

JCTC

Journal of Chemical Theory and Computation

An Evaluation of Explicit Receptor Flexibility in Molecular Docking Using Molecular Dynamics and Torsion Angle Molecular Dynamics

Roger S. Armen,[†] Jianhan Chen,[‡] and Charles L. Brooks III^{*,†}

Department of Chemistry, 930 N. University Ave, University of Michigan, Ann Arbor, Michigan 48109 and Department of Biochemistry, Kansas State University, Manhattan, Kansas 66506

Received May 21, 2009

Abstract: Incorporating receptor flexibility into molecular docking should improve results for flexible proteins. However, the incorporation of explicit all-atom flexibility with molecular dynamics for the entire protein chain may also introduce significant error and “noise” that could decrease docking accuracy and deteriorate the ability of a scoring function to rank native-like poses. We address this apparent paradox by comparing the successes of several flexible receptor models in cross-docking and multiple receptor ensemble docking for p38 α mitogen-activated protein kinase. Explicit all-atom receptor flexibility has been incorporated into a CHARMM-based molecular docking method using both molecular dynamics (MD) and torsion angle molecular dynamics (TAMD) for the refinement of predicted protein–ligand binding geometries. These flexible receptor models have been evaluated, and the accuracy and efficiency of TAMD sampling is directly compared to MD sampling. Several flexible receptor models are compared, encompassing flexible side chains, flexible loops, multiple flexible backbone segments, and treatment of the entire chain as flexible. We find that, although including side chain and some backbone flexibility is required for improved docking accuracy as expected, docking accuracy also diminishes as additional and unnecessary receptor flexibility is included into the conformational search space. Ensemble docking results demonstrate that including protein flexibility leads to improved agreement with binding data for 227 active compounds. This comparison also demonstrates that a flexible receptor model enriches high-affinity compound identification without significantly increasing the number of false positives from low-affinity compounds.

1. Introduction

It is widely recognized that a significant problem in the field of molecular docking is the incorporation of protein flexibility.^{1,2} In experimentally characterized flexible protein binding sites, conformational changes are observed upon binding different ligands that include minor changes of side chain positions, significant local side chain rearrangement, flexible loop rearrangements, larger backbone movements

of the binding pocket, and large hinge motion of domains.³ Incorporating protein flexibility in molecular docking is challenging because it is difficult to predict the extent of protein conformational changes upon ligand binding (challenges in sampling), and it is also more difficult to score flexible protein–ligand complexes accurately (challenges in scoring). It is clear from the literature that rigid receptor docking can be quite successful for some receptors.⁴ However, when rigid receptor docking is used for receptors that are known to be flexible, docking results from a single receptor conformation often suffer from too many false negatives⁵ that may be accommodated by minor conformational changes⁶ or slightly more open conformations of the

* Corresponding author phone: 734-647-6682; fax: 734-647-1604; e-mail: brookscl@umich.edu.

[†] University of Michigan.

[‡] Kansas State University.

binding pocket.⁷ Although many studies have shown that incorporating models of protein flexibility can improve docking for these cases, we demonstrate in this article that explicit protein flexibility with all atom molecular dynamics (MD) for the entire protein chain may also introduce significant error and “noise”, thereby decreasing docking accuracy and diminishing the ability of a scoring function to discriminate native-like binding poses. In contrast to this result, we also show that similar models that use MD or torsion angle molecular dynamics (TAMD) sampling for flexible side chains and flexible loops are significantly more successful than rigid docking.

In this study, we assess the incorporation of explicit all-atom protein flexibility using both MD and TAMD in CHARMM. Both methods allow flexibility of the protein and ligand simultaneously, which is essential for refinement in protein–ligand conformational space. However, TAMD (internal coordinate space) is more efficient in sampling compared to MD (Cartesian space) because there are significantly fewer degrees of freedom in internal coordinates. The use of conformational dynamics in internal coordinates for flexible receptor docking was pioneered by Abagyan and co-workers.^{8,9} Currently, explicitly exploring a large amount of conformational space for both the protein and ligand during the docking procedure is too computationally expensive for screening a large number ($>10^5$) of compounds. As many docking algorithms have been optimized for fast and accurate docking of a flexible ligand into a static protein structure, a computationally attractive approach to incorporate protein flexibility is to use multiple receptor conformations (MRCs) for rigid docking, which is now usually referred to as “receptor ensemble docking” (RED).¹⁰ In this study, we assess several different flexible receptor models using cross-docking accuracy and also compare RED results to experimental binding data.

P38 α mitogen-activated protein (MAP) kinase is a difficult test case for cross-docking in that it has a very flexible ligand binding site experimentally characterized by numerous crystal structures. Hydrogen exchange experiments monitored by mass spectrometry have shown that the unphosphorylated inactive state of p38 α is significantly more flexible than the doubly phosphorylated active state.¹¹ The majority of discovered inhibitors successfully bind to this more flexible inactive form of the active site, and some inhibitors bind to an alternative conformation of the binding site known as the “DFG-out” conformation.¹² This alternative conformation of the binding site is denoted the DFG-out conformation because the F169 side chain on the flexible activation loop flips out into the binding site revealing a deep hydrophobic pocket to which ligands are able to bind (Figure 1A). Inhibitors have now been observed binding in the DFG-out conformation in numerous tyrosine kinases and ser/thr kinases.¹³ It has been proposed that targeting the enzyme in the flexible inactive (nonphosphorylated) conformation is a good strategy for designing selective kinase inhibitors.¹⁴ Several of these inhibitors have been shown to be selective rather than promiscuous, and this has been attributed to binding interactions in the DFG-out conformation. In this paper, we examine an ensemble of 12 inactive (nonphos-

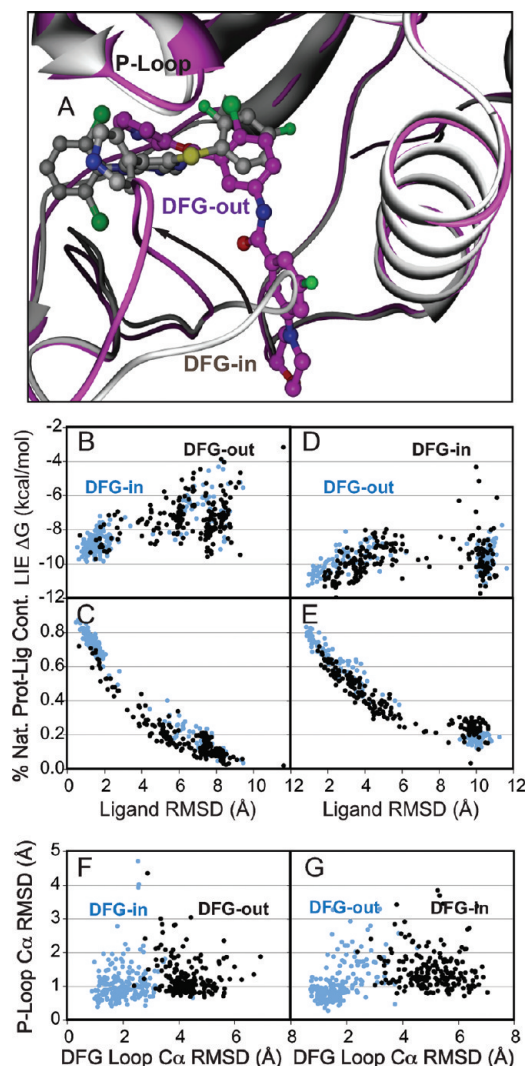


Figure 1. A successful cross-docking for a difficult example involving DFG loop rearrangement. (A) The superposition of a DFG-in protein–ligand complex (1ouy shown in gray) and a DFG-out protein–ligand complex (1w83 shown in magenta). (B) Ligand rmsd versus LIE₁ score plot of flexible receptor self-docking 1ouy (light blue), compared to flexible receptor cross-docking (black) of the ligand 1ouy into the receptor 1w83. (C) Ligand rmsd versus % native protein–ligand contacts for the same. (D) Ligand rmsd versus LIE₁ score plot of flexible receptor self-docking 1w83 (light blue), compared to flexible receptor cross-docking (black) of the ligand 1w83 into the receptor 1ouy. (E) Ligand rmsd versus % native protein–ligand contacts for the same. (F) C α rmsd of the P-loop versus C α rmsd of the DFG loop plot for flexible receptor self-docking 1ouy (light blue), compared to flexible receptor cross-docking (black) of the ligand 1ouy into the receptor 1w83. (G) C α rmsd of the P-loop versus C α rmsd of the DFG loop plot for flexible receptor self-docking 1w83 (light blue), compared to flexible receptor cross-docking (black) of the ligand 1w83 into the receptor 1ouy.

phorylated) p38 α receptor conformations: nine conformationally diverse crystal structures of DFG-in conformations (1a9u, 1bl6, 1bl7, 1di9, 1ouk, 1ouy, 1oz1, 1w84, and 1yqj)^{15–20} and three crystal structures of DFG-out conformations (1kv1, 1kv2, and 1w83).^{18,21}

2. Results and Discussion

2.1. Rigid Receptor Cross-Docking. With cross-docking simulations, it is possible to determine the sensitivity of docking results to changes in protein conformation by analyzing the effect on the docking accuracy and the prediction of $\Delta G_{\text{binding}}$. The entire 12×12 cross-dock was performed from two different starting conformations of the ligand. The first was started from the “swapped conformation” of the ligand from its cocrystal structure into a new crystal structure. The second was started from the “best rotation,” which is the lowest energy conformation from 1000 rigid body rotations into the new receptor. Cross-docking results starting from the swapped conformation are biased toward native-like poses and should represent the theoretical maximum cross-docking accuracy for the rigid receptor that atom clashes and conformational changes will allow. Rigid receptor docking results starting from the best rotation are unbiased and more directly analogous to the flexible receptor docking method, which also starts from a best rotation into a flexible receptor conformation. For cross-docking with a rigid receptor, the top five lowest energy docking poses were identified using the R-dependent dielectric (RDIE) scoring function. The reason why it is important to consider the top five scoring poses, and not only the best single scoring pose, is that scoring functions are not always able to correctly rank the most native-like poses, even in the case of self-docking. In comparing cross-docking starting from the swapped conformation and from the best rotation, the results are similar when only the top-scoring pose is considered. However, when the top five scoring poses are examined, it is possible to see that cross-docking from the swapped conformation had a higher percentage ($38\% \leq 2.0 \text{ \AA rmsd}$) of docking success compared to starting from the best rotation into the new receptor conformation ($25\% \leq 2.0 \text{ \AA rmsd}$). This is expected, as the conformational search starting from the swapped conformation should be more biased toward low rmsd native-like poses. The overall docking accuracy from rigid receptor cross-docking is similar for the lowest energy pose regardless of whether they are scored using either the RDIE or the LIE_1 (LIE = linear interaction energy) scoring function, but the LIE_1 scores perform slightly better if you consider all of the top five scoring poses. Three receptors (1kv1, 1kv2, and 1w83) are in the DFG-out conformation,^{18,21} and when the results from only the single top scoring poses are considered, these receptors are only able to accurately cross-dock their own ligands to within $\leq 2.0 \text{ \AA rmsd}$. Two smaller ligands, 1di9 (a quinazoline-based fragment) and 1w84 (an indole-based fragment), which bind in a similar conformation and are selective to the DFG-in conformation, are also able to bind to the DFG-out receptor conformations within $\leq 3.0 \text{ \AA rmsd}$ because the conformation of Phe 169 does not cause a direct atom clash. The ligands selective for the DFG-out conformation are not able to dock accurately in any of the nine DFG-in receptor conformations.

