and 2.0 kcal/mol/Å$^2$ in MD simulations. Both minimization and molecular

dynamics simulations were conducted with the Amber12 software package.[31]

A representative structure of PERK used in docking was obtained using the pairwise average linkage clustering method provided in the MaxCluster program.[32] A total of 450 snapshots (20 ps apart) were taken from the last 9 ns of MD simulation. RMSD of the protein structure was used as the measure of distance between two nodes in clustering, with a threshold value of 1.2 Å. A total of eight clusters are generated, and the median structure of the most populated cluster was chosen as the final model structure and used in subsequent docking work.

**Hit in Training Library.** A hit was defined as a compound demonstrating more than 50% inhibition at 1 $\mu$M concentration in the PERK kinase assay among a small library of 875 known kinase inhibitors. This yielded a total number of 15 hits. The remaining 860 compounds along with a larger library of 26 365 compounds were then considered as decoys or inactive compounds. Therefore, a library of 27 240 compounds was used in training the virtual screening.

**Docking and Pharmacophore Mapping.** A library of 27 240 compounds, including a known kinase-focused library, was processed by docking, using Gold5.0.1 and the goldscore scoring function.[33] For each compound, 10 GA runs were performed with a docking efficiency of 100%. It took approximately 3 days to dock the whole library on 20 2.4 GHz AMD Opteron cores for each target.

A ligand-based pharmacophore was generated and utilized in pharmacophore mapping using DiscoveryStudio3.5. A 3D database of the library was built first, with 255 conformations generated for each compound. The pharmacophore was generated based on the four most potent hits we found in the experimental high-throughput screening assay. A number of pharmacophores were first generated based on each of the compounds. Then each pharmacophore was examined by mapping it against all four compounds. The best-fitted pharmacophore was hence selected. This led to a five-feature pharmacophore, including two hydrogen-acceptor features and three hydrophobic features. An additional aromatic ring feature was added manually afterward to mimic the adenine ring of ATP. The most time-consuming part of pharmacophore mapping is the conformation-building step. Mapping the whole 3D database of 27 240 compounds (255 conformation each) took only 4 min on a workstation of four 2.4 GHz Intel Xeon cores.

**Enrichment Calculation.** Enrichment is defined as

$$\text{Enrichment} = \frac{\% \text{ hits}}{\% \text{ library}} \quad (1)$$

where percentage of hits means the percentage of the 15 true hits found by docking, while percentage of library indicates the percentage of the total number of compounds in the library.

Two other measures, true positive rate and false positive rate, used in the receiver operating characteristic (ROC) plot in this study, are defined as

$$\text{True positive rate (\% hits)} = \frac{\text{Hits in docking result}}{\text{All hits}} \quad (2)$$

$$\text{False positive rate (\% decoys)} = \frac{\text{Decoys in docking result}}{\text{All decoys}} \quad (3)$$

**Biochemical Screening of the Compounds Identified by Virtual Screening.** After virtual screening of the ZINC database, 50 commercially available compounds were purchased and assessed. PERK kinase activity assay was performed in 96-well microplates (OptiPlate-96, PerkinElmer LAS, Inc.). The reaction had a total volume of 100 $\mu$L, containing 25 mM HEPES (pH 7.5), 10 mM MgCl$_2$, 50 mM KCl, 2 mM DTT, 0.1 mM EGTA, 0.1 mM EDTA, 0.03% Brij 35, 5% DMSO, and 10 $\mu$g/mL BSA. The activity of 20 nM PERK was tested against 5 $\mu$M of eIF2alpha. Each reaction mixture was incubated in a 96 well plate at room temperature for 30 min. The reaction was initiated by the addition of 10 $\mu$L [$\gamma$-$^{32}$P] ATP, adjusting the final ATP concentration to 10 $\mu$M. The reaction was incubated at room temperature for 10 min and then quenched by transferring 80 $\mu$L of reaction mixture to each well of a P81 96-well filter plate (Unifilter, Whatman) containing 0.1 M phosphoric acid. The P81 filter plate was washed with 0.1 M phosphoric acid thoroughly, followed by the addition of a scintillation cocktail. A MicroBeta TriLux liquid scintillation counter (PerkinElmer) was used for screening plates. The top 4 compounds were further tested in a dose−response experiment under the same conditions. The inhibition of PERK activity was determined by measuring initial velocities in the presence of varying concentrations of four compounds.

