JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

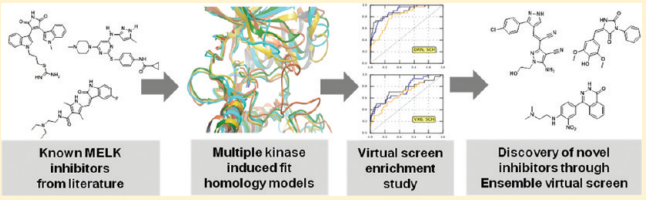
# Novel Inhibitor Discovery through Virtual Screening against Multiple Protein Conformations Generated via Ligand-Directed Modeling: A Maternal Embryonic Leucine Zipper Kinase Example

Kiran V. Mahasenan and Chenglong Li*

Division of Medicinal Chemistry and Pharmacognosy, College of Pharmacy, The Ohio State University, Columbus, Ohio 43210, United States

Supporting Information

**ABSTRACT:** Kinase targets have been demonstrated to undergo major conformational reorganization upon ligand binding. Such protein conformational plasticity remains a significant challenge in structure-based virtual screening methodology and may be approximated by screening against an ensemble of diverse protein conformations. Maternal embryonic leucine zipper kinase (MELK), a member of serine-threonine kinase family, has been recently found to be involved in the tumerogenic state of glioblastoma, breast, ovarian, and colon cancers. We therefore modeled several conformers of MELK utilizing the available chemogenomic and crystallographic data of homologous kinases. We carried out docking pose prediction and virtual screening enrichment studies with these conformers. The performances of the ensembles were evaluated by their ability to reproduce known inhibitor bioactive conformations and to efficiently recover known active compounds early in the virtual screen when seeded with decoy sets. A few of the individual MELK conformers performed satisfactorily in reproducing the native protein−ligand pharmacophoric interactions up to 50% of the cases. By selecting an ensemble of a few representative conformational states, most of the known inhibitor binding poses could be rationalized. For example, a four conformer ensemble is able to recover 95% of the studied actives, especially with imperfect scoring function(s). The virtual screening enrichment varied considerably among different MELK conformers. Enrichment appears to improve by selection of a proper protein conformation. For example, several holo and unliganded active conformations are better to accommodate diverse chemotypes than ATP-bound conformer. These results prove that using an ensemble of diverse conformations could give a better performance. Applying this approach, we were able to screen a commercially available library of half a million compounds against three conformers to discover three novel inhibitors of MELK, one from each template. Among the three compounds validated via experimental enzyme inhibition assays, one is relatively potent (**15**; $K_d$ = 0.37 $\mu$M), one moderately active (**12**; $K_d$ = 3.2 $\mu$M), and one weak but very selective (**9**; $K_d$ = 18 $\mu$M). These novel hits may be utilized to assist in the development of small molecule therapeutic agents useful in diseases caused by deregulated MELK, and perhaps more importantly, the approach demonstrates the advantages of choosing an appropriate ensemble of a few conformers in pursuing compound potency, selectivity, and novel chemotypes over using single target conformation for structure-based drug design in general.




## INTRODUCTION

Structure-based virtual screening (SBVS) methods, especially those utilizing molecular docking, have been gaining popularity in the past decade.[1] However, enrichment of current structure-based virtual screens through molecular docking suffers because of several reasons including inaccurate scoring functions and failure to incorporate protein flexibility.[2] Ligand-induced protein flexibility has been well established.[3] This plasticity can vary from local side chain adjustments to major domain movements. Even minor side chain motion in a binding site can have a substantial consequence on docking affinity calculation and pose prediction. In routine docking calculations, such conformational changes are ignored because of the computational complexity of the calculations involved. No matter how good the ligand conformational sampling method or scoring function is a correct binding pose could potentially never be found without sampling alternate receptor poses.[4] Although there has been promising progress in multicore, multithread computational infrastructure in the recent past, explicitly addressing protein flexibility remains a major challenge in structure-based in silico screening, especially when dealing with huge chemical libraries.[5] In actual biological systems, protein and ligand move together; hence, we should ideally simulate the movement of both protein and ligand simultaneously.[6] Unfortunately, such a method could be computationally resource intensive with the current infrastructure and thus unrealistic for high-throughput docking calculations. Consequently, a number of streamlined schemes have been explored to incorporate protein flexibility to balance the computational cost verses accuracy paradigm. One of the popular methods makes use of multiple conformations of the protein for

**Received:** January 22, 2012
**Published:** April 29, 2012

   dx.doi.org/10.1021/ci300040c | J. Chem. Inf. Model. 2012, 52, 1345−1355

molecular docking, commonly referred as "ensemble docking", the "relaxed complex method", etc.[7−10] The initiative here is to obtain several possible states of the protein in the folded dynamic conformational equilibrium and use all of them simultaneously for flexible ligand docking. The simulation is expected to sufficiently sample the conformational space of the target protein, and the desired ligand binding configuration is captured within this space. Alternatively, in a practical scenario, such induced fit plasticity may be accounted for through the utilization of multiple X-ray crystal structures of the protein of interest with diverse ligand bound states for the ensemble docking.[11−13] When such structural information is unavailable, it is possible that ligand-induced conformations may be modeled on the basis of homologous structures. For example, if a small molecule A inhibits two kinases X and Y with low micromolar or better experimental binding affinity, it is reasonable to assume that it binds in a similar way to both the proteins. Therefore, a model of the Y−A complex may be built by using experimental structural data of the X−A complex, provided X and Y share sufficient sequence homology. It may be argued that such a scenario is limited because it requires availability of inhibitor potency data as well as ligands cocrystallized to homologous proteins. However, a review of the current literature suggests that advancements in structural genomics/chemical biology initiatives together with efficient informatics tools and the development of commercial screening assay platforms have provided modelers with a plethora of data in the public domain for confident in silico modeling.[14−17] On the basis of this observation, we investigated the possibility of developing a SBVS platform using maternal embryonic leucine zipper kinase (MELK), a member of the AMP serine/threonine kinase family, as a trial example.

MELK has recently been suggested as one of the putative therapeutic targets for several types of cancers including, breast, colorectal, ovary, and brain tumors.[18,19] In 2008, two independent reports proved that expression levels of MELK were correlated with pathologic grade of brain tumors.[20,21] In primary cultures from human glioblastoma (GBM) and medulloblastoma, MELK knockdown by siRNA results in inhibition of the proliferation and survival of these tumors. Furthermore, MELK siRNA considerably inhibits not only proliferation but also survival of stem cells isolated from GBM samples in vitro. It was also observed that MELK expression levels were inversely correlated with survival of the patients. These results suggest that inhibiting this kinase may be an attractive strategy for the development of targeted treatment for cancers such as aggressive brain tumors. X-ray crystallographic structures of human MELK have not been solved at the time of this research work. However, useful chemogenomic data is available for the target from published kinome profiling results.[22−25] With an ultimate aim to find novel inhibitors of MELK, we carried out a structure-based molecular modeling approach. The main questions we tried to address in this study are as follows: (a) In the absence of MELK crystal structures or crystal structures of high sequence identity homologues, can we use available chemogenomic data to model inhibitor induced kinase states for the purpose of drug design? (b) Can utilizing multiple conformational states of MELK improve our understanding of inhibitor binding? (c) How do the protein conformers fare in recovering known inhibitors compared to decoys in a structure-based virtual screen? (d) Can we make use of such information in a structure-based virtual screening to discover novel MELK inhibitors? To answer these questions, we modeled several conformational states of MELK by taking advantage of publicly available chemogenomic and X-ray crystallographic data. The conformational ensembles were employed to probe the role of receptor flexibility in the development of an efficient structure-based virtual screening protocol for MELK. Finally, we applied the method for the discovery of novel MELK inhibitors.