2.2. Constructing Several Models to Represent Protein Flexibility. It was our goal to directly compare cross-docking results using several models of protein flexibility scaling from small side chain movements to large backbone movements.

We first considered a rigid backbone and flexible side chain model and then constructed several flexible protein models that were combinations of flexible side chains and flexible backbone segments. To determine which side chains should be considered flexible, the 12 reference protein–ligand complex structures were superimposed by the best fit to match the backbone C_α coordinates. After superposition, distances less than 6 \AA between heavy atoms of the ligand and the protein side chain were used to determine that 34 total residue side chains were considered to interact with the 12 ligands. In the approximation of a rigid backbone, for large residue side chains to adequately sample conformational space, the side chains of adjacent nearest neighbor residues should also be considered to be flexible to allow for concerted structural rearrangements. Therefore, the 34 residues comprising the binding site were linked together into six distinct segments of contiguous residues (res: 30–41, 49–55, 67–89, 103–115, 137–157, 165–177) that form the topology of the binding site. These six contiguous residue segments were then considered to all have flexible side chains, for a total of 74 residues (non-gly or -pro) with flexible side chains. The two most flexible segments in the superposition of the reference structures are residues 30–41, which will be denoted the “P-loop” (phosphate binding loop), and part of the activation loop (res: 165–177), which will be denoted the “DFG loop”. These two loops were the most obvious to include as flexible backbone segments in combination with flexible side chains. In total, eight flexible protein models were constructed, which are summarized in Table 1. Model 1 includes flexible side chains only. Models including backbone flexibility were then combined with these flexible side chains: model 2 includes backbone flexibility only in the DFG loop, while model 3 includes backbone flexibility in the two most flexible loops, the P-loop and the DFG loop. Model 4 includes backbone flexibility in all six contiguous residue segments that form the topology of the binding site. Subsequent models 5–7 were constructed and incrementally include flexibility into previously rigid segments. Model 7 is composed of one large flexible segment that defines the entire topology of the binding site (res: 30–177) but still allows 58% of the protein chain to be rigid. Model 8 includes full backbone and side chain flexibility of the entire protein (res: 5–351).

2.3. TAMD Flexible Receptor Cross-Docking. For flexible receptor docking using TAMD, the entire 12×12 cross-dock was performed for the eight different models of protein flexibility. Ligand heavy atom rmsd metrics were used to compare overall docking accuracy over all 144 cross-docking simulations (Table 2). Considering the percentage where the lowest scoring pose is $\leq 2.0 \text{ \AA rmsd}$, the docking accuracy seems to decrease proportionally to the log of the flexible receptor degrees of freedom. The docking accuracies of flexible models 1–6 are all significantly better than rigid docking. Flexible model 8, where the entire protein chain is considered to be flexible, is the only model that had a lower docking accuracy than rigid docking. The fact that the most flexible receptor model, model 8, exhibited the lowest docking accuracy agrees with observations from previous studies, suggesting that increased receptor flexibility may

Table 1. Description of Flexible Receptor Docking Models

name for flex protein model	N degrees of freedom	description	N flex backbone residues	N Flex side chains	residues composing flexible backbone segments
TAMD model 1	271	flex side chains only	0	75	none
TAMD model 2	295	1 flex loop + flex side chains	12	75	DFG loop: 165–177
TAMD model 3	317	2 flex loops + flex side chains	23	75	P-loop: 30–41, DFG loop: 165–177
TAMD model 4	430	6 flex backbone segments	83	75	30–41, 49–55, 67–89, 103–115, 137–157, 165–177
TAMD model 5	514	4 flex backbone seg.	103	N/A	30–41, 49–89, 103–115, 137–177
TAMD model 6	604	2 flex larger backbone seg.	125	N/A	30–115, 137–177
TAMD model 7	720	1 huge flex backbone seg.	147	N/A	30–177
TAMD model 8	1639	entire protein flexible	346	N/A	5–351
MD model 2	2817	1 flex loop + flex side chains	12	75	DFG loop: 165–177
MD model 3	3030	2 flex loops + flex side chains	23	75	P-loop: 30–41, DFG loop: 165–177
MD model 4	4149	6 flex backbone segments	83	75	30–41, 49–55, 67–89, 103–115, 137–157, 165–177
MD model 8	16788	entire protein flexible	346	N/A	5–351

Table 2. Cross-Docking Accuracy for a Rigid Receptor Compared to Eight Flexible Receptor Models

name for protein model	N degrees of freedom	description	best pose ≤ 2.0 Å (%)	best pose ≤ 3.0 Å (%)	any of top five ≤ 2.0 Å (%)	any of top five ≤ 3.0 Å (%)
rigid receptor 1	0	start with “swapped conf.”	15	24	38	51
rigid receptor 2	0	start with “best rotation”	15	24	25	38
rigid receptor 1	0	rescore with flex receptor LIE ₁	21	31	37	49
TAMD model 1	271	flex side chains only	40	58	61	85
TAMD model 2	295	1 flex loop + flex side chains	37	60	76	97
TAMD model 3	317	2 flex loops + flex side chains	35	51	84	94
TAMD model 4	430	6 flex backbone segments	31	51	60	84
TAMD model 5	514	4 flex backbone seg.	26	44	56	86
TAMD model 6	604	2 flex larger backbone seg.	26	47	58	83
TAMD model 7	720	1 huge flex backbone seg.	20	41	59	84
TAMD model 8	1,639	entire protein flexible	10	30	48	80
MD model 2	2,817	1 flex loop + flex side chains	35	65	65	88
MD model 3	3,030	2 flex loops + flex side chains	33	56	71	91
MD model 4	4,149	6 flex backbone segments	27	50	60	84
MD model 8	16,788	entire protein flexible	12	27	44	82

result in lower overall docking accuracy.²² However, when you consider the results including the top five scoring poses, model 8 still performs much better than rigid docking. This indicates that model 8 was able to successfully sample and then rank more native-like poses into the top five compared to rigid docking. For model 8, when low rmsd poses are in the top five, there is a diminished ability of the scoring function to correctly rank the most native pose as the best pose compared to the other flexible receptor models. For model 8, when a pose ≤ 2.0 Å rmsd is represented in the top five, it is correctly ranked as the lowest scoring pose only 21% of the time. In contrast, for model 1, where only the side chains are flexible, when a pose ≤ 2.0 Å rmsd is represented in the top five, it is correctly ranked as the lowest scoring pose 67% of the time. Models 2–4 are able to correctly rank low rmsd poses from the top five as the top pose on the order of 42–52% of the time. Therefore, the most flexible receptor model displays a reduced ability to score or rank native-like binding poses compared to the other models.

Considering only the percentage where any of the top five scoring poses is ≤ 2.0 Å rmsd in Table 2, the docking accuracy increases from rigid receptor (38%) to model 1 (61%), to model 2 (76%) and is the highest for model 3 (84%). Model 3 is the most successfully flexible receptor model for sampling and then ranking low rmsd poses into the top five (Supporting Information Figure S1). From a

careful comparison of all 144 cross-docks individually, it is clear that this improvement is due to the inclusion of backbone flexibility for the two loops. The increase in docking accuracy from 38% (rigid receptor) to 61% (model 1) reflects success for many of the less difficult cross-docking situations where only minor or major changes in side chain conformation are sufficient for successful docking. The increase in docking accuracy from 61% for model 1 to 76% for model 2 reflects success in more difficult cross-docking situations where rearrangements in the backbone conformation of the DFG loop and side chains are required for successful docking. Including flexibility in the two most flexible loops (model 3) was even more successful at ranking low rmsd poses (≤ 2.0 Å) in the top five. It was somewhat surprising that there was such a drop in docking accuracy from model 3 (two flexible loops) to model 4 (six flexible segments). From superposition of the cross-docking crystal structures, it was expected that model 4 would perform well for backbone rearrangements in the gatekeeper region (res: 103–115) required for success in a few specific cross-docks. Even though model 4 outperforms models 1–3 for a few of these specific cross-docking examples, it is less successful in many of the less challenging cross-docks that required only minor changes in side chain conformation.

Our investigation of models 4–8 was aimed at assessing ways of including binding site flexibility in the absence of multiple crystal structures or any evidence as to which parts

of the receptor binding site are flexible. If nothing was known about a receptor's flexibility, would it be better to apply backbone flexibility to multiple protein segments that form the binding site, or to all contiguous residues that form the binding site, or to the entire protein? It was surprising that, when the top five scoring poses were examined, there was not much of a difference between models 4–7. In contrast to models 1–3, models 4–7 include backbone flexibility in the hinge region (res: 106–112), which is the third most flexible backbone segment in the reference crystal structures, and atom clashes involving this region result in some rigid docking failures. It was therefore surprising that including backbone flexibility in this region did not have a significant impact. In general, our results reflect that the best docking success is achieved with the most minimal model that adequately describes the experimentally observed protein flexibility. Model 3, which has many flexible side chains and only two flexible loops, demonstrates the best performance in docking accuracy because it describes the flexibility of the receptor adequately with the fewest necessary degrees of freedom comprising the conformational search space. It would be possible to pursue this strategy for a new receptor target even in the absence of any knowledge of the flexibility of the binding site, by predicting the most flexible segments from MD simulations, normal-mode analysis,^{23,24} elastic network normal modes,²⁵ or other framework approach algorithms.²⁶ A comparison between flexible receptor models 1–8 also suggests that it would likely be better to consider only the segments forming the binding site to be flexible, rather than the entire protein. These principles can be widely extended to build efficient flexible receptor docking models for new targets of interest. For systems where large conformational changes are expected, our results suggest that it may be a good idea to use other conformational sampling approaches (including normal-mode analysis, fully flexible MD conformational sampling with solvent, etc.) in order to identify a small subset of diverse low-energy conformations. Then, when low energy conformations representing these various states are selected, local explicit flexibility in the binding site can be applied to improve docking accuracy. This approach is in contrast to docking into a large number of fully flexible structures (like flexible model 8) where the performance of a scoring function is diminished due to energetic “noise”.