## ■ RESULT

**Model Structure VS apo Structure in Docking.** As noted earlier, the apo PERK structure shows a closed G-loop region (Figure 2). Superimposing the MD-refined PERK structure with the original crystal structure clearly shows that residues Gly18 and Phe19 in the apo structure block the gate of the ATP-binding pocket (Figure 3a), and that they were lifted away in the MD refined structure. The RMSD plot of non-
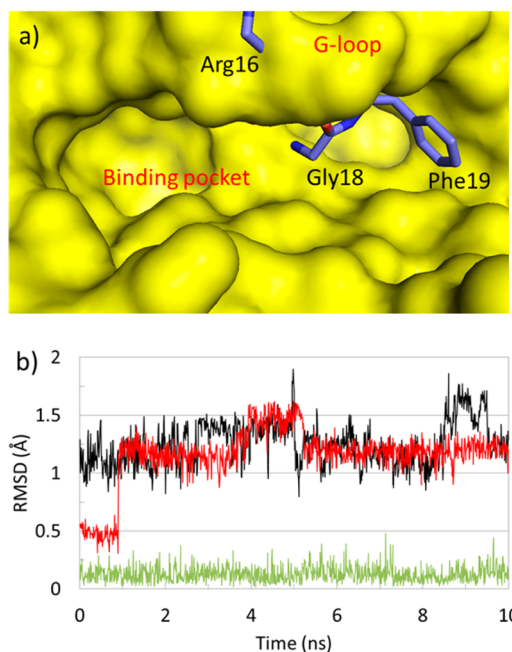


**Figure 3.** Plots of PERK: (a) Comparison of MD-refined PERK (yellow surface) with the crystal structure (blue) reveals that residues Gly18 and Phe19 blocked the gate of the ATP-binding pocket. (b) RMSD plot of the non-hydrogen atoms in Arg16 (dark), Gly18 (green), and Phe19 (red) from the MD simulation.
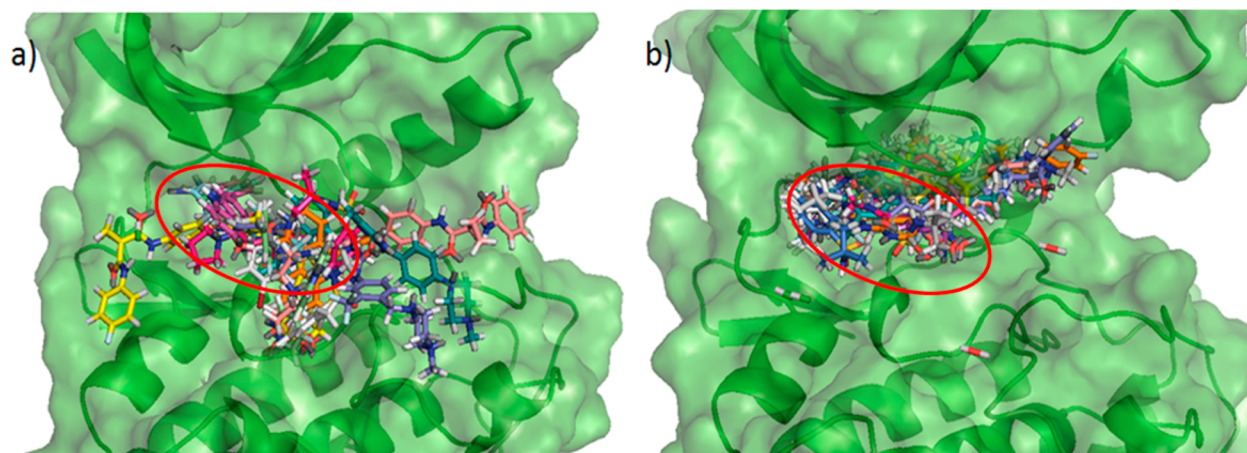
**Figure 4.** Docked compounds in (a) the apo PERK structure and (b) the modeled "open" structure. Compounds in apo structure show a more random likely distribution around the ATP-binding pocket (red cycle) due to the closure of the G- loop region, while the MD-refined model structure gives more meaningful docking poses in the pocket in general.

hydrogen atoms of Phe19 over time (aligned by backbone atoms), for example, also indicates a notable rearrangement of not only the backbone but also the side chain of the residue (Figure 3b).

This closure, observed in the apo PERK structure, blocks the binding of compounds at the ATP-binding site. A simple illustration is shown in Figure 4a, where a number of top ranked compounds from the docking with the apo and model structures are shown. The docked compounds in the apo structure appear to be more randomly distributed around the binding site in comparison with those docked "deeply" into the model structure as shown in Figure 4b. This simple visual comparison between the docking results of the apo and model structure demonstrates how unreliable the results could be using the apo structure. A previous study on cyclin-dependent kinase 2 using goldscore and chemscore has demonstrated that the enrichment in docking-based virtual screening is related to the quality of the binding poses predicted,[34] thus it is important to ensure that the docked poses are reasonable in VS.