## METHODS

**I. Data Collection.** (a) *Active compounds*: Compounds tested for MELK inhibition were collected from published literature.[22−25] In total, 30 compounds of a varying potency were collected (Table S1 of the Supporting Information). Compounds with at least 25% inhibition of MELK activity when tested at 1 $\mu$M were included in the data set as MELK inhibitors. Compound coordinates were either built with the Maestro interface of the Schrödinger Suite 2010 or downloaded from the PubChem database.[15] The downloaded structures were manually corrected for proper isomeric states. Further, tautomeric/protonation states were assigned for all compounds and energy minimized with the LigPrep program. (b) *Decoy compounds*: Multiple decoy sets were used for enrichment calculations because such an approach would potentially remove the bias that might be introduced by using only a specific decoy set. Five published decoy sets were collected for the virtual screen enrichment studies. Schrödinger's "drug-like" decoy library (SCH) with average molecular weight of 400 is a library of 1000 compounds pooled from chemical collections of pharmaceutical corporate databases. This collection has been used to evaluate the performance of the Glide program.[26] Cummins and Jaeger decoy set (CMN) consist of 1000 randomly selected molecules from the MDL Drug Data Report (MDDR).[27] The Bissantz and Rognan decoy compounds (ROG) were randomly chosen from the Advanced Chemical Directory (ACD) after applying filters.[28] This data set has been later refined by the Jain group for enrichment studies to compile a decoy set of 861 compounds.[29] They also compiled a decoy set of 1000 compounds (JAI) from the ZINC database. Because the target under study is a kinase, we also sought to test the screening efficiency against a kinase-focused collection of decoys. The kinase-specific CDK2 decoy set of 1780 compounds was obtained from DUD (Directory of Useful Decoys).[30] (c) *Virtual screening compounds for novel inhibitor discovery*: The ChemBridge EXPRESS-Pick virtual library was obtained from the vendor, ChemBridge (San Diego, CA, USA), and prepared for docking with the LigPrep program.

**II. Protein Modeling, Structure Validation, and Investigation of Protein Stability through Molecular Dynamics.** For each MELK inhibitor, a similarity search based on SMILES string with a Tanimoto coefficient cut off of 70% was carried out to find crystal structural complexes of kinases in the Protein Data Bank[31] with the same or similar compound (Figure 1). For each identified kinase-inhibitor complex, BLAST sequence analysis was carried out through NCBI in order to access the sequence similarity and identity. This similarity analysis was done to evaluate the homology of template protein with MELK. In total, 16 groups of templates were selected for induced fit MELK conformational modeling (Table S3 of the Supporting Information) for each of the 16 inhibitors in the active compounds data set.

If inhibitor information is unavailable for the kinase target under investigation, relying on an induced fit template might mislead the modeler, especially when the cocrystal template

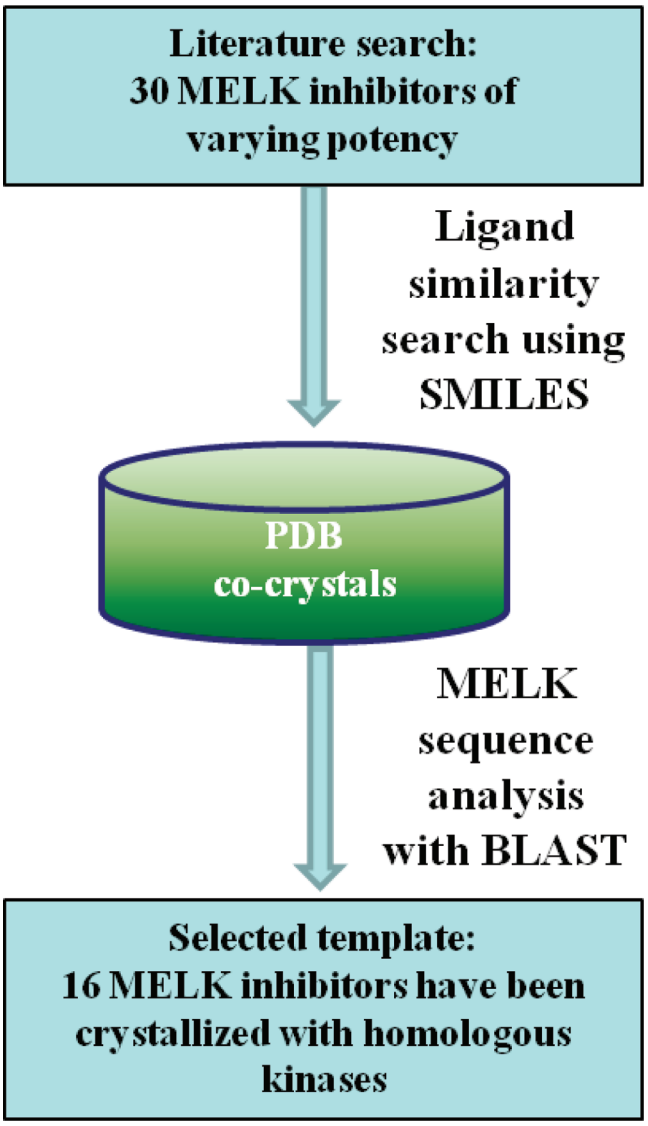



**Figure 1.** Schematic representation of template search and selection for induced fit MELK modeling.

ligand is inactive for the target protein. In such a scenario, the options available are to model the ATP bound conformer or apo kinase. Hence, we decided to investigate the utility of apo models for inhibitor discovery. However, it may be noted that utility of such a model would be limited for ligands that do not induce major induced fit shape change on the binding pocket. For modeling the unbound MELK conformers, templates were selected on the basis of conventional comparative modeling strategy in which a BLAST sequence similarity search was carried out using the human MELK kinase domain sequence as the query to the PDB. Six kinase crystal structures from the top hits had well-defined activation loops,[32] and hence, each was used to build a putative MELK model in an unbound state. As a convenient naming convention, the MELK induced fit conformers are referred by their PDB ligand ID, i.e., the staurosporine-bound conformation of MELK is referred to as STU, the VX-680-bound conformation is referred to as VX6, etc. Apo conformations are named after the PDB code of the template used in the modeling (Table S2). The comparative model building and refinement were carried out with Modeller9v8.[33] The initial models were built with the default Modeller alignment for each set of sequences. The automodel optimization and refinement protocol with default parameters were used in Modeller to generate 100 models for each group of templates. For induced fit conformational modeling, inhibitors were transferred as residues with CHARMM "undf" atom type (BLK) into the MELK model in the model building process. This would prevent amino acid residue side chains to occupy the ligand binding space thus allowing protein conformational adjustments ideal for binding of inhibitors with similar shape. Crystallographic water molecules were not considered in the modeling protocol (see the details in Supporting Information). From the 10 best models based on the Modeller DOPE scoring criterion, a single model was selected based on inspection and quality check with Ramachandran plot analysis as well as Protein Report through the Maestro v9 interface. The overall stability of each model was confirmed through a short molecular dynamics simulation. Relatively stable trajectories were obtained for all inhibitor-bound models (Supporting Information, Figure S2). We refer to the modeled inhibitor binding pose as the native state, although it has not been experimentally confirmed.

**III. Grid Preparation, Molecular Docking, and Virtual Screening Enrichment Studies.** The final protein models selected were subjected to preliminary preparation for docking employing the Schrödinger Protein Preparation Wizard protocol of Schrödinger Suite 2010. Hydrogen atoms were added followed by proper protonation state assignment. The structures were minimized with cut off criteria of 0.30 Å rmsd using the OPLS2005 force field. The prepared structures were carried forward for the docking calculations. Grid boxes capable of accommodating ligands with up to 20 Å of geometric length were created at the center of each ligand binding site with all other parameters kept at their default settings. For apo conformations, where no ligand is present, the MELK-ATP conformation was aligned to the binding pocket residues and the ATP coordinate center was used as the center of the grid box. All the docking and virtual screening calculations were performed with the Virtual Screening Workflow script included with the Schrödinger Suite 2010 (Glide v5.6). Ligands were docked flexibly, penalizing nonplanar conformations of amide bonds. No constraints or biases were introduced in the docking process. Because full force-field minimization was not performed during the docking stage, a postdocking minimization was applied. In this minimization, highly strained ligand geometries are optimized and rescored while keeping the protein fixed. The top scoring solution for each of the screened compounds was retained for the final analysis. Calculation and analysis of ligand heavy atom rmsd of the docked solution to the native pose was performed with in-house scripts as well as Schrödinger python scripts. Before rmsd calculations, all the protein conformers were aligned on the basis of the ATP binding site residues Leu86, Glu87, Tyr 88, Cys89, Pro90, Glu93, and Ile149. This represents a reasonable alignment of the most common pharmacophoric interaction of hinge region[32] backbone hydrogen bonds with the ligands (Figure S3 of the Supporting Information). For the final screening performance analysis, docking results of active compounds were merged with docking results of decoy sets and enrichment metrics were calculated.