Reduced docking accuracy in flexible model 8 was primarily due to significant binding site rearrangements resulting in low-scoring non-native ligand binding poses. Distortions in the geometry of the ligand binding residue side chains directly resulted in a loss of docking accuracy in flexible model 8. Over the entire cross-dock, we show that 60–70% of the native protein–ligand contacts were sufficient for docking in native-like binding modes. In the majority of the non-native binding pocket conformations, more than 50% of the ligand binding residues had undergone significant rearrangement and cannot form enough native-like ligand contacts. Specific cross-docks, where only model 8 failed, were examined in detail, and in many cases there were significant binding site conformational changes in the side chains that are not observed in any of the reference

crystal structures. In these cases, the overall backbone C_α rmsd is in the range of 1.5–2.0 Å from the starting and reference structures, but these backbone deviations are on the order of the maximum C_α rmsd between reference crystal structures (1.7 Å C_α rmsd for 1w83 and 1yqj). One similarity between these docking failures was that there were significant changes in either the backbone or side chain conformation of the majority of residues that form native contacts with the ligand. Two specific groups of conserved hydrophobic contacts in kinase scaffolds have been proposed as key contacts involved in catalytic and regulatory conformational changes: the regulatory spine (res: 75, 86, 148, and 169) and the catalytic spine (res: 38, 51, 113, 156, 157, 158, 212, and 216).^{27,28} An interesting observation was that, in many cases where only model 8 failed, conformational changes were also observed in the side chain contacts in the catalytic spine, where no changes in the catalytic spine are observed between any of the 12 reference crystal structures despite significant changes in the binding pocket. Model 8 is the only model where the entire catalytic spine is flexible, and these observed rearrangements only occurred for model 8. Therefore, it would have been possible to use this information to select low-energy conformations generated from model 8 that maintained native-like structure in the catalytic spine but allowed significant rearrangements of the binding site.

It is possible that some non-native distortions to the structure in fully flexible model 8 may be due to a lack of explicit or implicit solvent during the brief conformational sampling procedure to generate apo receptor conformations. Including implicit solvent during this conformational search and the subsequent refinement of the receptor ligand complexes would be possible, but at a great increase in computational expense. To assess this, we performed a few control simulations of this procedure using the GBMV implicit solvent model and found that the overall level of backbone (C_α rmsd) and side chain rearrangement is very similar. Apo receptor conformations generated with implicit solvent also exhibited similar binding site distortions in side chain conformations that contact the ligand, as well as side chain conformational changes involved in the catalytic spine (most notably residues Leu156, Val158, and Ile212) and in all four residues in the regulatory spine. One notable difference in the binding site conformations generated with implicit solvent is that Asp168 and Glu72 form less close contacts with Lys53, Arg67, and Arg173, presumably because the full charges on these side chains were solvated. For certain specific cross-docking failures, non-native contacts between Asp168 and Lys53 were formed, but there are also examples of close interactions between these residues in some of the reference cocrystal structures (1ouk and 1w84) compared to the others. It remains unclear if the added computational expense of the implicit solvent during receptor conformation generation would actually improve docking accuracy compared to our other models. Performing the entire cross-dock again with this procedure was beyond the scope of this manuscript. A large ensemble of fully flexible protein structures produced with high-quality explicit solvent simulations would likely retain ensemble members that contain such non-native distortions. Docking into this entire ensemble

would still be “noisy”, and therefore, we propose that a better strategy would be to select a few low-energy representative states from such an ensemble and then apply local flexibility to the binding site as mentioned above. Our results for this specific kinase system may not be generalized to all other flexible receptors, but our conclusions should hold in the regime where the amount of sampling required for visiting diverse low-energy conformational states (such as DFG-in/DFG-out) is sufficient to rearrange the majority of ligand interacting residues in the binding site.

TAMD flexible receptor docking succeeds in many difficult cases using models 2 and 3 where rigid receptor docking fails. When the cross-docking results are compared to the experimental binding data using the LIE_2 scoring functions (Supporting Information Table S1), the flexible receptor results are in much better agreement (1.1 average unsigned error (AUE) kcal/mol), as expected. The poorer agreement with the binding data for the rigid receptor is the result of lower docking accuracy, especially for atom clashes with high-affinity compounds (false negatives). There is a big improvement in docking accuracy for the three receptors in the DFG-out conformation using the TAMD flexible receptor models, where models 2 and 3 perform the best. Receptors in the DFG-out conformation are now able to successfully dock ligands that bind in the DFG-in conformation as well as the converse, ligands selective for DFG-out are also able to successfully dock to DFG-in receptors. A specific example of one of these difficult cross-docks is swapping the ligands from 1ouy (DFG-in) and 1w83 (DFG-out), shown in Figure 1A. The swapped conformation of each ligand has atom clashes with the DFG loop, requiring significant DFG loop conformational changes for successful cross-docking into a native-like conformation. For flexible receptor cross-docking of the ligand 1ouy into the receptor 1w83 (Figure 1B and C), the lowest-energy pose for this cross-dock was 1.5 Å rmsd for the ligand, and the complex contained 71% of the native protein–ligand contacts. For cross-docking of the ligand 1w83 into the receptor 1ouy, the lowest-energy pose was 2.4 Å rmsd for the ligand, and the complex contained 66% of the native protein–ligand contacts. Among the top five nonredundant poses was another pose with 1.8 Å rmsd for the ligand which contained 69% of the native protein–ligand contacts in the complex.

Despite the fact that the DFG loop does not sample the entire range of DFG loop conformations, TAMD models 2 and 3 are still able to successfully cross-dock ligands selective for the DFG-out conformation into receptors in the DFG-in conformation. Upon careful examination of these structures, many individual examples demonstrate that smaller movements of the DFG loop backbone (on the order of 3–6 Å), as well as changes in side chain conformation, are sufficient to remove the primary atom clashes between the native conformation of the ligand and the DFG loop. Thus, for cross-docking these examples, even though the generated conformations of the DFG loop are only able to come as close as 2.5–3.0 Å C_α rmsd to the reference backbone conformation of the DFG-loop (Figure 1E and F), the ligands are still capable of docking in the correct conformation because up to 70% of the native protein–ligand

contacts are properly formed and contribute the strongest anchor points for binding in the correct orientation. TAMD sampling using model 3 was also successful at correctly predicting significant backbone and side chain conformational changes in the P-loop. Figure 2 shows two separate examples of concerted changes in the two flexible loops that lead to successful cross-docking in difficult cases. In both examples in Figure 2, Tyr 35 from the P-loop undergoes a successful rearrangement from the starting conformation (purple) to the final conformation (red), which is in turn closer to the reference target conformation (blue), allowing the ligand (red) to dock in the correct conformation. Arrows are shown to highlight the movement of Tyr 35 from the starting conformation closer to the reference conformation. In the first example (Figure 2A and B), the ligand 1bl7 (DFG-in) is able to dock accurately into the DFG-out conformation because of the backbone and side chain rearrangement of Phe 169. In the second example (Figure 2C and D), the ligand 1yqj (DFG-in) is able to dock accurately into the 1di9 (DFG-in) receptor conformation. Tyr 35 on the P-loop moves over (red) to form a more native-like contact with the ligand, as in the reference structure (blue).

2.4. MD Flexible Receptor Cross-Docking. TAMD sampling was directly compared to MD sampling over the entire cross-dock set for flexible receptor models 2, 3, 4, and 8. Considering only the first docking accuracy metric in Table 2 (the percentage where the lowest scoring pose is ≤ 2.0 Å rmsd), TAMD docking accuracy is better than MD for models 2–4. MD sampling reduced the docking accuracy for cross-docks requiring significant DFG loop conformational changes. For many difficult cross-docks, TAMD sampling was able to identify a native-like low-energy conformation, and MD sampling was insufficient. For these examples, scatter plots of ligand rmsd versus LIE_1 energy show that TAMD sampling produced low-energy native-like conformations, and MD sampling was unable to sample native-like conformations at all (Supporting Information Figure S2). In other cases, MD sampling was sufficient to sample native-like conformations yet displayed a reduced ability to score or rank native-like binding poses compared to TAMD (non-native conformations score better than the native). Despite these deficiencies of MD sampling compared to TAMD sampling, MD model 2 was used for ensemble docking of 227 active compounds to determine if it was sufficient to demonstrate significant improvement over rigid ensemble docking. This choice was made as a compromise between accuracy and speed. While our rigid receptor docking protocol runs within an hour for each ligand to generate the top-ranked poses, fully flexible TAMD model 8 (or MD model 8) has comparative a runtime on the order of 34 h. MD model 2 has a slightly faster runtime on the order of 26 h when using the *confsfix* routine in CHARMM to speed up the nonbonded evaluations. These runtimes are not yet optimized, and all runtimes were evaluated on a single 2.66 GHz Intel Xeon quad-core processor.

2.5. LIE Scoring Functions Used for Receptor Ensemble Docking. In analyzing the results from docking the series of 227 active compounds into the ensemble of 12 receptor conformations, LIE_1 (derived from native scoring) was

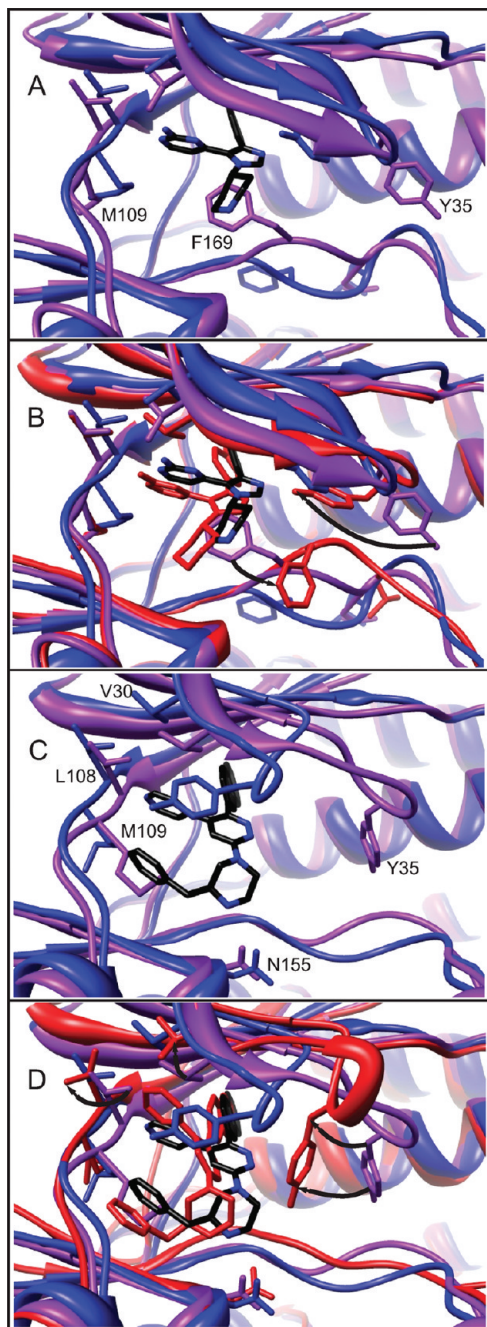


Figure 2. Successful cross-docking examples involving concerted rearrangements of backbone and side chain of the P-loop and the DFG loop. For all images, the reference receptor conformation is shown in blue, superimposed with the starting receptor conformation for a difficult cross-dock shown in purple. The reference ligand position is shown in black. In B and D, the new lowest-energy flexible receptor conformation is shown in red, along with the correctly predicted native-like ligand conformation (rmsd < 2.0 Å). In both examples, Tyr 35 from the P-loop undergoes a successful rearrangement from the starting conformation (purple) to the final conformation (red), which is in turn closer to the reference target conformation (blue), allowing the ligand (red) to dock in the correct conformation. Arrows are shown to highlight the movement of Tyr 35 from the starting conformation closer to the reference conformation. In A and B, the ligand 1bl7 (DFG-in) is able to dock accurately into the DFG-out conformation because of the backbone and side-chain rearrangement of Phe 169. In the second example (C and D), the ligand 1yqj (DFG-in) is able to dock accurately into the 1di9 (DFG-in) receptor conformation.

applied to rank the ligand conformations for both rigid and flexible receptor docking. After the top five scoring poses are identified, calculated energy components from MD are used for comparison to experimental binding data using the LIE₂ model that incorporates desolvation. For clarity in presenting results from ensemble docking, only results from the single lowest-scoring pose will be used, rather than including all results from the top five. Our objective in comparing to binding data (fitting LIE₂ to the experiment) is to assess if the top-scoring ligand poses derived from flexible receptor docking are able to provide better agreement with binding data than rigid receptor docking.