To achieve a quantitative understanding of the docking results, statistical measures like enrichment factor and ROC plot were calculated by comparing the *in silico* results with the *in vitro* results. Presumably, if the same docking protocol was used in docking, a better kinase structure would give a better prediction. Thus, we ranked the docked compounds, and at select rankings we calculated the ratio of identified hits to total (15) hits. This value is defined as the true positive rate (eq 2). We also calculated the ratio of the decoys to the total number of decoys to obtain the false positive rate (eq 3). Plotting the true positive rate against the false positive rate gives the so-called ROC plot (Figure 5). A steeper curve in the ROC plot indicates a better prediction of the true hits against the decoys. Generally, the modeled PERK structure gives a line that is significantly above the line of the apo structure. When 50% of the hits are found from the top ranked compounds, only 6% of the decoys are picked up using the model structure. In contrast, 22% of the decoys are picked up by the docking using apo structure when 50% of the hits are found, resulting in a 4-fold performance boost with the model structure. Another indicator of the predictive power called area under the curve (AUC) was also calculated. The AUC was measured to be 0.90 and 0.75 for the model and apo PERK, respectively. Given that a value of 0.5 means a random result with no selectivity and a value of 1.0 for
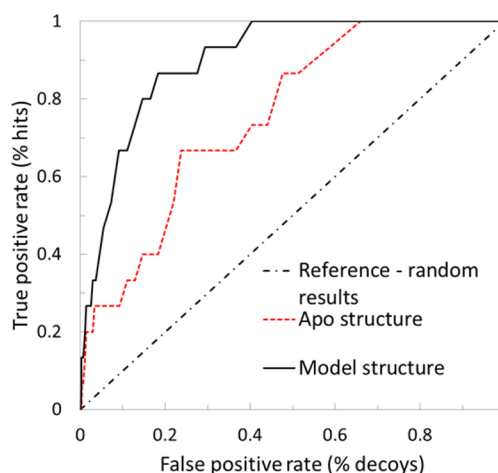


**Figure 5.** The ROC plot of the docking results using the apo (red dashed line) and model (dark solid line) structures. The steeper the line, the better predictive power the model has (a perfect line should have an AUC value of 1.00), and the diagonal dark dash dot line indicates random results (with an AUC value of 0.50). The calculated AUC values are 0.90 and 0.75 for the model and apo structures, respectively.

a perfect model, the docking result using our model structure is indeed noticeably better than that of the apo structure.

Other than the ROC plot, the enrichment of the top 20% of docking results is also examined. The model structure generally doubles the enrichment of the apo structure in the top 20% of the ranked library, i.e., the chance of finding one of the 15 hits in the top 20% prediction increased twice from the apo to the model structure (Table 1). Using the apo structure, the first hit was found in the top 100 compounds. In contrast, the first hit was captured in the top 60 compounds, and a total of two hits were captured in the top 100. For comparison, in an earlier virtual screening study[2] using the cocrystal structure of FGFR1 kinase, an enrichment of about 8 was reported when the top 1000 compounds were selected from the docked library of about 40 000 compounds, including 41 actives, which is comparable to our enrichment of 7.3 and 9.1 for apo and model structures, respectively, when the same number of compounds are selected, respectively. Our docking results are also comparable with another bench mark study[35] using the

**Table 1. Enrichment of Different VS Protocols (the Actual Numbers of True Hits Found in VS Are Shown in the Parentheses)**

| number of top ranked compounds (% library) | 50 (<1%) | 100 (<1%) | 500 (~2%) | 1000 (~4%) | 2000 (~7%) | 4000 (~15%) | 5138 (~20%) |
|---|---|---|---|---|---|---|---|
| docking (apo) | 0 (0) | 18.2 (1) | 10.9 (3) | 7.3 (4) | 3.6 (4) | 2.7 (6) | 2.5 (7) |
| docking (model) | 0 (0) | 36.3 (2) | 14.5 (4) | 9.1 (5) | 7.3 (8) | 5.4 (12) | 4.6 (13) |
| pharmacophore mapping | 36.3 (1) | 18.2 (1) | 10.9 (3) | 10.9 (6) | 7.3 (8) | 5.0 (11) | 4.6 (13) |
| pharmacophore + docking (model) | 72.6 (1) | 36.3 (2) | 14.5 (4) | 14.5 (8) | 9.1 (10) | 5.4 (12) | 4.6 (13) |
| docking (model)[a] | 0.0 (0) | 24.8 (1) | 9.9 (2) | 7.4 (3) | 6.2 (5) | 5.0 (8) | 4.3 (9) |
| pharmacophore mapping[a] | 0.0 (0) | 0.0 (0) | 9.9 (2) | 7.4 (3) | 6.2 (5) | 4.3 (7) | 4.3 (9) |
| pharmacophore + docking (model)[a] | 49.5 (1) | 24.8 (1) | 9.9 (2) | 12.4 (5) | 7.4 (6) | 5.0 (8) | 4.3 (9) |

[a]The four hits used to generate the pharmacophore model are excluded in statistics.

DUD data set[36] with different scoring functions, including ChemScore, ChemGauss, and PLP, which can identify about 20−33% of the true hits within the top 5% of the ranked libraries against kinase targets like ABL, EGFR, P38, and VEGFR2. The docking finds about 27% and 33% of the true hits when about 4% of the ranked compounds in the library are selected using the apo and model structures respectively in our study (Table 1).