**IV. Enrichment Metrics.** There has been scientific debate over the advantages and disadvantages of various enrichment metrics.[34,35] The classic enrichment metric, the enrichment factor (EF), has several limitations. As EF is highly dependent

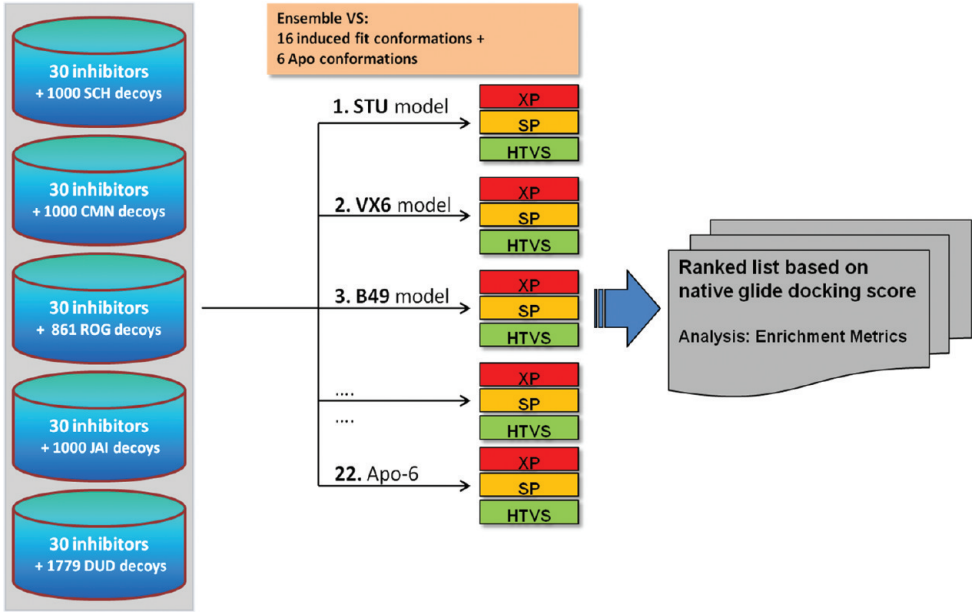



**Figure 2.** Ensemble VS scheme. Five decoy sets were screened with Glide XP, SP, and HTVS separately to calculate the enrichment.

on the ratio of active molecules to decoys used in the screening database. We selected multiple decoy sets of varying number; our analysis is not well-suited for EF calculation. Instead, we used the area under the ROC plot as enrichment metric for overall screening performance.[34] This method has been widely used as an enrichment metric and is considered a good indicator of screening performance. Although it has been reported to be not completely independent of the databases size, the area under the ROC plot suits calculations in the presented study.

## RESULTS AND DISCUSSION

**1. Accuracy of Docking/Scoring Methods in Pose Prediction.** Performance of a docking/scoring program for pose reproduction is conventionally measured by the root-mean-square deviation (rmsd) of the docked ligand pose from the experimental binding pose. Because the MELK binding modes of the ligands were unavailable experimentally, we considered the true binding mode of inhibitors to be the same as that found in the homologous template kinase structures. Although it is possible for the inhibitors to bind in an alternate binding mode to MELK, our assumption is validated by the observation of multiple templates used for each induced fit MELK model (Table S3 of the Supporting Information). For each group of induced fit templates, the inhibitor binds to different kinases in comparable binding mode, except for L20. Hence, for the purpose of our study, we refer to the inhibitor binding mode similar to that found in the homologous template kinase as the MELK native binding mode.

*1a. Self-Docking Studies.* To test the docking/scoring method in reproducing the native binding mode, we investigated whether each of the Glide docking methods, XP, SP, and HTVS, could redock inhibitors to their respective bound MELK conformer. Docking results were benchmarked by the ability of each protein conformer to dock the top scored solution of each ligand within 2 or 4 Å rmsd of the native pose. It is reasonably assumed in this calculation that the generation of ligand conformations successfully enumerates the native ligand conformation among all the poses.[36] In such a case, if the scoring function could capture key interactions, all the ligands are expected to dock back to their own binding pocket with low rmsd to the native pose. The self-docking results shows that XP was successful in reproducing the native poses within 4 Å in 95% of the cases, whereas SP and HTVS were successful in 62% and 50% of the cases, respectively (Figure 3). Although

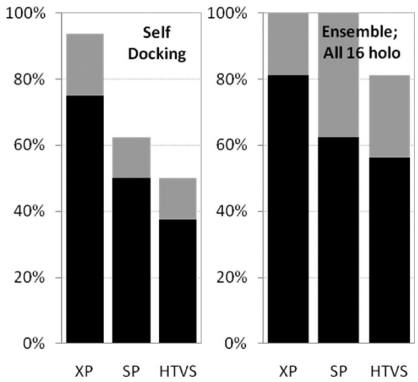



**Figure 3.** Ability of each of the MELK conformations in recovering native pose of the ligand in self-docking and ensemble docking. *Y*-axis is the percentage of lowest scoring solutions of each ligand that falls within rmsd of 2 Å (black) or 4 Å (gray) to the native MELK model ligand pose.

recovering the true binding mode may not be necessarily required for superior virtual screen enrichment, identifying the correct pharmacophoric interactions is critical for structure-guided lead optimization.

*1b. Ensemble Docking Studies.* The inability for a ligand to redock to its native pose could be due to several reasons. One issue may be that the ligand sampling does not generate the bioactive conformation, not giving the docking algorithm a chance to score the proper pose. If the sampling was able to generate the correct ligand conformation, it is possible that scoring function could fail to rank the bioactive conformation as the best solution. In yet another scenario, errors in modeling the correct protein conformation could prevent identification of

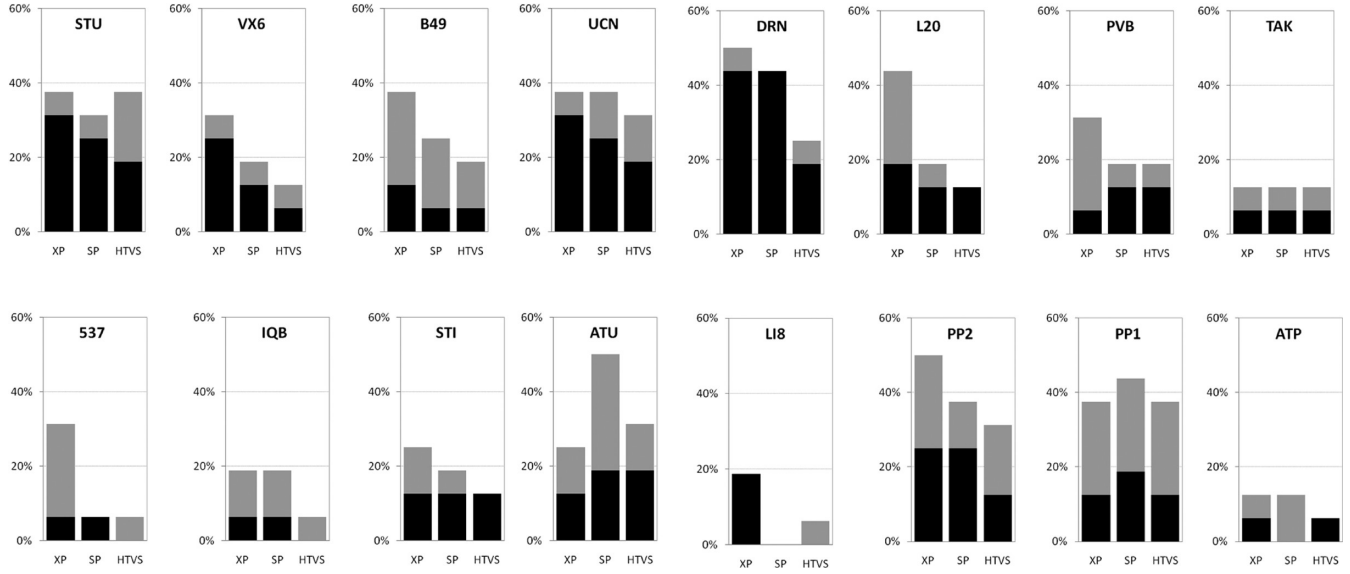



**Figure 4.** Recovery of native ligand poses by 16 MELK holo conformations. *Y*-axis is the percentage of best-scoring solutions that fall within RMSDs of 2 Å (black) or 4 Å (gray) of the native ligand poses.