In analyzing results from ensemble docking using the LIE₂ model, much better fits to the experimental data were achieved by adding an entropic term related to the number of freely rotatable bonds²⁹ in each given ligand that becomes immobilized upon binding to the receptor: $[\Delta G_{\text{binding}} = \alpha(\Delta_{\text{VDW}}) + \beta(\Delta_{\text{ELEC+GBE}}) + \gamma(N_{\text{rotatable_bonds}})]$. This term had not been included previously in LIE₂ because the cross-docking test set showed minimal improvement upon adding the entropic term. Fits to the LIE₂ model yield values of γ that are within the range of other values (0.1–0.6 kcal/mol) reported in the literature for the entropic contribution to $\Delta G_{\text{binding}}$ from a single freely rotatable bond in a ligand.^{29,30} The parameter γ was found to be (rigid $\gamma = 0.19$, flex $\gamma = 0.35$) when the docking results for all 12 receptors were fit simultaneously to binding data. Comparisons to binding data are calculated in two ways: first, all 12 receptors were fit simultaneously to the binding data using a single LIE₂ model, and second, each receptor was fit individually to the experimental data with its own LIE₂ model. The second approach has smaller errors in kilocalories per mole, as expected, but an important finding is that these two fitting approaches both show similar qualitative trends and few differences in comparing rigid and flexible receptors.

2.6. Docking Accuracy for RED. Although there are not crystal structures for all of the 227 active compounds, a reasonable approximation can be made for the ligand binding conformations assuming that the vast majority of active compounds will adopt binding modes similar to known crystal structures of similar compounds. There are six major series of related compounds: (1) 2-aminopyrimidine carbamates,³¹ (2) 4-azaindoles,¹⁹ (3) aminopyrazoles,³² (4) benzimidazolones,³³ (5) indole and pyridine fragments,¹⁸ and (6) triazolopyridine oxazoles.³⁴ All ligands that bind in the DFG-in conformation share at least two anchor points or pharmacophores (Figure 3): (1) a hydrogen bond or close interactions with the backbone and side chains in the gatekeeper region of the active site (res: 107–110H) and (2) a deep hydrophobic pocket that binds aromatic groups (T106, L75, and L104). These two anchor points can be used to determine the overall orientation of each ligand docking pose. Series 2, 4, and 6 also have a third unique anchor point which is a hydrophilic interaction with the charged NH₃ group of K53. Ligand binding poses were considered to bind in a native-like orientation if they satisfied two or more of these anchor points, and this criterion was used to approximate the overall docking accuracy for each receptor (Table 3). Hydrogen bonding to the hinge region in Figure

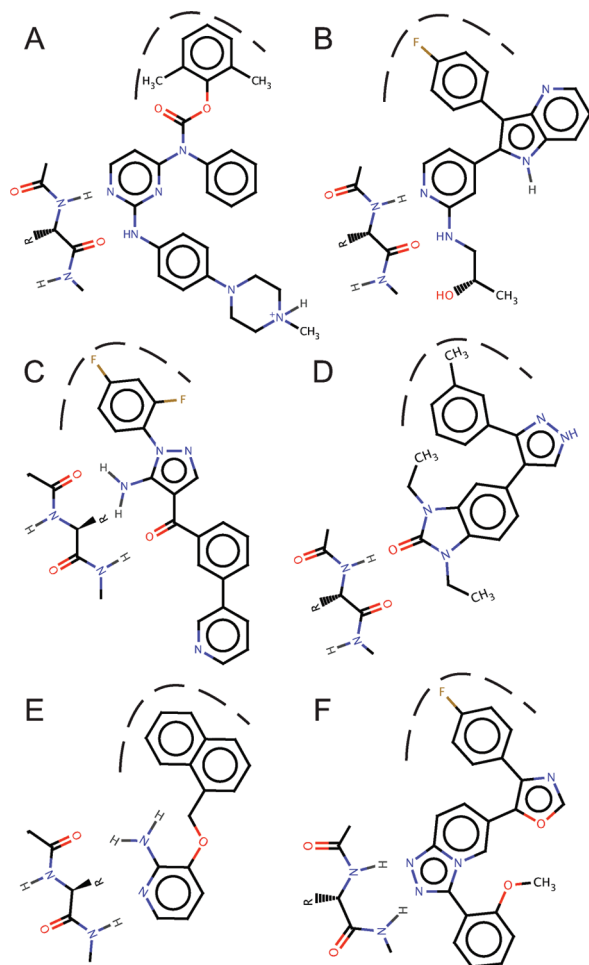


Figure 3. Two common protein–ligand interactions (anchor points), shared by all six series of active compounds binding in the DFG-in receptor conformation. An interaction with the hydrophobic pocket (dashed line) and hydrogen bonding to residues in the gatekeeper region are common to all six series of active compounds: (A) 2-aminopyrimidine carbamates, (B) 4-azaindoles, (C) aminopyrazoles, (D) benzimidazolones, (E) indole and pyridine fragments, and (F) triazolopyridine oxazoles.

3F shows the peptide plane flip (from structure 1zzl) that is also present in the structure for 1ouk and 1ouy. All compounds are able to hydrogen-bond to the hinge region in either conformation, although the ideal ligand interactions may not be formed.

The highest docking accuracy using a rigid receptor model was found for receptors 1ouk (67%) and 1ouy (64%). The receptor 1ouk also had the best docking accuracy in the cross-docking study according to ligand rmsd metrics. This is an interesting finding in light of a recent solution NMR study which concluded that the predominant ensemble in solution is most similar to the crystal structures of 1ouk and 1ouy.³⁵ This NMR study compared a set of 73 measured residual dipolar couplings (RDCs) from the apo conformation in solution to various crystal structure conformations.³⁵ Receptor conformations 1ouk and 1ouy exhibit a peptide plane flip over the gatekeeper residues M109 and G111, which participate in backbone hydrogen bonding with ligands. Comparison of the measured RDCs upon binding the pyridinylimidazole inhibitor 1a9u demonstrated evidence for

a structural transition over residues M109 and G111 from the receptor conformation (1ouk) to the crystallographic conformation (1a9u).³⁵ Therefore, prior to binding, the predominant conformation of the apo receptor was most similar to 1ouk and 1ouy. This experimental observation rationalizes the observed high docking accuracy and good agreement with binding data for rigid receptor docking with 1ouk and 1ouy. It is also interesting that the lowest rigid receptor docking accuracy is observed for the receptor conformation of 1a9u, and this can be attributed to the peptide plane flip in the gatekeeper region as well as a very closed conformation of the P-loop. Low docking accuracy for the ligand 1a9u in cross-docking can also be explained by NMR evidence that this ligand binds in more than one conformation.^{35,36}

There are several receptors that perform very poorly for rigid docking that show improvements in docking accuracy upon adding flexibility. The docking accuracy for individual rigid receptors ranges from as low as 11% for 1a9u to as high as 67% for 1ouk. For 5 of 12 receptors with relatively low rigid docking accuracy (1a9u, 1bl6, 1bl7, 1w83, and 1w84), the flexible docking model improves the docking accuracy. However, for five other receptors that had a comparatively high rigid receptor docking accuracy (1di9, 1ouk, 1ouy, 1oz1, and 1yqj), including flexibility slightly diminished overall docking accuracy and overall agreement with binding data. In this way, flexibility improved agreement with experimental results for receptors that perform poorly in rigid docking and reduced agreement with experimental results for the receptors that had the best results for rigid docking. However, despite this small decrease in overall docking accuracy for these receptors, flexible receptor docking still achieved overall improvement in ranking compounds when compared to binding data.

2.7. RED Comparison to Binding Data. Comparison to experimental binding data shows that the flexible receptor model has an improved ability to rank compounds. When all 12 receptors are fit to binding data simultaneously, 8 of 12 have a better linear correlation (R^2) with the experimental data with a flexible receptor. When the receptors are fit individually, 7 of the 12 receptors showed improved agreement in terms of error (AUE) in kilocalories per mole and linear correlation (R^2) with a flexible receptor. In a comparison of each individual receptor to experimental binding data, the flexible receptor model demonstrated improved enrichment of high-affinity compounds over low-affinity compounds (Table 3, Supporting Information Table S2). Results from rigid and flexible docking are compared by calculating the enrichment factor (EF) for each receptor on the basis of the number of high affinity hits in the top n ranked compounds. EFs were calculated [$EF = (Hits_n/NC_n)/(Hits_{total}/NC_{total})$], where $Hits_n$ is the number of true high-affinity actives retrieved in NC_n (the top n ranked compounds), $Hits_{total}$ is the total number of high-affinity actives, and NC_{total} is the total number of compounds. Enrichment factors were calculated for two sets of the highest-affinity compounds. The first group of high-affinity compounds was 46 of 227 ligands that have $\Delta G_{exp} < -11.0$ kcal/mol, and