**Pharmacophore As a Virtual Screening Filter.** As an alternative to docking, a ligand-based pharmacophore model was also generated by using DiscoveryStudio3.5 based on four hits identified in the biochemical assay. The automated process generated a five-feature pharmacophore with two hydrogen acceptor features and three hydrophobic features. An additional aromatic ring feature was added manually for the purpose of mimicking the adenine region of ATP, making it a six-feature pharmacophore (Figure 6). This is similar to a generic kinase
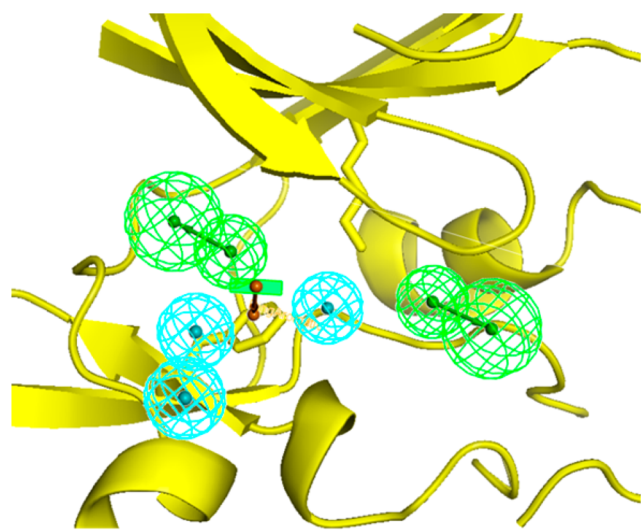


**Figure 6.** A ligand-based six-feature pharmacophore manually overlaid within the PERK active site (cartoon in yellow). Green arrows and spheres represent hydrogen acceptor features. Blue spheres represent hydrophobic features, and a brown arrow with a green base means an aromatic ring feature.

inhibitor pharmacophore reported,[37] which divided the ATP-binding site into five regions, the adenine region, the sugar pocket, the phosphate binding region, and the hydrophobic regions I and II. For each screened compound, the highest fit value among the 255 conformations was assigned as the final fit value. Compounds with fit values of less than zero were considered as inactive, thus were ignored. This collected 5138 compounds, which represents roughly 20% of the whole training library. These compounds were then ranked according

to their fit value. It is noted that among the 5138 compounds, 9 out of 11 hits (the four hits used to generate the pharmacophore are excluded) were found, which is comparable with the docking result (Table 1). The corresponding ROC plot is shown in Figure 7 as "pharmacophore mapping;" the top
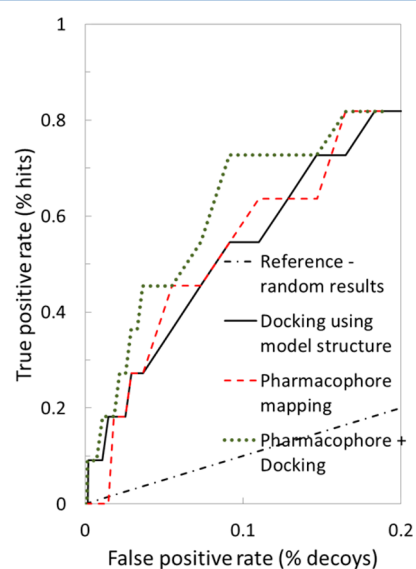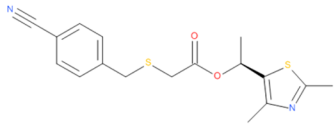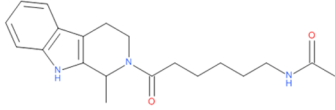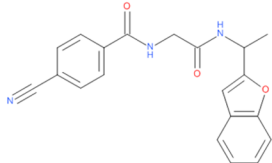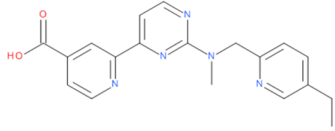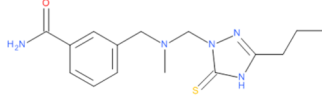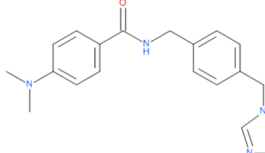


**Figure 7.** Comparison of ROC plots from different approaches. Note that the four hits employed to generate the pharmacophore model are removed in all data. Docking with the model structure is shown in dark solid line. The red dashed line indicates the pharmacophore mapping approach. The green dotted lines represents the consensus model of pharmacophore mapping and docking, and the dark dotted dash line is the random reference. The respect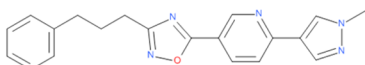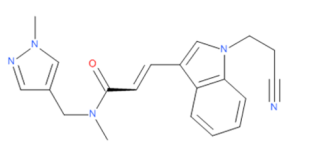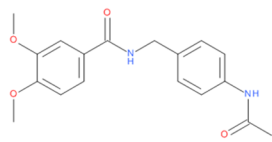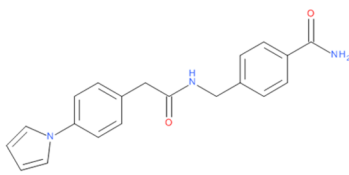ive AUC value for each line is 0.09, 0.09, 0.11, and 0.04 while a value of 0.20 is for a perfect model and 0.04 for random selections.