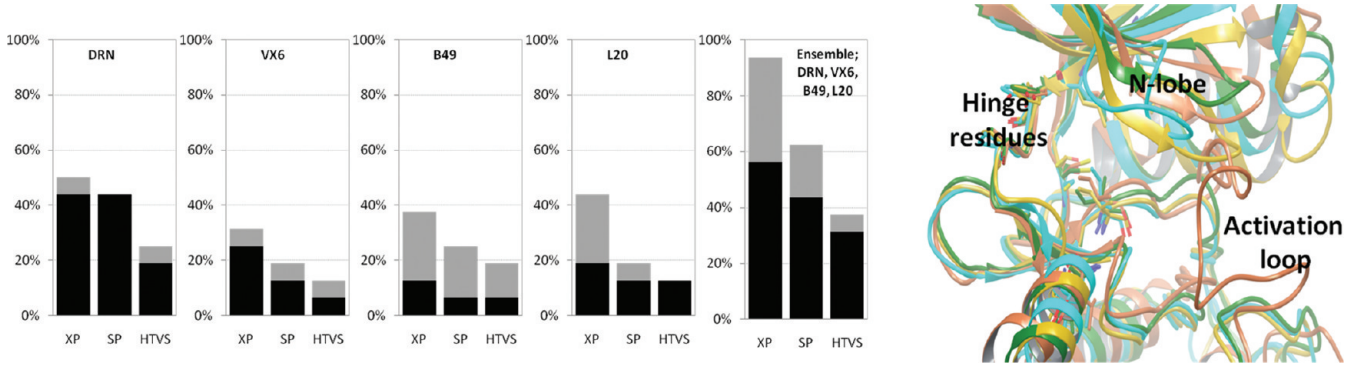



**Figure 5.** Performance of a minimal ensemble. (a) The ability of four individual protein conformations (DRN, VX6, B49, and L20) to recover the native binding mode in comparison to the ensemble of the conformational states. While the best performance of any individual conformation in the ensemble is 50% (DRN), an ensemble of the four states reproduced about 95% of the native ligand poses within 4 Å. (b) The four conformations aligned by the binding site residues shows the flexibility of the N-lobe, especially the P-loop, which are essential in determining the binding pocket shape.

the correct docking pose. To investigate the influence of binding site flexibility on molecular docking, a cross-docking experiment was performed. In this cross-docking experiment, we probed the influence of binding pocket flexibility on top-scoring docking poses of the ligands. The 16 ligands with known binding modes were extracted from their respective models, coordinates were randomized, and the ligands were docked to each of the 16 holo MELK models. Only the top-scoring docking pose for each ligand was retained for analysis. Furthermore, the pose for each ligand with the lowest rmsd to its native pose was specifically examined to determine whether any of the cross-docked solutions other than the self-docked solution could perform better at reproducing the native bioactive conformation. Indeed, the results show that ligands could dock to alternate protein conformations with lower rmsd than when docked to their cognate receptor conformation (Figure 3). This is especially noticeable when using the SP or HTVS docking algorithms. While SP self-docking reproduced about 60% of the ligands within 4 Å of their native poses, using an ensemble of protein conformations improved the success to 100% of the cases. For some of the ligands, self-docking did not produce a near native pose because of the inability of the docking-scoring method. With slight changes in the binding pocket with an alternate protein conformer, the same method was able to recover a near native state. From Figure 3, it is evident that the less sophisticated SP and HTVS methods benefitted more from ensemble docking..

From the cross-docking outcome, it is evident that binding pocket plasticity considerably affects docking performance (Figure 4). None of the individual conformations in our study were able to recover more than 50% of ligands within 4 Å of their native states; there is no "one-fits-all" protein conformation among the ensemble of conformations in this study. A few of the induced fit conformations like STU, DRN, and UCN reproduced a higher percentage of native ligand interactions where other conformations, such as ATP and TAK, performed poorly. The best recovery of native poses was by DRN, which recovered nearly 50% of the compounds within 4 Å rmsd of the native poses, indicating that multiple chemotypes were able to dock to the particular protein conformation in native-like fashion and with high scoring. In line with the previous reports, we found that the cofactor (ATP)-bound


dx.doi.org/10.1021/ci300040c | J. Chem. Inf. Model. 2012, 52, 1345−1355

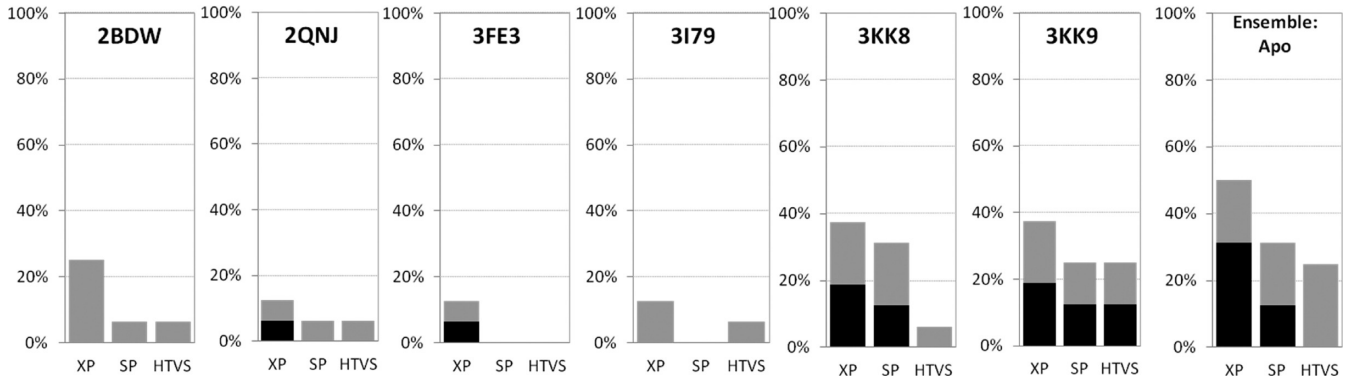



**Figure 6.** Ability of six unliganded MELK models to reproduce correct docking poses. *Y*-axis is the percentage of lowest scoring solutions that docked within rmsd of 2 Å (black) or 4 Å (gray) to the native MELK model ligand. pose.

enzyme conformer failed to produce a satisfactory outcome. This is attributed to movement of the flexible P-loop, the residues of which directly interacts with phosphate groups of ATP as well as magnesium ions. This restricts the binding site, making it difficult for other ligands to bind. To identify the structural aspects of the models that lead to poor docking performance, the models were superimposed and visually inspected (Figure S4 of the Supporting Information). In two of the underperforming MELK conformers (ATP and PVB), Ile17 of the N-lobe $\beta$-sheet flap was found to cause unfavorable steric interactions in the binding pocket, forcing inhibitors to extend toward the solvent accessible surface. Many of the inhibitors have planar aromatic rings at this position, which is unoccupied in the kinase-ATP complex. In the successful MELK conformations like STU, VX6, DRN, and UCN, Ile17 is oriented such that the planar aromatic rings of the inhibitors interact favorably with CH/$\pi$ hydrogen bonds to MELK residues which enhance binding affinity (Figure S4 of the Supporting Information).[37] These outcomes emphasize the importance of obtaining the proper binding site shape for successful molecular docking.