Table 3. Receptor Ensemble Docking Results: Docking Accuracy, Comparison to Binding Data, and Enrichment Factor (EF₁Max)^a

id	PDB		docking accuracy	fit all simultaneously				fit each individually			
				AUE (kcal/mol)	rms (kcal/mol)	R ²	EF ₁ Max	AUE (kcal/mol)	rms (kcal/mol)	R ²	EF ₁ Max
1	1a9u	rigid	0.11	1.90	2.24	0.04	0.71	1.82	2.16	0.03	0.22
1	1a9u	flex	0.22	2.11	2.54	0.12	1.76	1.68	2.11	0.10	1.57
2	1bl6	rigid	0.20	1.87	2.26	0.05	0.73	1.76	2.16	0.04	0.65
2	1bl6	flex	0.41	1.86	2.23	0.14	1.71	1.65	2.07	0.14	1.62
3	1bl7	rigid	0.32	1.88	2.24	0.04	0.58	1.87	2.23	0.04	0.60
3	1bl7	flex	0.38	1.87	2.29	0.10	0.73	1.67	2.09	0.09	0.68
4	1di9	rigid	0.27	1.78	2.16	0.19	2.42	1.43	1.80	0.18	2.25
4	1di9	flex	0.20	2.03	2.50	0.13	1.47	1.64	2.09	0.12	1.46
5	1kv1	rigid	0.15	2.37	3.13	0.01	0.55	2.32	2.94	0.01	0.52
5	1kv1	flex	0.04	2.83	3.61	0.02	0.83	1.57	1.99	0.12	1.33
6	1kv2	rigid	0.26	2.25	2.90	0.00	0.36	2.23	2.71	0.01	0.33
6	1kv2	flex	0.03	2.05	2.63	0.09	0.76	1.40	1.75	0.22	1.69
7	1ouk	rigid	0.67	1.53	1.85	0.15	0.93	1.49	1.84	0.16	1.11
7	1ouk	flex	0.33	1.95	2.33	0.10	1.22	1.64	2.05	0.10	1.19
8	1ouy	rigid	0.64	1.57	1.92	0.12	0.93	1.57	1.92	0.12	0.89
8	1ouy	flex	0.46	1.97	2.40	0.07	1.31	1.75	2.20	0.06	1.22
9	1oz1	rigid	0.49	1.49	1.87	0.21	1.38	1.41	1.78	0.19	1.07
9	1oz1	flex	0.18	1.79	2.21	0.11	1.16	1.61	2.06	0.12	1.33
10	1w83	rigid	0.05	1.95	2.52	0.00	0.35	1.67	2.11	0.07	1.33
10	1w83	flex	0.08	1.85	2.43	0.06	0.73	1.52	1.93	0.14	1.67
11	1w84	rigid	0.20	1.65	2.08	0.18	1.46	1.53	1.90	0.14	0.95
11	1w84	flex	0.21	1.51	1.89	0.19	2.17	1.49	1.88	0.21	2.50
12	1yqj	rigid	0.50	1.62	1.96	0.08	0.49	1.49	1.85	0.17	0.93
12	1yqj	flex	0.48	1.74	2.11	0.11	1.21	1.60	1.98	0.12	1.30

^a The enrichment factor EF₁ was calculated for the first group ($\Delta G_{\text{exp}} < -11.0$ kcal/mol) of high-affinity compounds. The maximum observed EF₁ (EF₁ Max) within the top 20% of ranked compounds is reported here. The maximum theoretical value of EF₁ is 4.9. Additional enrichment factors analysis is reported in Supporting Information Table S2.

the second group was 89 of 227 ligands with $\Delta G_{\text{exp}} < -10.0$ kcal/mol. The enrichment factor EF₁ was calculated for this first group 1 ($\Delta G_{\text{exp}} < -11.0$), and the enrichment factor EF₂ was calculated for group 2 ($\Delta G_{\text{exp}} < -10.0$ kcal/mol). Given these definitions, the maximum theoretical value of EF₁ is 4.9, and the maximum theoretical value of EF₂ is 2.6. A third group of 57 low-affinity compounds was selected from the 227 using the criterion $\Delta G_{\text{exp}} > -8.0$ kcal/mol. This third group of low-affinity compounds was used to calculate the number of false positives, assuming that group 2 (true positives) could be successfully separated from group 3. A large majority of the individual receptors (10 of 12 receptors) had improved enrichment factors with a flexible receptor (Figure 4), and the overall enrichment over the ensemble of 12 receptors is better for the flexible receptor model (Supporting Information Tables S2 and S3). The flexible receptor model also showed significant enrichment compared to the rigid receptor model in separating these 227 active compounds from a large number (>9000) of decoy molecules (Figure 5A and B). These decoy molecules, which are specific for p38 α , were taken from the database of useful decoys (DUD) and represent a very difficult decoy set because the ligands were selected to be similar to the actives in several molecular properties.³⁷ As 28 of our 227 active compounds are greater in molecular weight (MW) than the DUD decoy molecules, we also calculated the enrichment factors considering only 153 active ligands that share the same molecular weight range ($450 > \text{MW} > 320$) as the DUD decoys to eliminate any source of artificial enrichment (Figure 5C and D).³⁸ Either approach to calculating the enrichment factor shows significant improvement of the

flexible receptor model compared to the rigid receptor for either DFG-in or DFG-out conformations (Figure 5).

As our results show that certain receptors perform well for one ligand series, while not so well for others, we investigated several combinations of minimal receptor ensembles. A minimal receptor ensemble of four receptors (1a9u, 1di9, 1ouk, and 1oz1) was constructed using the best rigid receptor for each series respectively. When these rigid receptor docking results are fit simultaneously to the LIE₂ model, it achieves much better agreement with binding data than any individual receptor (Figure 6) and shows that it is possible for this LIE₂ model to rank all six series simultaneously using a MRC ensemble approach. For the six major series of related compounds, the receptor with the best agreement was used for each series respectively: (1) 2-aminopyrimidine carbamates (1ouk), (2) 4-azaindoles (1a9u), (3) aminopyrazoles (1a9u), (4) benzimidazolones (1oz1), (5) indole and pyridine fragments (1di9), and (6) triazolopyridine oxazoles (1ouk). The best linear fit for this model through all points (AUE = 1.1 kcal/mol, rms = 1.4 kcal/mol, and $R^2 = 0.41$) is shown in black in Figure 6A. Figure 6 shows that the best fit receptor for each series has a slightly different slope, even when fit simultaneously. The agreement with the binding data for each series is as follows: (1) 2-aminopyrimidine carbamates ($R^2 = 0.30$), (2) 4-azaindoles ($R^2 = 0.24$), (3) aminopyrazoles ($R^2 = 0.19$), (4) benzimidazolones ($R^2 = 0.16$), (5) indole and pyridine fragments ($R^2 = 0.72$), and (6) triazolopyridine oxazoles ($R^2 = 0.17$).

2.8. Comparison to Other Flexible Receptor Docking Studies with p38 α . As p38 α is considered to be a representative ser/thr kinase with strong induced fit effects, it has been

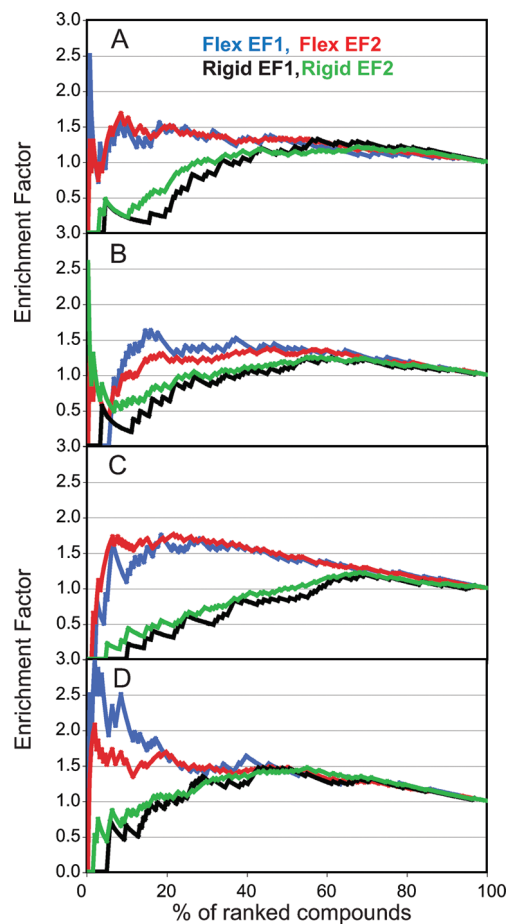


Figure 4. Enrichment factors for rigid and flexible receptor docking into receptors (A) 1a9u, (B) 1bl6, (C) 1kv2, and (D) 1w84. These enrichment factors EF₁ and EF₂ are calculated assuming that groups of high-affinity compounds can be separated from groups of low-affinity compounds.

used previously in several other flexible receptor docking studies.^{39–42} A small subset of our cross-dock can be compared to other studies with ICM³⁹ and Glide⁴¹ for docking accuracy, where a docking success is defined as the single lowest-energy structure ≤ 2.0 Å rmsd. For a very small 4×4 comparison with ICM,³⁹ our flexible model 1 was the best with a 38% accuracy, which was lower than ICM with a 63% accuracy; however, this lower performance is predominantly from poor docking accuracy of the ligand for 1a9u, which has been shown to bind experimentally in more than one binding mode.^{35,36} For a larger comparison with Glide⁴¹ (3×3 DFG-out; 9×9 DFG-in), flexible model 1 was the best with 46% accuracy, where Glide was 60% accurate.

Early database enrichments were also measured in these two studies using ICM³⁹ and Glide,⁴¹ and our early enrichment in the top 1–5% of the ranked database (Figure 5C and D) is roughly comparable given that these two studies have different lists of actives and decoys. These two studies also both used on the order of 1000 decoys, and we have used more than 9000 decoys. In the ICM study for the receptor 1a9u, enrichment factors for the top 1%, 2%, and 10% were 16.7, 8.3, and 5.0, respectively, whereas our corresponding EFs from Figure 5C for 1a9u are 7.1, 3.0, and 1.1, respectively. The early enrichments for our rigid

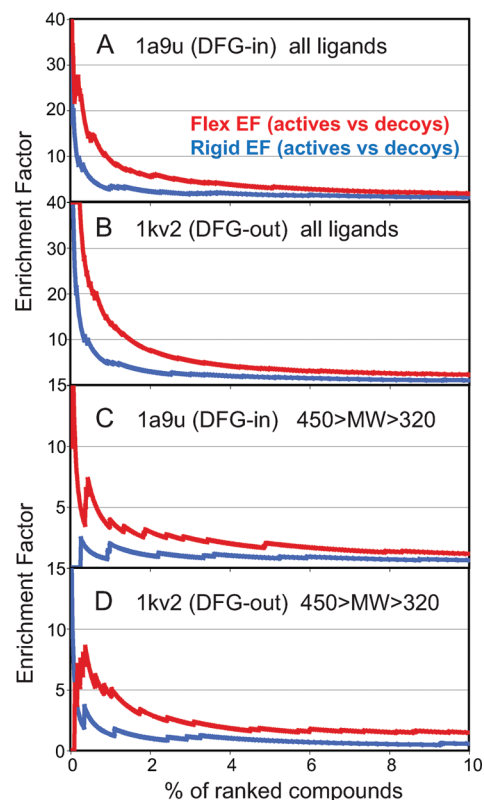


Figure 5. Enrichment factors for separating active compounds from more than 9000 decoy molecules. Enrichment of all 227 actives from decoys for receptor (A) 1a9u and (B) 1kv2. Enrichment of 153 actives in the same molecular weight range (320–450) as the decoys for receptor (C) 1a9u and (D) 1kv2.