20% of docking results using our model structure is also presented in the same plot as "docking using model structure." The two curves generally overlap with each other. The AUC value was measured as 0.09 and 0.09 for curves of pharmacophore and docking, respectively. Note that the x-axis in this plot was truncated at 0.2, thus an AUC value of 0.2 corresponds to a perfect model while a value of 0.04 represents random results. This suggests that pharmacophore mapping and docking performed equally well in this particular study.

It has been suggested that utilizing the ligand-biased receptor-based virtual screening could lead to better enrichment if a cocrystal structure is known.[38] Other studies also suggested that using combinations of docking and similarity-based approaches can increase the enrichment of VS.[25,26] Thus, we explored the potential of a pharmacophore-based approach to facilitate a receptor-structure-dependent docking method.

**Table 2. Ten Compounds (out of 50 from Virtual Screening) Confirmed to Be Active in the Biochemistry Assay[a]**

| | % of Inhibition | IC$_{50}$ (μM) | Structure |
|---|---|---|---|
| Compound 1 | 74% | 2.6 | |
| Compound 2 | 60% | 8.7 | |
| Compound 3 | 72% | 36 | |
| Compound 4 | 70% | 36 | |
| Compound 5 | 72% | / | |
| Compound 6 | 57 % | / | |
| Compound 7 | 53 % | / | |
| Compound 8 | 51 % | / | |
| Compound 9 | 49 % | / | |
| Compound 10 | 48 % | / | |

[a]The assay condition is set to have, for each 100 μL well, 20 nM of active PERK, 5 μM of EIF2α, 25 μM of the compounds, and 10 μM of radiolabeled ATP.

We reranked the 5138 selected compounds collected by pharmacophore mapping, using docking. This combination yielded a slightly better result than using either docking or pharmacophore mapping alone. This is supported by a bigger AUC value of 0.11, as well as a ROC curve that is always above the curves of either docking or pharmacophore mapping (Figure 7). To capture 50% of the hits, only 3% of the decoys were picked up in the combined approach while 6% of decoys were picked to obtain the same amount of hits in structure-based docking. Additionally the first captured hit was among the top 40 compounds while the first hit in docking was found in the top 60 compounds.

Comparing the consensus approach with the pharmacophore mapping, it is noted that the former has significantly moved the 15 true hits higher in ranking. For instance, the enrichment obtained in the consensus approach is generally 50−100% more than that obtained in pharmacophore mapping (Table 1). However, since the pharmacophore model was generated based on four hits, there is likely a bias toward these compounds in pharmacophore-based screening. In order to make a fair comparison, we thus removed all four hits in the result and calculated the enrichment again. Not surprisingly, even if we excluded the four hits from the result of the consensus approach, there was still one hit left in the top 50 compounds. However, neither docking nor pharmacophore model happen to predict any hit at this range. Furthermore, if we look at the top 1000, 2000, or 4000 ranked compounds by VS, the consensus approach always returns more hits than the other two approaches. Therefore, the consensus approach of combining a pharmacophore model with structure-based docking can be a better choice than using either approach alone, and a better enrichment can be expected. The possible reason behind this may be due to the fact that docking and ligand-based pharmacophore approaches explore different chemical/physical spaces, i.e. docking depends on the structure of a receptor as well as the ligands while the pharmacophore is solely dependent on the ligand. If designed carefully, the two approaches could complement each other in a consensus scheme.

**Screening of the ZINC Database.** On the basis of the success of our virtual screening protocol in the training library, we then applied it to screen the ZINC lead-like database, which includes about 3 million lead-like compounds. We first imported the database into DiscoveryStudio, which then generated a conformational library consisting of 255 conformations for each compound in the database. Subsequently, all conformations of the 3 million compounds were screened by mapping to the pharmacophore. The highest scored conformation for each compound was selected and then used to rank the 3 million compounds. The top 10 000 compounds were then selected and subjected to structure-based docking. The same docking procedure was used. The top 10% of the docked compounds, i.e. the top 1000, were then further profiled by clustering them into 100 clusters in order to filter out similar compounds in the library. We anticipate this may increase the chemical diversity in a smaller pool of selections. The center compound of each cluster was selected as the representative compound for that cluster. Another possible advantage of this is that if any center compound shows promising activity, we could come back and investigate more compounds in that cluster.

We purchased 50 commercially available compounds out of the 100 representative compounds and then tested them in a kinase assay. For each 50 $\mu$L well, there were 20 nM of active PERK, 5 $\mu$M of eIF2$\alpha$, 25 $\mu$M of the compounds, and 10 $\mu$M of ATP. The initial assay shows 10 active compounds exhibiting more than 50% inhibition (Table 2). All 10 compounds fit our pharmacopore well. As an example, the overlay of compound 6 with the pharmacophore is shown in Figure 8. Then dose responses were obtained for four of the compounds. The IC$_{50}$ of two of the compounds is less than 10 $\mu$M (2.6 and 8.7 $\mu$M, respectively).
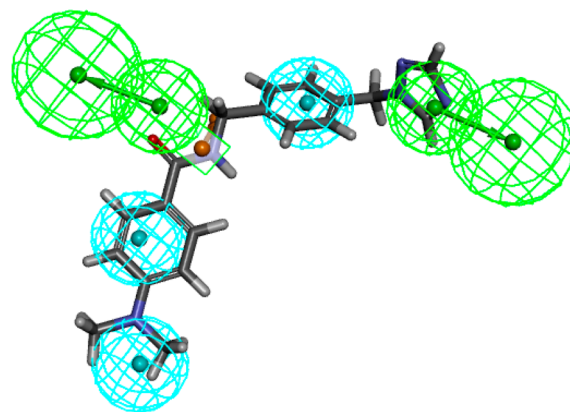


**Figure 8.** Overlay of compound 6 and the pharmacophore. Green indicates a hydrogen bond feature, and blue suggests a hydrophobic feature.