*1c. Probe for a Minimal Ensemble for Docking.* Because screening for multiple receptor conformations increases the chance that near-native docking poses are recovered, we decided to test a minimal, yet diverse set of conformers for their ability to account for the correct binding of the ligands in our data set. We manually selected four structurally distinct MELK conformers, DRN, VX6, B49 and L20, from the pool on the basis of backbone rmsd of the binding site residues, visual inspection of the p-loop, and individual docking performance to carry out an ensemble docking analysis. With the four-structure ensemble, the Glide XP docking algorithm was able to recover 95% of the ligands within 4 Å of their native states, accounting for all of the ligands except Imatinib, which binds to the distinct inactive type-II kinase conformation (Figure 5). This shows that the inability of any single protein conformation to reproduce a correct docking mode is primarily due to the rigid protein approximation. If an appropriate protein conformation is provided, the docking-scoring functions are sufficiently capable of sampling the ligand conformations and recognizing the correct solutions. It is also interesting to note that a few conformational states of the protein are sufficient to account for binding of most of the ligands. A practical aspect that emerges from this observation is the utility of multiple protein conformations in docking calculation in order to probe the near-native bioactive conformation of kinase inhibitors for successful lead optimization. Although involvement of an expert is still essential to identify the true pharmacophoric protein−ligand interactions when no experimental information is known a priori, this approach provides the modeler with better insights to protein−ligand molecular recognition.

*1d. Docking to Unbound Conformations.* To expand upon the observation that inhibitor-induced MELK conformers performed well in recovering native binding modes of ligands, we probed the ability of conformations modeled after unbound states to reproduce correct binding modes. Six MELK models were built on the basis of unique, unliganded kinase crystal structures. Interestingly, two of these models (3KK8, 3KK9) performed comparable to or better than many of the induced fit models (Figure 6). On the basis of structural clustering with backbone atoms, these two conformations were grouped with the kinase active state along with other holo active conformers (Table 1). This demonstrates that if an unliganded active kinase conformer is modeled, that would be sufficient to explain binding interactions of several of the inhibitors. However, this outcome is not surprising because the majority of inhibitors discovered until now target the active kinase state.[38] The compounds VX-680, Purvalanol-B, and Imatinib did not dock to any of the apo conformation within 4 Å rmsd. Imatinib binds to a distinct, so-called "DFG-out", inactive conformation, whereas the VX-680-bound MELK conformation shows a major change in the P-loop orientation with a characteristic folded $\beta$-sheet "flap".

**Table 1. Structural clustering of MELK conformers[a]**

| Model | Holo | Apo |
|---|---|---|
| Cluster 1: Activation loop oriented similar to phosphorylated state (Active state) | STU, VX6, DRN, UCN, L20, TAK, IQB, ATU, LI8, PP1, PP2, ATP | 3KK9, 3KK8 |
| Cluster 2: Activation loop randomly disoriented (Inctive state) | B49, PVB, 537, STI | 2BDW, 2QNJ, 3FE3, 3I79 |

[a]Clustering based on backbone atoms classified the conformers broadly into two categories: the active kinase state and the inactive state.

Unliganded conformers would be of interest in a drug discovery project scenario when chemogenomic data is unavailable or insufficient for the target. If ligands are able to dock to the modeled apo state, it is likely that such ligands bind through "conformational selection" of the protein rather than by inducing major conformational changes in the binding site.


dx.doi.org/10.1021/ci300040c | J. Chem. Inf. Model. 2012, 52, 1345−1355

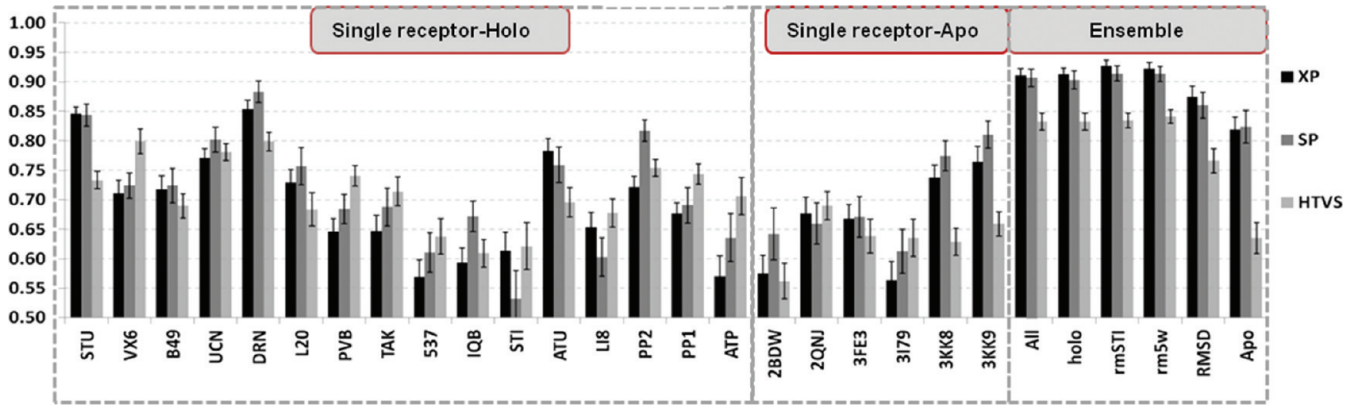



**Figure 7.** Virtual screen enrichment metric, ROC, calculated for individual MELK conformers and different combinations of conformers. The values represented are the mean of metrics calculated for five virtual screens, each utilizing different decoy sets. The number of decoys range from 861 to 1779, and 30 inhibitors were seeded with each of the decoy sets in the screening (Table S3 of the Supporting Information). All the self-docking solutions were removed for the ensemble analysis to avoid any potential bias. The maximum enrichment possible for ROC is 1.00, whereas a random enrichment will give a value around 0.5. The ensembles of kinase conformers consist of all the conformers in the study (All), 16 induced fit conformers (holo), induced fit without type-II kinase conformer (rmSTI), without worst performing five conformers (rm5w), selection of four conformers (DRN, VX6, B49, L20) that produce a minimal ensemble for acceptable pose reproduction (rmsd), and unliganded conformers (Apo).

Although protein plasticity is generally believed to be induced by the inhibitor, recent reports offer insightful views on this idea.[39,40] The protein structure in its native folded state exists as a population of dynamic conformational states that are separated by relatively high free energy barriers. According to one school of thought, a ligand may select the protein conformer to which it binds most favorably. After the ligand binding occurs via conformational selection, the complex could undergo further induced fit rearrangements. Hence, in computational docking, consideration of this local adjustment may be required to achieve better pose prediction. One plausible approach would be to apply the available induced fit docking methods to account for the side-chain flexibility once the appropriate unliganded conformers are modeled.[41,42]

*1e. Significance of Modeling Multiple Conformations in Structure-Based Computational Molecular Design.* One of the relatively major conformational changes in the MELK binding site is observed with VX6. In this model, Phe22 in the P-loop undergoes a major shift from that observed in other models to adopt a folded state, thus allowing favorable $\pi-\pi$ interaction with VX-680. According to one recent statistical and computational analysis of X-ray crystallographic structures, inhibitors that bind to or induce such a folded kinase state are most likely to exhibit better kinome selectivity profiles, particularly if they favorably interact with the conserved Phe or Tyr residues of the P-loop.[43] Hence, such distinctive target states might be helpful in SBVS for the discovery of novel scaffolds with better selectivity profiles compared to receptor conformations modeled after structures bound to nonselective ligands, such as staurosporine.

**2. Virtual Screen Enrichment Performance.** From the cross-docking data, it is evident that utilizing a few diverse protein conformations improves the ability to identify binding poses more similar to the bioactive conformation when compared with single rigid receptor docking. Further, we sought to study whether this would also correlate with enhanced virtual screen enrichment in recovering true ligands from a virtual library when seeded with decoy compound sets. Such an improved enrichment would be vital for practical virtual screening purposes where it is more important to score and rank an active ligand at the top of the list than to find the correct binding mode. In order to obtain an unbiased representation of the enrichment performance, we used five separate, published decoys sets to calculate the performance (Table S3 of the Supporting Information). The final results were calculated as the average of the performance of these five screening results.

Comparison of the ensemble VS data to the single receptor VS shows that when protein flexibility is accounted for in the docking calculation enrichment is improved (Figure 7). If we know the single most enriched receptor conformation a priori, it would be ideal in terms of computational cost. However, in a scenario where we do not have such information, using an ensemble of protein conformations could give a better performance. One potential caveat of utilizing multiple conformations has been previously reported. The rate of false positive can increase if inactive compounds dock with higher docking scores than actives compounds in any of the single protein conformations. Although we did not notice this to be critical in our study, we observed a high false positive rate in the "DFG-out" model, STI. Therefore, together with incorporation of receptor flexibility through multiple reliable protein models, a capable scoring function is essential for successful ensemble SBVS.