docking results for receptor 1kv2 are also reasonable compared to rigid docking performed in the DUD data set paper with DOCK,³⁷ where the EF at 1% was 2.1 with DOCK and our respective EF at 1% was 1.4, although our EF values range from above 10 to 3.3 for multiple actives found within the top 1%. In a recent report comparing ICM and Glide for enrichment of 111 actives against 6450 decoys, the area under a receiver operating characteristic (ROC) curve was reported for enrichment.⁴² In the ROC curve, the signal (fraction of true positives) is plotted on the y axis against the noise (fraction of false positives) on the x axis, and the area under the curve (AUC) indicates the overall signal-to-noise, where 1.0 is theoretically perfect performance and 0.5 is a random performance. In the above-mentioned study, the ROC AUC values for single receptor docking were 0.64 for Glide and 0.70 for ICM, and the ROC AUCs for multiple receptor docking were 0.81 for Glide and 0.88 for ICM.⁴²

We have calculated the ROC AUC for our actives with the DUD decoy set (Figure 5C and D) for the molecular weight range of 320–450 and have found that the ROC AUC for the flexible receptor is 0.47 (1a9u) and 0.44 (1kv2). These are both less than 0.5 and are therefore worse than random when considering ranking the entire database, despite the fact that these data sets show reasonable early enrichment within the top 1–10% of the data set. However, the ROC AUCs are still better for the flexible receptor compared to the rigid receptors 0.42 (1a9u) and 0.36 (1kv2). One reason for this poor apparent performance is due to the large number

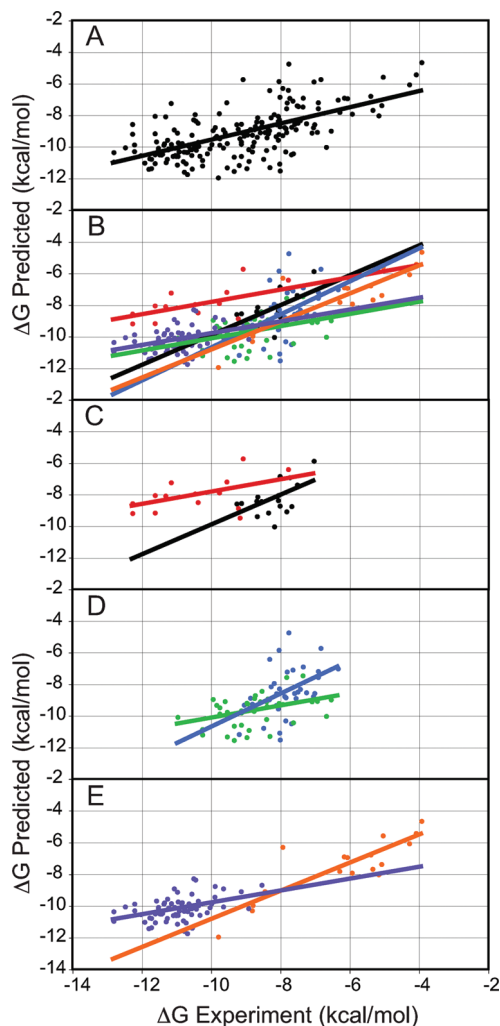


Figure 6. Comparison to binding data, using a minimal ensemble of four rigid receptor conformations (1a9u, 1di9, 1ouk, and 1oz1). The best linear fit for this model through all points is shown in A in black. The best linear fit is shown in B, C, D, and E for each of the six series: (1) 2-aminopyrimidine carbamates (red), (2) 4-azaindoles (black), (3) aminopyrazoles (light blue), (4) benzimidazolones (green), (5) indole and pyridine fragments (orange), and (6) triazolopyridine oxazoles (purple).

of low MW actives in our set, where high MW decoys score better than low MW actives. When only the actives and the decoys in the high MW range of 375–450 are compared (52 actives and 4432 decoys), then our enrichment studies have a ROC AUC of 0.64 (1a9u) and 0.64 (1kv2) for the flexible receptors. These ROC AUC values are both better than random over the entire database and more similar to values reported for single receptor docking for Glide and ICM.⁴² Another reason for the lower ROC AUC for the entire MW range of 320–450 may be that the decoys in the DUD data set may be more similar to p38 α actives than the decoy set in the study using Glide and ICM.⁴² Our results show that, for DUD, reasonable, early enrichment is observed in EF factors but that much better overall enrichment occurs in a comparison of actives and decoys in the MW range of 375–450 compared to the MW range of 320–375. We suspect that this may be because greater MW ligands have

more specific contacts with the receptor compared to their decoys and that lower MW decoys are more similar to low MW actives.³⁸ In general, although our results for the flexible receptor models are much better than our rigid receptor results, these flexible receptor models are still not performing as well as ICM and Glide.

3. Computational Methods

3.1. Linear Interaction Energy Scoring Function for Docking. For ranking docking poses of the same ligand, using the total energy of the protein–ligand complex with a RDIE is sufficient,⁴³ as long as the receptor is rigid. When flexibility is introduced into the protein during the conformational search, the total energy of a given protein–ligand complex is no longer a reliable metric to distinguish low-energy native-like poses. This lower predictive power is due to “noise” from conformational changes in intraprotein interactions, which are energetically equivalent or greater than conformational changes in protein–ligand interaction energy. For this reason, when protein flexibility is introduced into the conformational search problem, a scoring function that is based on energetic components from protein–ligand interaction is more effective. In a previous study, CHARMM protein–ligand interaction energies performed well compared to eight other commonly used scoring approaches for the discrimination of docked and misdocked structures.⁴³ Although this approach works well for rigid docking, for flexible protein docking, we have employed LIE models⁴⁴ for our CHARMM protein–ligand interaction energies of the generic form $\Delta G_{\text{binding}} = \alpha(\Delta_{\text{component1}}) + \beta(\Delta_{\text{component2}}) + \gamma(\Delta_{\text{component3}})$ as the sum of potential energy components with weighted coefficients. Following the work of Caflisch and co-workers,⁴⁴ we have used a LIE model of the form $\Delta G_{\text{binding}} = \alpha(\Delta_{\text{VDW}}) + \beta(\Delta_{\text{(ELEC+GBE)}}$ for predicting the absolute $\Delta G_{\text{binding}}$ in order to compare ligands of various sizes and shapes (to rank a series of ligands). In the current study of protein flexibility, we use a less computationally expensive LIE model of the form $\Delta G_{\text{binding}} = \alpha(\Delta_{\text{VDW}}) + \beta(\Delta_{\text{ELEC}})$ for ranking binding poses of the same ligand where both the protein and ligand are flexible simultaneously. This simplified LIE score avoids the computationally expensive calculation of generalized Born (using GBMV) solvation energy components for every docking pose. Using this model assumes that differences in solvation energy should be a minor component when considering poses of the same ligand. Therefore, our approach to flexible receptor docking uses a hierarchical two-step scoring approach to reduce computational expense,^{45,46} where a simplified “LIE₁” model is used to select the best-scoring poses, which are then rescored using a more sophisticated and computationally expensive “LIE₂” model that includes desolvation.

In order to construct these LIE scoring functions for flexible receptor docking, energetic components acquired from “native scoring” and from self-docking using a rigid receptor were compared. Native scoring LIE components were calculated from both the minimized static native structure and MD simulations started from these structures. Three metrics of comparison are used to compare predicted

$\Delta G_{\text{binding}}$ to experimental $\Delta G_{\text{binding}}$ values: average unsigned error (AUE) in kilocalories per mole, root-mean-square error (rms) in kilocalories per mole, and the best fit linear correlation coefficient (R^2). When the parameters α and β are fit to the energy components from native scoring, the more sophisticated LIE_2 model including the desolvation effects [$\Delta G_{\text{binding}} = \alpha(\Delta_{\text{VDW}}) + \beta(\Delta_{\text{ELEC+GBE}})$] demonstrates an improved agreement with experimental results as expected compared to LIE_1 without solvation [$\Delta G_{\text{binding}} = \alpha(\Delta_{\text{VDW}}) + \beta(\Delta_{\text{ELEC}})$]. The best fit (AUE = 1.0 kcal/mol) was from the LIE_2 model applied to ensemble averages over MD simulation. Given the reasonable transferability of these LIE_1 fits between native scoring and self-docking, we used them as a scoring function for flexible receptor docking.⁴⁴ As an additional check, all of the docking poses collected from self-docking were rescored using the LIE_1 scoring function, and native-like ligand poses with heavy atom rmsd ≤ 2.0 Å were still obtained as the lowest-energy poses. Cross-docking results described below show that using LIE_1 scores results in improved cross-docking accuracy compared to using the total energy of the complex, for both rigid receptor docking and for flexible receptor docking.

3.2. Preparation of Cross-Docking Set. Our molecular docking approach uses the program CHARMM (Chemistry at HARvard Molecular Mechanics) for an all-atom potential energy description of protein–ligand complexes.⁴⁷ The CHARMM force field, originally parametrized by Momany and Rone⁴⁸ has been extended to describe ligands in the Ligand–Protein Database (LPDB) and was used to build potential energy functions for all ligands in this study.⁴⁹ The crystal structure coordinates of 12 protein–ligand complex structures of p38 α MAP kinase were downloaded from the PDB database.^{15–21} All structural water molecules were removed, and standard protonation states for neutral pH were applied to all ionizable protein residues. Appropriate protonation states were applied to ligands by calculating pK_a 's for titratable groups using both Marvin (<http://www.chemaxon.com>) and the Web server SPARC (<http://ibmlc2/chem.uga.edu/sparc/>).⁵⁰ Hydrogen atoms were built for ligands using open Babel (<http://openbabel.org>), and all protein hydrogen atoms were built using CHARMM. For ligands with multiple possible tautomers (such as 1a9u), only the single tautomer that allows a close favorable interaction (between the NH_3 group of Lys 53 and imidazole nitrogen) inferred from the cocrystal structure was used.

The activation loop residues D168 and F169 form the characteristic inactive “DFG-out conformation” in the structures of 1kv1 and 1kv2 and 1w83. The structures 1kv1 and 1kv2 are missing electron density for the rest of the flexible residues 170–184 in the activation loop. Flexible loop conformations for this missing electron density were built using Modeler and the MODloop server (<http://modbase.compbio.ucsf.edu/modloop/modloop.html>).⁵¹ For 1kv1 and 1kv2, several conformations of the activation loop were considered, but the lowest-energy models similar to the DFG-out conformation of 1w83 were selected. The all-atom protein–ligand complexes were minimized for 200 steps in vacuo, and this was taken as the experimental reference conformation of all structures.