## CONCLUSION

The docking-based virtual screening method normally requires a quality crystal structure, which may not always be available. In the case of PERK, the only available crystal structure, at the time of our study, was an apo structure. The structure presents a closed G-loop region at the ATP-binding site of the kinase, which hinders the docking of the inhibitors into the ATP-binding site. We examined two approaches that can potentially resolve the issue. In our first approach, by using another kinase PKR as a template, we artificially lifted the G-loop so that the ATP-binding site could accommodate an ATP molecule. MD simulation was applied to relax the system to obtain a model cocrystal structure of PERK with an "open" active site. Applying the model structure in virtual library docking yielded a significantly improved enrichment, generally twice the enrichment of using the apo crystal structure. This in turn suggests that a well-modeled protein structure can be a better target than an inadequate crystal structure in structure-based docking.

The other approach we investigated was the ligand-based pharmacophore mapping. On the basis of the four inhibitory compounds found in experimental kinase assays, a six-feature pharmacophore was built using the ligand-based pharmacophore generation module in DiscoveryStudio3.5. Using the pharmacophore mapping method to screen the same library shows a similar performance to the docking method in terms of enrichment. However, in regards to efficiency, pharmacophore mapping is 1000 fold faster than docking, given the fact that docking of about 27 000 compounds took about 3 days using 20 2.4 GHz AMD Opteron cores, while pharmacophore mapping (with a prebuilt conformational database of all the compounds) only took 4 min on four 2.4 GHz Intel Xeon cores. Reranking the pharmacophore mapping results using the docking scores shows a slightly better prediction than using docking or pharmacophore mapping alone. The reranked result generally predicts more hits in the upper region of the ranking list, thus showing a higher probability of finding a hit in a smaller number of the ranked compounds. The improvement though is not enormous, yet notably enough to validate our argument. The limitation of the pharmacophore approach, however, is that some known inhibitors must be available. This may be achievable by acquiring information from the literature or by performing experiments, but within an affordable scale. With the huge saving in resources and the competitive accuracy,

the pharmacophore approach can serve as a cost-effective predocking filter for virtual library screening.

Upon our preliminary study of the combination of docking and pharmacophore modeling, we proposed a consensus virtual screening approach which uses pharmacophore mapping as a fast filter to generate a much reduced compound pool for docking, then makes the final decision based on the docking result (Figure 1). This consensus approach was then applied to screen the ZINC lead-like database,[39,40] which includes about 3 million compounds. On the basis of the VS using the consensus approach, we purchased 50 compounds to test them in vitro. Ten out of 50 compounds show activity while two exhibit an IC$_{50}$ of less than 10 $\mu$M, which further provides validity of this consensus approach. We anticipate that more potent compounds, i.e. subnanomolar IC$_{50}$, may be found if a more kinase specific library was screened.

## ■ AUTHOR INFORMATION

### Corresponding Authors
*E-mail: dalby@austin.utexas.edu.
*E-mail: pren@mail.utexas.edu.

### Notes
The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

## ■ REFERENCES

(1) Hu, X.; Compton, J. R.; AbdulHameed, M. D. M.; Marchand, C. L.; Robertson, K. L.; Leary, D. H.; Jadhav, A.; Hershfield, J. R.; Wallqvist, A.; Friedlander, A. M.; Legler, P. M. 3-Substituted Indole Inhibitors Against Francisella tularensis FabI Identified by Structure-Based Virtual Screening. *J. Med. Chem.* 2013, 56, 5275−5287.

(2) Ravindranathan, K. P.; Mandiyan, V.; Ekkati, A. R.; Bae, J. H.; Schlessinger, J.; Jorgensen, W. L. Discovery of novel fibroblast growth factor receptor 1 kinase inhibitor by structure-based virtual screening. *J. Med. Chem.* 2010, 53, 1662−1672.

(3) Teli, M. K.; Rajanikant, G. K. Computational Repositioning and Experimental Validation of Approved Drugs for HIF-Prolyl Hydroxylase Inhibition. *J. Chem. Inf. Model.* 2013, 53, 1818−1824.

(4) Sahner, J. H.; Groh, M.; Negri, M.; Haupenthal, J.; Hartmann, R. W. Novel small molecule inhibitors targeting the "switch region" of bacterial RNAP: Structure-based optimization of a virtual screening hit. *Eur. J. Med. Chem.* 2013, 65, 223−231.