**3. Comparison of Performance of Scoring Functions.** We analyzed the comparative ability of the Glide XP, SP, and HTVS docking-scoring method to reproduce the native pose as the best-scoring docking solution when the same protein conformation was used in the docking experiment. The most rigorous computational docking method XP performed slightly better than the faster SP in recovering the native poses within 2 Å rmsd.(Figure 4 and Figure 6) In all cases except the PVB, ATU, and PP1 conformations, XP performed better than SP, although not by a remarkable margin. In general, we observe that if a particular protein conformation is used in the docking, XP fared better than SP. One notable exception in the performance was with the LI8 conformer with which SP could not recover any ligands within even 4 Å rmsd while XP recovered about 20% of the ligands within 2 Å rmsd. This result shows the difference in the docking-scoring method when using the same protein conformation. As expected, the fastest and most crude docking method, Glide HTVS, was clearly inferior


dx.doi.org/10.1021/ci300040c | J. Chem. Inf. Model. 2012, 52, 1345−1355

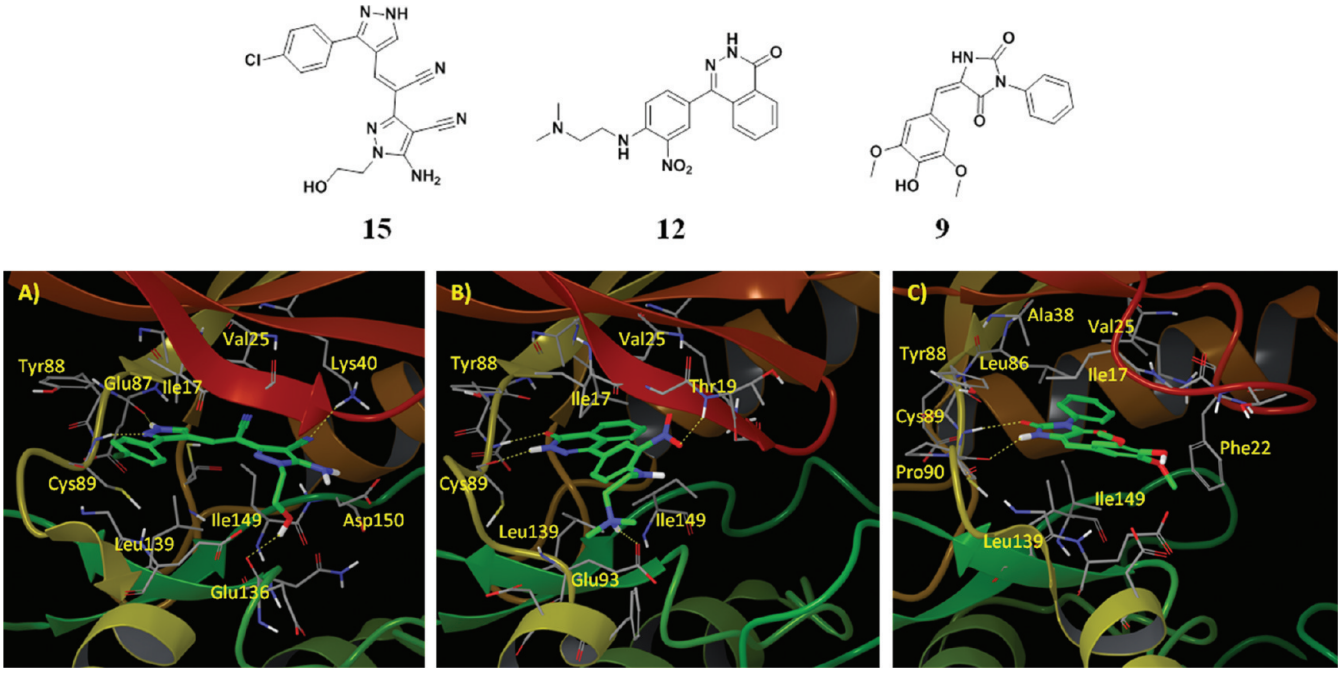



**Figure 8.** Three novel MELK inhibitors discovered through ensemble virtual screening, **15**, **12**, and **9**, and their predicted binding poses A, B, and C, respectively. Important residues in the binding pocket are labeled.

to the docking performance of SP and XP. The overall best enrichment was given by Glide SP for the DRN conformation. Considering that the XP docking algorithm requires several fold more computational time than the SP method, the latter method appears to be most suitable for large scale screening, especially when utilizing multiple protein conformers.

**4. Discovery of Novel MELK Inhibitors by Ensemble Virtual Screening Protocol.** Because the ensemble screening study with known inhibitors yielded results with high enrichment factors, we sought to test the method for discovery of novel inhibitors of MELK. We chose to screen a readily available commercial compound library against multiple MELK conformers. The overall screening process was carried out in sequential stages to balance speed and accuracy. The preliminary ensemble VS was carried out with three diverse MELK conformers with good enrichment factors: DRN, VX6, and B49. In the preliminary stage, about half a million compounds were rapidly screened to the ATP binding pocket of all three selected conformers using Glide HTVS docking. The top 10% of the compounds from HTVS were carried forward to be further screened by the more exhaustive Glide SP docking algorithm. High scored 10% of the SP docked compounds were retained as a focused library of potential MELK inhibitors. In order to further narrow down the compounds for purchase and enzyme inhibition assay, we performed Glide XP ensemble docking with all the holo MELK conformers except the ATP and STI conformers. After visual inspection of the docked poses, we purchased 23 compounds for experimental testing of MELK inhibition. We used commercial kinome profiling service by KINOMEscan because the assay evaluates the direct competitive binding of compounds to the ATP pocket.[22] The preliminary screening was performed by a single concentration duplicate assay. In the preliminary screening, five compounds inhibited MELK activity by at least 25% at a single concentration of 10 $\mu$M ( Table S4 of the Supporting Information). Finally binding affinity was determined for selected three inhibitors (Figure 8).

All of the three inhibitors show a pair of hydrogen bonds with the hinge residues, typical for many of the ATP-competitive kinase inhibitors. The most potent inhibitor, **15** ($K_d$ = 0.37 $\mu$M), appears to bind to conformers similar to the staurosporin-induced conformation. The pyrazole ring of **15** shows hydrogen bonds with the backbone of hinge residues Glu87 and Cys89. Both the nitrile groups contact residues in the interior of the pocket. One of the nitrile groups makes a hydrogen bond with the conserved catalytic residue Lys40. Another hydrogen bond is formed by the hydroxyl group to backbone carbonyl oxygen of Glu136. Asp150 is in the vicinity of the amino group of **15** to make Coulombic interactions. Favorable hydrophobic interactions are formed with Leu139 and Ile149, which forms the floor of the binding pocket, as well as Ile17 and Val25, which are part of the roof N-terminal lobe. Because **15** does not appears to interact with the solvent accessible specificity region residues, it is amenable to optimization in order to gain potency and selectivity. On the contrary, the protonated dimethylamino group of inhibitor **12** ($K_d$ = 3.5 $\mu$M) extends to this region showing electrostatic interactions with Glu93, an interaction similar to that formed by sunitinib. Hydrogen bonds are formed between the phthalazinone ring of **12** and the hinge residues as well as between the nitro group and backbone of Thr19. The most selective among all the inhibitors was **9** ($K_d$ = 18 $\mu$M), which showed no appreciable activity to the other kinases we tested in the preliminary screen (Table S4 of the Supporting Information). One of the methoxy groups shows favorable interaction with Phe22 of the P-loop folded conformation (VX6), a residue position that has recently been highlighted as a possible determinant of selectivity for many of the known kinase inhibitors.[43] A pair of hinge residue hydrogen bonds is formed by the imidazolidinedione moiety, while the phenyl group extends toward the interior of the binding pocket,


dx.doi.org/10.1021/ci300040c | J. Chem. Inf. Model. 2012, 52, 1345−1355

making favorable hydrophobic interactions. These compounds could be potentially optimized for further investigation for various disorders caused by MELK deregulation such as GBM. Because the three inhibitors were selected from three different MELK conformers, the same VS protocol could only yield one inhibitor if a single template were used. In order to check the novelty of these three computationally discovered MELK inhibitors, we performed a similarity search to the Chemical Abstracts Service (CAS) database through the SciFinder Web interface. The search results show that these compounds or similar compounds (with at least 80% similarity) have not been previously proven as kinase inhibitors, confirming that such computational methods can discovery novel chemical scaffolds for the desired bioactivity.