3.3. Rigid Receptor Docking (CDOCKER). In our rigid receptor-docking model, a 1.0 Å grid is used to describe the static protein conformation of the binding site, where the interaction energies of 20 types of probe atoms are calculated for every point on the grid. The 3D grid is calculated to extend 8 Å in all directions from any atom in the ligand. Following a simulated annealing conformational search of the flexible ligand, the grid is removed and the all-atom protein–ligand representation is minimized, fixing the coordinates of the protein, using the standard all-atom potential function with a distance-dependent dielectric (RDIE). This interaction energy is taken as the score for the final ligand pose. Other details of previous CDOCKER setups and protocols have been published elsewhere.^{52–55} For each given protein–ligand complex, 10 docking trials were performed, composed of 1000 individual docking attempts (100 random conformations \times 10 random rotations) for a total of 10 000 docking attempts per complex. Within each docking trial, the interaction energy scores of the final ligand poses are sorted and clustered to identify the “top five” nonredundant clusters of docking poses. After these “top five” docking poses are identified, then a short MD simulation of the entire protein–ligand complex is performed for each of the “top five” in order to more accurately predict $\Delta G_{\text{binding}}$. In these short MD simulations, the electrostatic contribution to the solvation energy is approximated by using a generalized Born implicit solvent potential energy term (GBMV).^{56,57} For the “top five” docking poses, MD simulations are performed using a 16 Å nonbonded cutoff and the GBMV implicit solvent model. These MD simulations are performed at 300 K using harmonic restraints on the C_α atoms of the protein backbone, starting with 500 (2 fs) MD steps for equilibration. The harmonic restraints are then released for an additional 1000 (2 fs) steps of production MD, which is used to calculate the ensemble average protein–ligand interaction energies. The first 50 steps are skipped, and then ensemble average interaction energy properties are calculated from 950 structures for the LIE_2 model.

3.4. TAMD Flexible Receptor Docking. The TAMD method uses an efficient Newton–Euler inverse mass operator algorithm for solving the equations of motion in internal coordinates.⁵⁸ Molecules are represented with a branched tree structure of rigid body clusters connected by torsional hinges. Torsional cross-terms are constructed from local molecular fragments, using a soft-core potential to introduce implicit bond and angle flexibility into the rigid geometry approximation. The use of the soft-core potential effectively removes high-energy barriers on the IC potential energy surface and facilitates more efficient sampling of conformational space. Additional details about the TAMD method have been published elsewhere.⁵⁹

The flexible receptor docking method differs from the rigid docking method in that it never uses a grid representation of the protein. An overview of the flexible docking procedure is shown in Figure 7. In the first step of the flexible receptor docking procedure, random conformations of the ligand are generated for a total of 200 ligand conformations. The next step is to perform TAMD conformational sampling of the

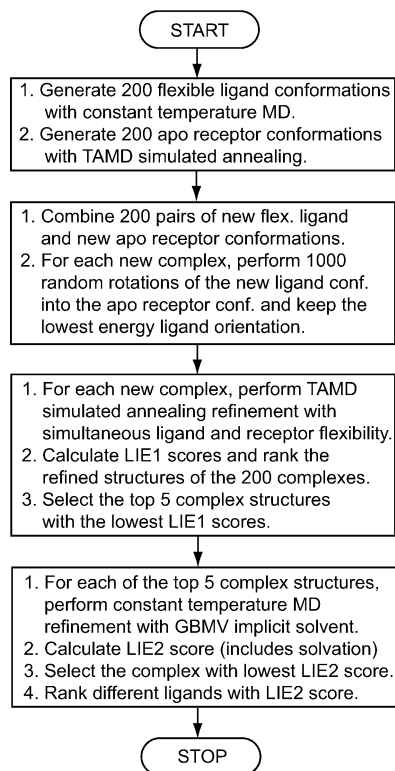


Figure 7. Overview of the flexible receptor docking protocol.

receptor in the absence of the ligand to generate a diverse ensemble of 200 flexible receptor conformations. During this conformation generation procedure, a 14 Å nonbonded cutoff is used along with a RDIE and a soft-core potential. When TAMD is used for simulated annealing, quite large temperatures can be effectively employed for rapid conformational sampling, depending on which model of protein flexibility is used. Models with fewer flexible degrees of freedom can use larger temperatures for sampling without integrator stability problems. For flexible model 2, heating is performed from 2000 to 2500 K, while for flexible model 8, heating is performed from 500 to 700 K and then the system is cooled from 700 K down to 300 K. These protocols yield overall protein rmsd's that are in the range of 0.3–3.0 Å C_{α} rmsd from the starting receptor conformations.

In the next step, the 200 random flexible receptor conformations are docked with 200 random flexible ligand conformations. For each new random pair of flexible ligand and receptor conformations, 1000 ligand rotations are performed around a reference ligand's center of mass, to identify the optimal rotation conformation of the flexible ligand in the flexible receptor conformation. For this search for an optimal rotation in the binding site, a soft-core potential and all-atom models are used for both the ligand and the receptor. Next, the structure of the protein–ligand complex is refined using TAMD simulated annealing, which allows for simultaneous flexibility of the protein and ligand. The heating and cooling cycle is carried out for all complexes using 1000 steps of 2 fs TAMD, where the system is heated from 500 to 700 K over 1000 steps and then cooled from 700 K back down to 300 K over 1000 steps. During this simulated annealing cycle, significant rearrangements of both

the protein and ligand are possible. Although some complexes do not move far from their starting structure because they contain favorable protein–ligand interactions, other less favorable complexes undergo significant conformational rearrangements. For flexible protein model 2, changes characterized by atom movements as large as 6–8 Å are observed for both ligand and flexible protein side chains in numerous examples. The final TAMD generated protein–ligand complex structures are then minimized for 1000 steps using the soft-core potential. Then, the soft-core potential is removed and replaced with the standard potential function with a distance-dependent dielectric (RDIE). Interaction energies and potential energy components are calculated from this final minimized protein–ligand complex and are used to calculate the LIE₁ score to energy-rank the TAMD generated flexible receptor complexes. Similar to rigid receptor docking, these LIE₁ scores for the final flexible protein–ligand complex conformation are sorted and clustered to identify the “top five” nonredundant clusters of docking poses. As for rigid receptor docking, then a short MD simulation of the entire protein–ligand complex is performed using the GBMV implicit solvent model for each of the “top five” in order to calculate the potential energy components for the LIE₂ score.

3.5. RED for a Series of 227 Active Compounds. In order to extend our study from the $N = 12$ cross-dock, we also considered the docking of 227 active compounds that have IC₅₀ binding data for human p38 α reported in the Binding Database (www.bindingdb.org).⁶⁰ Approximate binding free energies were calculated from IC₅₀ data using solution information and the relationship ($K_i = IC_{50}(K_m/(S + K_m))$). The covalent structure and initial low-energy 3D conformations of each ligand were generated from smiles strings using Marvin. Appropriate protonation states were applied to ligands by calculating pK_a's for potential titratable groups using both Marvin and the Web server SPARC. Decoy molecules specific for p38 α were taken from the DUD, and the protonation states from DUD were used.³⁷ All 227 ligands were docked to the ensemble of 12 receptor conformations using both the rigid and flexible receptor-docking model. For flexible receptor docking, flexible protein model 2 was used with MD flexible receptor conformational sampling in order to determine if MD sampling alone was sufficient to achieve improvements over rigid docking.

4. Conclusions

Several models were constructed to represent the flexibility of the binding site for p38 α MAP kinase, which exhibits significant binding site flexibility in ligand-bound cocrystal structures. These flexible receptor models scaled from the inclusion of (1) flexible side chains (model 1), (2) one or more flexible backbone segments (models 2 through 7), and (3) treatment of the entire protein as flexible (model 8). As constructed, these models scale in the number of total protein degrees of freedom that describe the conformational search space. Models 2 and 3 demonstrate superior performance in cross-docking accuracy because they describe the flexibility of the system adequately with the fewest degrees of freedom. We show that the fully flexible protein model 8 suffers from

reduced docking accuracy. In fully flexible protein model 8, the scoring function docking accuracy is diminished and more “noisy” due to extensive binding site rearrangements involving the majority of the ligand interacting residues. While some of these conformational changes are native and observed in crystal structures leading to docking successes, other non-native rearrangements result in low-scoring non-native ligand binding poses.

In a direct comparison of TAMD sampling to MD sampling for these various models of protein flexibility, TAMD sampling was found to be both more efficient and more accurate. However, even using MD sampling, the percentage of cross-docking success is significantly greater with the flexible receptor models 2 and 3 than for the rigid receptor models. This comparison demonstrates that the flexible receptor models succeed in many difficult cases where the rigid receptor model fails and that reasonable “native-like” receptor conformations are sampled regardless of the starting conformation. Using MD sampling and flexible receptor model 2, we performed RED using the same 12 receptor conformations and 227 active compounds. The comparison to experimental binding data showed that the flexible receptor model showed improved agreement with binding data and an improved ability to rank compounds. In a comparison of each individual receptor, the flexible receptor model also showed improved enrichment of true positives in top-ranked compounds. The flexible receptor methodology described in this paper is most applicable for proteins that show strong induced fit effects.

Abbreviations

fs, femtosecond; ps, picosecond; MD, molecular dynamics; TAMD, torsion angle molecular dynamics; LPDB, Ligand-Protein Database; RDIE, distance-dependent dielectric function; LIE, linear interaction energy; ROC, receiver operating characteristic; AUC, area under the curve; MW, molecular weight; AUE, average unsigned error; rmsd, root mean squared deviation.

Acknowledgment. This work has been supported by the National Institutes of Health to C.L.B. (GM037554) and to R.S.A. (GM076836), in the form of an NRSA Postdoctoral Fellowship. Molecular graphics images were produced using the UCSF Chimera⁶¹ package from the Resource of Bio-computing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081). Marvin was used for drawing all chemical structures (Marvin, v. 4.0.5; ChemAxon: Budapest, Hungary, 2007; <http://www.chemaxon.com>).