(5) Rea, V. E. A.; Lavecchia, A.; Di Giovanni, C.; Rossi, F. W.; Gorrasi, A.; Pesapane, A.; de Paulis, A.; Ragno, P.; Montuori, N. Discovery of new small molecules targeting the vitronectin binding site of the urokinase receptor that block cancer cell invasion. *Mol. Cancer Ther.* 2013, 12, 1402−1416.

(6) Kaoud, T. S.; Yan, C.; Mitra, S.; Tseng, C.-C.; Jose, J.; Taliaferro, J. M.; Tuohetahuntila, M.; Devkota, A.; Sammons, R.; Park, J.; Park, H.; Shi, Y.; Hong, J.; Ren, P.; Dalby, K. N. From in Silico Discovery to Intracellular Activity: Targeting JNK−Protein Interactions with Small Molecules. *ACS Med. Chem. Lett.* 2012, 3, 721−725.

(7) Taylor, S. S.; Kornev, A. P. Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem. Sci.* 2011, 36, 65−77.

(8) Schnieders, M. J.; Kaoud, T. S.; Yan, C.; Dalby, K. N.; Ren, P. Computational insights for the discovery of non-ATP competitive inhibitors of MAP kinases. *Curr. Pharm. Des.* 2012, 18, 1173−1185.

(9) Yan, C.; Kaoud, T.; Lee, S.; Dalby, K. N.; Ren, P. Understanding the specificity of a docking interaction between JNK1 and the scaffolding protein JIP1. *J. Phys. Chem. B* 2011, 115, 1491−1502.

(10) Wang, H.; Blais, J.; Ron, D.; Cardozo, T. Structural determinants of PERK inhibitor potency and selectivity. *Chem. Biol. Drug Des.* 2010, 76, 480−495.

(11) Wang, J.; Chen, L.; Sinha, S. H.; Liang, Z.; Chai, H.; Muniyan, S.; Chou, Y.-W.; Yang, C.; Yan, L.; Feng, Y.; Li, K. K.; Lin, M.-F.; Jiang, H.; Zheng, Y. G.; Luo, C. Pharmacophore-based virtual screening and biological evaluation of small molecule inhibitors for protein arginine methylation. *J. Med. Chem.* 2012, 55, 7978−7987.

(12) Singh, J.; Chuaqui, C. E.; Boriack-Sjodin, P. A.; Lee, W.-C.; Pontz, T.; Corbley, M. J.; Cheung, H.-K.; Arduini, R. M.; Mead, J. N.; Newman, M. N.; Papadatos, J. L.; Bowes, S.; Josiah, S.; Ling, L. E. Successful shape-based virtual screening: the discovery of a potent inhibitor of the type I TGFbeta receptor kinase (TbetaRI). *Bioorg. Med. Chem. Lett.* 2003, 13, 4355−4359.

(13) Chao, W.-R.; Yean, D.; Amin, K.; Green, C.; Jong, L. Computer-aided rational drug design: a novel agent (SR13668) designed to mimic the unique anticancer mechanisms of dietary indole-3-carbinol to block Akt signaling. *J. Med. Chem.* 2007, 50, 3412−3415.

(14) De Luca, L.; Ferro, S.; Damiano, F. M.; Supuran, C. T.; Vullo, D.; Chimirri, A.; Gitto, R. Structure-based screening for the discovery of new carbonic anhydrase VII inhibitors. *Eur. J. Med. Chem.* 2014, 71, 105−111.

(15) Yang, S.-Y. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discovery Today* 2010, 15, 444−450.

(16) Walter, P.; Ron, D. The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. *Science* 2011, 334, 1081−1086.

(17) Tsai, Y. C.; Weissman, A. M. The Unfolded Protein Response, Degradation from the Endoplasmic Reticulum, and Cancer. *Genes Cancer* 2010, 1, 764−778.

(18) Tabas, I.; Ron, D. Integrating the mechanisms of apoptosis induced by endoplasmic reticulum stress. *Nat. Cell Biol.* 2011, 13 (3), 184−190.

(19) Lin, J.; LaVail, M., Misfolded Proteins and Retinal Dystrophies. In *Retinal Degenerative Diseases*; Anderson, R. E., Hollyfield, J. G., LaVail, M. M., Eds.; Springer: New York, 2010; Vol. 664, pp 115−121.

(20) Fonseca, S. G.; Gromada, J.; Urano, F. Endoplasmic reticulum stress and pancreatic $\beta$-cell death. *Trends Endocrinol. Metab.* 2011, 22 (7), 266−274.

(21) Carrasco, D. R.; Sukhdeo, K.; Protopopova, M.; Sinha, R.; Enos, M.; Carrasco, Daniel E.; Zheng, M.; Mani, M.; Henderson, J.; Pinkus, G. S.; Munshi, N.; Horner, J.; Ivanova, E. V.; Protopopov, A.; Anderson, K. C.; Tonon, G.; DePinho, R. A. The Differentiation and Stress Response Factor XBP-1 Drives Multiple Myeloma Pathogenesis. *Cancer Cell* 2007, 11 (4), 349−360.