## CONCLUSIONS

Studying the structural basis of molecular recognition between an enzyme and its inhibitor is vital for rational lead optimization for improved potency, selectivity, and ADMET profile. One of the most reliable strategies to understand this pharmacophoric interaction is by cocrystallization of inhibitors with target proteins. In the absence of any such experimental structural data, comparative modeling can be helpful, provided that homologous proteins have been crystallized. In this study, we modeled several inhibitor-bound conformational states of MELK and explored their use in structure-based drug design. We first investigated the influence of different protein conformations and scoring functions on successful pose prediction. A few of the conformers demonstrated a better ability to dock known inhibitors. Further, we carried out virtual screening enrichment studies with three docking-scoring methods and five published decoy sets to validate the utility of the MELK homology models and to determine an optimal protocol in SBVS. All the models displayed better enrichment than random screening, with enrichment performance varying with protein conformation. Our results suggest that utilizing an ensemble of conformations can produce a more robust enrichment compared to single receptor screening. Interestingly, a few of the apo models showed comparable enrichment and pose prediction to many of the holo conformers. The conformers with lower enrichment may not always represent an unsuitable model for screening. Rather, these may represent a more unique conformation for the protein which could be exploited to discover novel selective inhibitor scaffolds. Our results show that by selecting appropriate crystal structure templates, a kinase domain can be modeled for the purpose of inhibitor design even when it shares as sequence identity as low as 35% with the template. Thus, the kinome space that is accessible to structure-based drug design expands well beyond that of kinases for which experimental structures have been determined. This approach and the results obtained in our experiments could assist in the development of new classes of MELK inhibitors as well as for the optimization of the known kinase scaffolds for selectivity and potency.

As mentioned earlier, we modeled apo conformers to evaluate their utility if no inhibitor data is available for enrichment studies for the kinase target. We observe that the performance of phosphorylated apo kinase conformers were comparable to several induced fit conformers. This outcome could be because the active compound data set has been derived from experiments using the phosphorylated kinase enzyme. Our results also suggest that the pan-kinase inhibitor staurosporine induced conformer can be quite useful. This is likely due to the fact that its flat and spacious six membered ring occupies major portion of the active site, shaping up optimal site for inhibitors that make hydrogen bonding to the hinge region residues. It would be beneficial to preselect the best conformers for an improved performance of SBVS. However, from our analysis and previous studies that investigated a wide spectrum of target classes suggest that it is not possible to know which conformer would be highly enriched a priori.[44] Recent progress in this field suggests some novel approaches in identifying ideal set of receptor conformers for SBVS.[45−47]

## ASSOCIATED CONTENT

### Supporting Information

Information regarding MELK inhibitors, PDB templates used in the MELK modeling, details of the modeling process, rmsd plots of the MELK-inhibitor complex molecular dynamics simulations and information regarding the decoy data sets used for enrichment studies are reported in the Supporting Information. This material is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION

### Corresponding Author

*E-mail: li.728@osu.edu. Phone: 614-247-8786.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors thank Ryan Pavlovicz for his help in discussions as well as manuscript preparation. Schrödinger LLC support, especially Dr. Jeffrey A. Saunders and Dr. Stuart Murdock, are acknowledged for modifying python scripts used in this work. Decoy data sets were downloaded from Schrödinger, Dr. Ajay Jain lab, and DUD Web sites. Dr. Ichiro Nakano, Dr. May Mok, and Dr. Chunyu Gu are acknowledged for their help in various phases of this project. Part of the work has been done using computing resources like Glenn cluster provided by the Ohio Supercomputer Center.

## REFERENCES

(1) Ripphausen, P.; Nisius, B.; Peltason, L.; Bajorath, J. Quo vadis, virtual screening? A comprehensive survey of prospective applications. *J. Med. Chem.* **2010**, *53*, 8461−7.

(2) Yuriev, E.; Agostino, M.; Ramsland, P. A. Challenges and advances in computational docking: 2009 in review. *J. Mol. Recognit.* **2011**, *24*, 149−64.

(3) Teague, S. J. Implications of protein flexibility for drug discovery. *Nat. Rev. Drug Discov.* **2003**, *2*, 527−41.

(4) Erickson, J. A.; Jalaie, M.; Robertson, D. H.; Lewis, R. A.; Vieth, M. Lessons in molecular recognition: The effects of ligand and protein flexibility on molecular docking accuracy. *J. Med. Chem.* **2004**, *47*, 45−55.

(5) C, B. R.; Subramanian, J.; Sharma, S. D. Managing protein flexibility in docking and its applications. *Drug Discovery Today* **2009**, *14*, 394−400.

(6) Dror, R. O.; Pan, A. C.; Arlow, D. H.; Borhani, D. W.; Maragakis, P.; Shan, Y.; Xu, H.; Shaw, D. E. Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 13118−23.

(7) Bolstad, E. S.; Anderson, A. C. In pursuit of virtual lead optimization: The role of the receptor structure and ensembles in accurate docking. *Proteins* **2008**, *73*, 566−80.