Supporting Information Available: Supporting Information is available in the form of two additional figures (S1 and S2) and three tables (S1, S2, and S3). This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- Carlson, H. A. *Curr. Opin. Chem. Biol.* **2002**, *6*, 447–52.
- Abagyan, R.; Totrov, M. *Curr. Opin. Chem. Biol.* **2001**, *5*, 375–82.
- Gerstein, M.; Krebs, W. *Nucleic Acids Res.* **1998**, *26*, 4280–90.
- Kuntz, I. D.; Meng, E. C.; Shoichet, B. K. *Acc. Chem. Res.* **1994**, *27*, 117–123.
- Polgar, T.; Baki, A.; Szendrei, G. I.; Keseru, G. M. *J. Med. Chem.* **2005**, *48*, 7946–59.
- Zavodszky, M. I.; Kuhn, L. A. *Protein Sci.* **2005**, *14*, 1104–14.
- McGovern, S. L.; Shoichet, B. K. *J. Med. Chem.* **2003**, *46*, 2895–907.
- Mazur, A. K.; Dorofeev, V. E.; Abagyan, R. A. *J. Comput. Phys.* **1991**, *92*, 261–272.
- Totrov, M.; Abagyan, R. *Proteins: Struct., Funct., Genet.* **1997**, 215–220.
- Totrov, M.; Abagyan, R. *Curr. Opin. Struct. Biol.* **2008**, *18*, 178–84.
- Sours, K. M.; Kwok, S. C.; Rachidi, T.; Lee, T.; Ring, A.; Hoofnagle, A. N.; Resing, K. A.; Ahn, N. G. *J. Mol. Biol.* **2008**, *379*, 1075–93.
- Akella, R.; Moon, T. M.; Goldsmith, E. J. *Biochim. Biophys. Acta* **2008**, *1784*, 48–55.
- Ghose, A. K.; Herbertz, T.; Pippin, D. A.; Salvino, J. M.; Mallamo, J. P. *J. Med. Chem.* **2008**, *51*, 5149–71.
- Liu, Y.; Gray, N. S. *Nat. Chem. Biol.* **2006**, *2*, 358–64.
- Wang, Z.; Canagarajah, B. J.; Boehm, J. C.; Kassisa, S.; Cobb, M. H.; Young, P. R.; Abdel-Meguid, S.; Adams, J. L.; Goldsmith, E. J. *Structure* **1998**, *6*, 1117–28.
- Shewchuk, L.; Hassell, A.; Wisely, B.; Rocque, W.; Holmes, W.; Veal, J.; Kuyper, L. F. *J. Med. Chem.* **2000**, *43*, 133–8.
- Fitzgerald, C. E.; Patel, S. B.; Becker, J. W.; Cameron, P. M.; Zaller, D.; Pikounis, V. B.; O’Keefe, S. J.; Scapin, G. *Nat. Struct. Biol.* **2003**, *10*, 764–9.
- Gill, A. L.; Frederickson, M.; Cleasby, A.; Woodhead, S. J.; Carr, M. G.; Woodhead, A. J.; Walker, M. T.; Congreve, M. S.; Devine, L. A.; Tisi, D.; O’Reilly, M.; Seavers, L. C.; Davis, D. J.; Curry, J.; Anthony, R.; Padova, A.; Murray, C. W.; Carr, R. A.; Jhoti, H. *J. Med. Chem.* **2005**, *48*, 414–26.
- Trejo, A.; Arzeno, H.; Browner, M.; Chanda, S.; Cheng, S.; Comer, D. D.; Dalrymple, S. A.; Dunten, P.; Lafargue, J.; Lovejoy, B.; Freire-Moar, J.; Lim, J.; McIntosh, J.; Miller, J.; Papp, E.; Reuter, D.; Roberts, R.; Sanpablo, F.; Saunders, J.; Song, K.; Villasenor, A.; Warren, S. D.; Welch, M.; Weller, P.; Whiteley, P. E.; Zeng, L.; Goldstein, D. M. *J. Med. Chem.* **2003**, *46*, 4702–13.
- Tamayo, N.; Liao, L.; Goldberg, M.; Powers, D.; Tudor, Y. Y.; Yu, V.; Wong, L. M.; Henkle, B.; Middleton, S.; Syed, R.; Harvey, T.; Jang, G.; Hungate, R.; Dominguez, C. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 2409–13.
- Pargellis, C.; Tong, L.; Churchill, L.; Cirillo, P. F.; Gilmore, T.; Graham, A. G.; Grob, P. M.; Hickey, E. R.; Moss, N.; Pav, S.; Regan, J. *Nat. Struct. Biol.* **2002**, *9*, 268–72.
- Erickson, J. A.; Jalaie, M.; Robertson, D. H.; Lewis, R. A.; Vieth, M. *J. Med. Chem.* **2004**, *47*, 45–55.
- Cavasotto, C. N.; Kovacs, J. A.; Abagyan, R. A. *J. Am. Chem. Soc.* **2005**, *127*, 9632–40.
- May, A.; Zacharias, M. *Proteins* **2008**, *70*, 794–809.
- Tama, F.; Wriggers, W.; Brooks, C. L., III. *J. Mol. Biol.* **2002**, *321*, 297–305.

- (26) Gohlke, H.; Kuhn, L. A.; Case, D. A. *Proteins* **2004**, *56*, 322–37.
- (27) Kornev, A. P.; Haste, N. M.; Taylor, S. S.; Eyck, L. F. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 17783–8.
- (28) Kornev, A. P.; Taylor, S. S.; Ten Eyck, L. F. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 14377–82.
- (29) Salaniwal, S.; Manas, E. S.; Alvarez, J. C.; Unwalla, R. J. *Proteins* **2007**, *66*, 422–35.
- (30) Gohlke, H.; Klebe, G. *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 2644–76.
- (31) Martin, M. W.; Newcomb, J.; Nunes, J. J.; McGowan, D. C.; Armistead, D. M.; Boucher, C.; Buchanan, J. L.; Buckner, W.; Chai, L.; Elbaum, D.; Epstein, L. F.; Faust, T.; Flynn, S.; Gallant, P.; Gore, A.; Gu, Y.; Hsieh, F.; Huang, X.; Lee, J. H.; Metz, D.; Middleton, S.; Mohn, D.; Morgenstern, K.; Morrison, M. J.; Novak, P. M.; Oliveira-dos-Santos, A.; Powers, D.; Rose, P.; Schneider, S.; Sell, S.; Tudor, Y.; Turci, S. M.; Welcher, A. A.; White, R. D.; Zack, D.; Zhao, H.; Zhu, L.; Zhu, X.; Ghiron, C.; Amouzegh, P.; Ermann, M.; Jenkins, J.; Johnston, D.; Napier, S.; Power, E. *J. Med. Chem.* **2006**, *49*, 4981–91.
- (32) Goldstein, D. M.; Alfredson, T.; Bertrand, J.; Browner, M. F.; Clifford, K.; Dalrymple, S. A.; Dunn, J.; Freire-Moar, J.; Harris, S.; Labadie, S. S.; La Fargue, J.; Lapierre, J. M.; Larrabee, S.; Li, F.; Papp, E.; McWeeney, D.; Ramesha, C.; Roberts, R.; Rotstein, D.; San Pablo, B.; Sjogren, E. B.; So, O. Y.; Talamas, F. X.; Tao, W.; Trejo, A.; Villaseñor, A.; Welch, M.; Welch, T.; Weller, P.; Whiteley, P. E.; Young, K.; Zipfel, S. *J. Med. Chem.* **2006**, *49*, 1562–75.
- (33) Dombroski, M. A.; Letavic, M. A.; McClure, K. F.; Barberia, J. T.; Carty, T. J.; Cortina, S. R.; Csiki, C.; Dipesa, A. J.; Elliott, N. C.; Gabel, C. A.; Jordan, C. K.; Labasi, J. M.; Martin, W. H.; Peese, K. M.; Stock, I. A.; Svensson, L.; Sweeney, F. J.; Yu, C. H. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 919–923.
- (34) McClure, K. F.; Letavic, M. A.; Kalgutkar, A. S.; Gabel, C. A.; Audoly, L.; Barberia, J. T.; Braganza, J. F.; Carter, D.; Carty, T. J.; Cortina, S. R.; Dombroski, M. A.; Donahue, K. M.; Elliott, N. C.; Gibbons, C. P.; Jordan, C. K.; Kuperman, A. V.; Labasi, J. M.; Laliberte, R. E.; McCoy, J. M.; Naiman, B. M.; Nelson, K. L.; Nguyen, H. T.; Peese, K. M.; Sweeney, F. J.; Taylor, T. J.; Trebino, C. E.; Abramov, Y. A.; Laird, E. R.; Volberg, W. A.; Zhou, J.; Bach, J.; Lombardo, F. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4339–44.
- (35) Honndorf, V. S.; Coudeville, N.; Laufer, S.; Becker, S.; Griesinger, C. *Angew. Chem., Int. Ed. Engl.* **2008**, *47*, 3548–51.
- (36) Vogtherr, M.; Saxena, K.; Hoelder, S.; Grimme, S.; Betz, M.; Schieborr, U.; Pescatore, B.; Robin, M.; Delarbre, L.; Langer, T.; Wendt, K. U.; Schwalbe, H. *Angew. Chem., Int. Ed. Engl.* **2006**, *45*, 993–7.
- (37) Huang, N.; Shoichet, B. K.; Irwin, J. J. *J. Med. Chem.* **2006**, *49*, 6789–801.
- (38) Verdonk, M. L.; Berdini, V.; Hartshorn, M. J.; Mooij, W. T.; Murray, C. W.; Taylor, R. D.; Watson, P. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 793–806.
- (39) Cavasotto, C. N.; Abagyan, R. A. *J. Mol. Biol.* **2004**, *337*, 209–25.
- (40) Subramanian, J.; Sharma, S. *ChemMedChem* **2008**, *3*, 336–44.
- (41) Rao, S.; Sanschagrin, P. C.; Greenwood, J. R.; Repasky, M. P.; Sherman, W.; Farid, R. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 621–7.
- (42) Lee, H. S.; Choi, J.; Kufareva, I.; Abagyan, R.; Filikov, A.; Yang, Y.; Yoon, S. *J. Chem. Inf. Model.* **2008**, *48*, 489–97.
- (43) Ferrara, P.; Gohlke, H.; Price, D. J.; Klebe, G.; Brooks, C. L., III. *J. Med. Chem.* **2004**, *47*, 3032–47.
- (44) Kolb, P.; Huang, D.; Dey, F.; Caffisch, A. *J. Med. Chem.* **2008**, *51*, 1179–1188.
- (45) Hoffmann, D.; Kramer, B.; Washio, T.; Steinmetzer, T.; Rarey, M.; Lengauer, T. *J. Med. Chem.* **1999**, *42*, 4422–33.
- (46) Yoon, S.; Welsh, W. J. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 88–96.
- (47) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
- (48) Momany, F. A.; Rone, R. *J. Comput. Chem.* **1992**, *13*, 888–900.
- (49) Roche, O.; Kiyama, R.; Brooks, C. L., III. *J. Med. Chem.* **2001**, *44*, 3592–8.
- (50) Hilal, S. H.; Karickhoff, S. W.; Carreira, L. A. *Quant. Struct.-Act. Relat.* **1995**, *14*, 348–355.
- (51) Fiser, A.; Sali, A. *Bioinformatics* **2003**, *19*, 2500–1.
- (52) Vieth, M.; Hirst, J. D.; Dominy, B. N.; Daigler, H.; Brooks, C. L. *J. Comput. Chem.* **1998**, *19*, 1623–1631.
- (53) Vieth, M.; Hirst, J. D.; Kolinski, A.; Brooks, C. L. *J. Comput. Chem.* **1998**, *19*, 1612–1622.
- (54) Wu, G. S.; Robertson, D. H.; Brooks, C. L.; Vieth, M. *J. Comput. Chem.* **2003**, *24*, 1549–1562.
- (55) Bursulaya, B. D.; Totrov, M.; Abagyan, R.; Brooks, C. L., III. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 755–63.
- (56) Lee, M. S.; Feig, M.; Salsbury, F. R., Jr.; Brooks, C. L., III. *J. Comput. Chem.* **2003**, *24*, 1348–56.
- (57) Feig, M.; Onufriev, A.; Lee, M. S.; Im, W.; Case, D. A.; Brooks, C. L., III. *J. Comput. Chem.* **2004**, *25*, 265–84.
- (58) Jain, A.; Vaidehi, N.; Rodriguez, G. *J. Comput. Phys.* **1993**, *106*, 258–268.
- (59) Chen, J.; Im, W.; Brooks, C. L., III. *J. Comput. Chem.* **2005**, *26*, 1565–78.
- (60) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. *Nucleic Acids Res.* **2007**, *35*, D198–201.
- (61) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605–12.