(22) Papandreou, I.; Denko, N. C.; Olson, M.; Van Melckebeke, H.; Lust, S.; Tam, A.; Solow-Cordero, D. E.; Bouley, D. M.; Offner, F.; Niwa, M.; Koong, A. C. Identification of an Ire1alpha endonuclease specific inhibitor with cytotoxic activity against human multiple myeloma. *Blood* 2011, 117 (4), 1311−1314.

(23) Li, B.; Gao, B.; Ye, L.; Han, X.; Wang, W.; Kong, L.; Fang, X.; Zeng, Y.; Zheng, H.; Li, S.; Wu, Z.; Ye, L. Hepatitis B virus X protein (HBx) activates ATF6 and IRE1-XBP1 pathways of unfolded protein response. *Virus Res.* 2007, 124, 44−49.

(24) Axten, J. M.; Medina, J. R.; Feng, Y.; Shu, A.; Romeril, S. P.; Grant, S. W.; Li, W. H. H.; Heerding, D. A.; Minthorn, E.; Mencken, T.; Atkins, C.; Liu, Q.; Rabindran, S.; Kumar, R.; Hong, X.; Goetz, A.; Stanley, T.; Taylor, J. D.; Sigethy, S. D.; Tomberlin, G. H.; Hassell, A. M.; Kahler, K. M.; Shewchuk, L. M.; Gampe, R. T. Discovery of 7−Methyl-5-(1-{[3-(trifluoromethyl)phenyl]acetyl}-2,3-dihydro−1H−indol-5-yl)−7H−pyrrolo[2,3−d]pyrimidin-4-amine (GSK2606414), a Potent and Selective First-in-Class Inhibitor of Protein Kinase R (PKR)-like Endoplasmic Reticulum Kinase (PERK). *J. Med. Chem.* 2012, 55, 7193−7207.

(25) Planesas, J. S. M.; Claramunt, R. M.; Teixidó, J.; Borrell, J. I.; Pérez-Nueno, V. I. Improving VEGFR-2 Docking-Based Screening by Pharmacophore Postfiltering and Similarity Search Postprocessing. *J. Chem. Inf. Model.* 2011, 51 (4), 777−787.

(26) Zhang, Y.; Yang, S.; Jiao, Y.; Liu, H.; Yuan, H.; Lu, S.; Ran, T.; Yao, S.; Ke, Z.; Xu, J.; Xiong, X.; Chen, Y.; Lu, T. An Integrated Virtual Screening Approach for VEGFR-2 Inhibitors. *J. Chem. Inf. Model.* **2013**, *53*, 3163−3177.

(27) Park, J.; Dalby, K. N. High through put screening of PERK inhibitors. Unpublished work. College of Pharmacy, University of Texas at Austin: Austin, TX, 2012.

(28) Cui, W.; Li, J.; Ron, D.; Sha, B. The structure of the PERK kinase domain suggests the mechanism for its activation. *Acta Crystallogr., Sect.D* **2011**, *67*, 423−428.

(29) Dar, A. C.; Dever, T. E.; Sicheri, F. Higher-order substrate recognition of eIF2alpha by the RNA-dependent protein kinase PKR. *Cell* **2005**, *122*, 887−900.

(30) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(31) Case, D. A.; Darden, T. A.; Cheatham, I. T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER12*; University of California: San Francisco, CA, 2012.

(32) Siew, N.; Elofsson, A.; Rychlewski, L.; Fischer, D. MaxSub: an automated measure for the assessment of protein structure prediction quality. *Bioinformatics* **2000**, *16*, 776−785.

(33) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267* (3), 727−748.

(34) Verdonk, M. L.; Berdini, V.; Hartshorn, M. J.; Mooij, W. T. M.; Murray, C. W.; Taylor, R. D.; Watson, P. Virtual screening using protein-ligand docking: avoiding artificial enrichment. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 793−806.

(35) Dixit, A.; Verkhivker, G. M. Integrating Ligand-Based and Protein-Centric Virtual Screening of Kinase Inhibitors Using Ensembles of Multiple Protein Kinase Genes and Conformations. *J. Chem. Inf. Model.* **2012**, *52* (10), 2501−2515.

(36) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking Sets for Molecular Docking. *J. Med. Chem.* **2006**, *49* (23), 6789−6801.

(37) Traxler, P.; Furet, P. Strategies toward the design of novel and selective protein tyrosine kinase inhibitors. *Pharmacol. Ther.* **1999**, *82*, 195−206.

(38) Cross, S.; Baroni, M.; Carosati, E.; Benedetti, P.; Clementi, S. FLAP: GRID Molecular Interaction Fields in Virtual Screening. Validation using the DUD Data Set. *J. Chem. Inf. Model.* **2010**, *50* (8), 1442−1450.

(39) Irwin, J. J.; Shoichet, B. K. ZINC - A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45* (1), 177−182.

(40) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: A free tool to discover chemistry for biology. *J. Chem. Inf. Model.* **2012**, *52* (7), 1757−1768.