(8) Lin, J. H.; Perryman, A. L.; Schames, J. R.; McCammon, J. A. The relaxed complex method: Accommodating receptor flexibility for drug design with an improved scoring scheme. *Biopolymers* **2003**, *68*, 47−62.
(9) Lin, J. H.; Perryman, A. L.; Schames, J. R.; McCammon, J. A. Computational drug design accommodating receptor flexibility: the relaxed complex scheme. *J. Am. Chem. Soc.* **2002**, *124*, 5632−3.
(10) McCammon, J. A. Target flexibility in molecular recognition. *Biochim. Biophys. Acta* **2005**, *1754*, 221−4.
(11) Dong, Z.; Li, Y.; Kim, D. J.; Ma, W.; Lubet, R. A.; Bode, A. M., Discovery of novel checkpoint kinase 1 inhibitors by virtual screening based on multiple crystal structures. *J. Chem. Inf. Model.* **2011**.
(12) Craig, I. R.; Essex, J. W.; Spiegel, K. Ensemble docking into multiple crystallographically derived protein structures: an evaluation based on the statistical analysis of enrichments. *J. Chem. Inf. Model.* **2010**, *50*, 511−24.
(13) Rao, S.; Sanschagrin, P. C.; Greenwood, J. R.; Repasky, M. P.; Sherman, W.; Farid, R. Improving database enrichment through ensemble docking. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 621−7.
(14) Wendt, B.; Uhrig, U.; Bos, F. Capturing structure-activity relationships from chemogenomic spaces. *J. Chem. Inf. Model.* **2011**, *51*, 843−51.
(15) Li, Q.; Cheng, T.; Wang, Y.; Bryant, S. H. PubChem as a public resource for drug discovery. *Drug Discovery Today* **2010**, *15*, 1052−7.
(16) Wang, Y.; Bolton, E.; Dracheva, S.; Karapetyan, K.; Shoemaker, B. A.; Suzek, T. O.; Wang, J.; Xiao, J.; Zhang, J.; Bryant, S. H. An overview of the PubChem bioAssay resource. *Nucleic Acids Res.* **2010**, *38*, D255−66.
(17) Xie, X. Q. Exploiting PubChem for virtual screening. *Expert. Opin. Drug Discov.* **2010**, *5*, 1205−1220.
(18) Gray, D.; Jubb, A. M.; Hogue, D.; Dowd, P.; Kljavin, N.; Yi, S.; Bai, W.; Frantz, G.; Zhang, Z.; Koeppen, H.; de Sauvage, F. J.; Davis, D. P. Maternal embryonic leucine zipper kinase/murine protein serine-threonine kinase 38 is a promising therapeutic target for multiple cancers. *Cancer Res.* **2005**, *65*, 9751−61.
(19) Lin, M. L.; Park, J. H.; Nishidate, T.; Nakamura, Y.; Katagiri, T. Involvement of maternal embryonic leucine zipper kinase (MELK) in mammary carcinogenesis through interaction with Bcl-G, a pro-apoptotic member of the Bcl-2 family. *Breast Cancer Res.* **2007**, *9*, R17.
(20) Marie, S. K.; Okamoto, O. K.; Uno, M.; Hasegawa, A. P.; Oba-Shinjo, S. M.; Cohen, T.; Camargo, A. A.; Kosoy, A.; Carlotti, C. G., Jr.; Toledo, S.; Moreira-Filho, C. A.; Zago, M. A.; Simpson, A. J.; Caballero, O. L. Maternal embryonic leucine zipper kinase transcript abundance correlates with malignancy grade in human astrocytomas. *Int. J. Cancer* **2008**, *122*, 807−15.
(21) Nakano, I.; Masterman-Smith, M.; Saigusa, K.; Paucar, A. A.; Horvath, S.; Shoemaker, L.; Watanabe, M.; Negro, A.; Bajpai, R.; Howes, A.; Lelievre, V.; Waschek, J. A.; Lazareff, J. A.; Freije, W. A.; Liau, L. M.; Gilbertson, R. J.; Cloughesy, T. F.; Geschwind, D. H.; Nelson, S. F.; Mischel, P. S.; Terskikh, A. V.; Kornblum, H. I. Maternal embryonic leucine zipper kinase is a key regulator of the proliferation of malignant brain tumors, including brain tumor stem cells. *J. Neurosci. Res.* **2008**, *86*, 48−60.
(22) Karaman, M. W.; Herrgard, S.; Treiber, D. K.; Gallant, P.; Atteridge, C. E.; Campbell, B. T.; Chan, K. W.; Ciceri, P.; Davis, M. I.; Edeen, P. T.; Faraoni, R.; Floyd, M.; Hunt, J. P.; Lockhart, D. J.; Milanov, Z. V.; Morrison, M. J.; Pallares, G.; Patel, H. K.; Pritchard, S.; Wodicka, L. M.; Zarrinkar, P. P. A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **2008**, *26*, 127−32.
(23) Bain, J.; Plater, L.; Elliott, M.; Shpiro, N.; Hastie, C. J.; McLauchlan, H.; Klevernic, I.; Arthur, J. S.; Alessi, D. R.; Cohen, P. The selectivity of protein kinase inhibitors: A further update. *Biochem. J.* **2007**, *408*, 297−315.
(24) Hoang, T. M.; Favier, B.; Valette, A.; Barette, C.; Nguyen, C. H.; Lafanechere, L.; Grierson, D. S.; Dimitrov, S.; Molla, A. Benzo[e]-pyridoindoles, novel inhibitors of the aurora kinases. *Cell Cycle* **2009**, *8*, 765−72.
(25) Akue-Gedu, R.; Rossignol, E.; Azzaro, S.; Knapp, S.; Filippakopoulos, P.; Bullock, A. N.; Bain, J.; Cohen, P.; Prudhomme, M.; Anizon, F.; Moreau, P. Synthesis, kinase inhibitory potencies, and in vitro antiproliferative evaluation of new Pim kinase inhibitors. *J. Med. Chem.* **2009**, *52*, 6369−81.
(26) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739−49.
(27) Cummings, M. D.; DesJarlais, R. L.; Gibbs, A. C.; Mohan, V.; Jaeger, E. P. Comparison of automated docking programs as virtual screening tools. *J. Med. Chem.* **2005**, *48*, 962−76.
(28) Bissantz, C.; Folkers, G.; Rognan, D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *J. Med. Chem.* **2000**, *43*, 4759−67.
(29) Pham, T. A.; Jain, A. N. Parameter estimation for scoring protein-ligand interactions using negative training data. *J. Med. Chem.* **2006**, *49*, 5856−68.
(30) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking sets for molecular docking. *J. Med. Chem.* **2006**, *49*, 6789−801.
(31) Kirchmair, J.; Markt, P.; Distinto, S.; Schuster, D.; Spitzer, G. M.; Liedl, K. R.; Langer, T.; Wolber, G. The Protein Data Bank (PDB), its related services and software tools as key components for in silico guided drug discovery. *J. Med. Chem.* **2008**, *51*, 7021−40.
(32) Zhang, J.; Yang, P. L.; Gray, N. S. Targeting cancer with small molecule kinase inhibitors. *Nat. Rev. Cancer* **2009**, *9*, 28−39.
(33) Eswar, N.; Webb, B.; Marti-Renom, M. A.; Madhusudhan, M. S.; Eramian, D.; Shen, M. Y.; Pieper, U.; Sali, A., Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* **2007**, *Chapter 2*, Unit 2 9.
(34) Truchon, J. F.; Bayly, C. I. Evaluating virtual screening methods: good and bad metrics for the "early recognition" problem. *J. Chem. Inf. Model.* **2007**, *47*, 488−508.
(35) Kirchmair, J.; Markt, P.; Distinto, S.; Wolber, G.; Langer, T. Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection--what can we learn from earlier mistakes? *J. Comput.-Aided Mol. Des.* **2008**, *22*, 213−28.
(36) Watts, K. S.; Dalal, P.; Murphy, R. B.; Sherman, W.; Friesner, R. A.; Shelley, J. C. ConfGen: a conformational search method for efficient generation of bioactive conformers. *J. Chem. Inf. Model.* **2010**, *50*, 534−46.
(37) Ozawa, T.; Tsuji, E.; Ozawa, M.; Handa, C.; Mukaiyama, H.; Nishimura, T.; Kobayashi, S.; Okazaki, K. The importance of CH/pi hydrogen bonds in rational drug design: An ab initio fragment molecular orbital study to leukocyte-specific protein tyrosine (LCK) kinase. *Bioorg. Med. Chem.* **2008**, *16*, 10311−8.
(38) Liu, Y.; Gray, N. S. Rational design of inhibitors that bind to inactive kinase conformations. *Nat. Chem. Biol.* **2006**, *2*, 358−64.
(39) Bakan, A.; Bahar, I. The intrinsic dynamics of enzymes plays a dominant role in determining the structural changes induced upon inhibitor binding. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 14349−54.
(40) Bahar, I.; Chennubhotla, C.; Tobi, D. Intrinsic dynamics of enzymes in the unbound state and relation to allosteric regulation. *Curr. Opin. Struct. Biol.* **2007**, *17*, 633−40.
(41) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **2006**, *49*, 534−53.
(42) Koska, J.; Spassov, V. Z.; Maynard, A. J.; Yan, L.; Austin, N.; Flook, P. K.; Venkatachalam, C. M. Fully automated molecular mechanics based induced fit protein-ligand docking method. *J. Chem. Inf. Model.* **2008**, *48*, 1965−73.
(43) Guimaraes, C. R.; Rai, B. K.; Munchhof, M. J.; Liu, S.; Wang, J.; Bhattacharya, S. K.; Buckbinder, L. Understanding the impact of the P-loop conformation on kinase selectivity. *J. Chem. Inf. Model.* **2011**, *51*, 1199−204.
(44) Fan, H.; Irwin, J. J.; Webb, B. M.; Klebe, G.; Shoichet, B. K.; Sali, A. Molecular docking screens using comparative models of proteins. *J. Chem. Inf. Model.* **2009**, *49*, 2512−27.

(45) Xu, M.; Lill, M. A. Utilizing experimental data for reducing ensemble size in flexible-protein docking. *J. Chem. Inf. Model.* **2012**, *52*, 187−98.
(46) Soderhjelm, P.; Tribello, G. A.; Parrinello, M. Locating binding poses in protein-ligand systems using reconnaissance metadynamics. *Proc. Natl. Acad. Sci.* **2012**, *109*, 5170−5175.
(47) Park, I.-H.; Li, C. Dynamic Ligand-Induced-Fit Simulation via Enhanced Conformational Samplings and Ensemble Dockings: A Survivin Example. *J. Phys. Chem. B* **2010**, *114*, 5144−5153.