Journal of Medicinal Chemistry

Article

pubs.acs.org/jmc

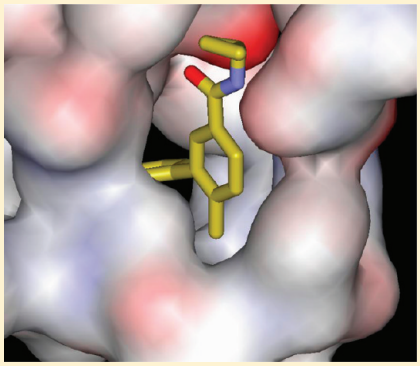
Methyl Effects on Protein–Ligand Binding

Cheryl S. Leung, Siegfried S. F. Leung, Julian Tirado-Rives, and William L. Jorgensen*

Department of Chemistry, Yale University, New Haven, Connecticut 06520, United States

Supporting Information

ABSTRACT: The effects of addition of a methyl group to a lead compound on biological activity are examined. A literature analysis of >2000 cases reveals that an activity boost of a factor of 10 or more is found with an 8% frequency, and a 100-fold boost is a 1 in 200 event. Four cases in the latter category are analyzed in depth to elucidate any unusual aspects of the protein–ligand binding, distribution of water molecules, and changes in conformational energetics. The analyses include Monte Carlo/free-energy perturbation (MC/FEP) calculations for methyl replacements in inhibitor series for p38α MAP kinase, ACK1, PTP1B, and thrombin. Methyl substitutions *ortho* to an aryl ring can be particularly effective at improving activity by inducing a propitious conformational change. The greatest improvements in activity arise from coupling the conformational gain with the burial of the methyl group in a hydrophobic region of the protein.



INTRODUCTION

The importance of methyl groups in modulating biological activity for small molecules is well documented.[1] Consistent with this, the most fundamental change in structure–activity studies is replacement of a hydrogen atom by a methyl group. The present study began by wondering what would be the maximum improvement in biological activity that could be produced by such a change and what would be the structural circumstances leading to it. One might envision a perfectly sized, hydrophobic pocket in a protein binding site ready to accept a methyl group. However, apart from electrostatics and steric complementary between the ligand and its receptor, conformational energetics, desolvation, and water placement in a binding site can also play important roles in the ligand binding process. Thus, we set out first to survey the literature for examples of the impact of methyl replacements on activity and then to examine the most beneficial cases with available crystallographic data. To this end, molecular simulations have been carried out to gain detailed insights and to ascertain if the computations could reproduce the observations. Among the many computational approaches that have been developed to study protein–ligand binding, free energy calculations in the context of Monte Carlo (MC) or molecular dynamics (MD) simulations are particularly powerful because they offer a rigorous way to compute binding affinities and relate thermodynamic quantities to molecular structures.[2−6] Specifically, free-energy perturbation (FEP) and thermodynamic integration (TI) methods with MC or MD sampling generally provide accurate predictions and have emerged as valuable in helping guide lead optimization.[7−14]

The typical view is that addition of a methyl group makes a molecule more hydrophobic and more prone to binding to biomolecules. The solvent-accessible surface area increases by ca. 30 Å^2, and this increase, or the corresponding increase in number of water neighbors, scales linearly with free energies of hydration, ΔG_{hyd}, for alkanes.[15] However, for lower alkanes owing to enthalpy–entropy compensation, there is little variation in ΔG_{hyd}, i.e., in the free energy of transfer from the gas phase to aqueous solution.[15a] A similar pattern is seen upon methylation of aromatic molecules, as demonstrated by the experimental ΔG_{hyd} values in Table 1.[16] For example, the first six entries show that $\Delta G_{hyd} = -0.85 \pm 0.05$ kcal/mol for benzene and mono-, di-, and trimethylated analogues. Thus, the gas/water partitioning of such molecules is normally affected little by addition of a methyl group or two. However, lipophilicity and the hydrophobic effect are taken to refer to transfer of a molecule from aqueous solution to a more lipid-like environment such as an alkane liquid, micelle, or interior of a protein. The expected effects of methylation now become apparent as in the experimental data in Table 1 for transfer from aqueous to hexadecane solution. Here, ΔG_{tran} equals the difference in free energies of solvation in hexadecane, ΔG_{C16}, and in water, ΔG_{hyd}.

The free energy of solvation in hexadecane becomes significantly more favorable with increasing methylation. In most cases, the benefit to ΔG_{tran} is about 0.7 kcal/mol per added methyl group. This can be taken as an estimate of the possible gain for transfer of a solvent-exposed methyl group on a ligand into a hydrophobic region of a protein. The situation for transfer to a hydrophobic pocket on a protein surface is more complex as it may involve displacement of bound water molecules.[17,18] Nevertheless, a similar estimate, 0.8 kcal/mol, for addition of an sp^3 carbon atom to the free energy of protein–ligand binding arose from an analysis of 200 crystal structures with associated inhibition data.[19] The increment came from partitioning the observed free energies of binding into functional group components.

Received: March 17, 2012
Published: April 13, 2012

Table 1. Solvent Accessible Surface Areas, and Experimental Free Energies of Solvation in Hexadecane and Water, and Free Energies of Transfer for Some Aromatic Molecules

solute	SASA[a]	ΔG_{C16}[b]	ΔG_{hyd}[b]	ΔG_{tran}[b]
benzene	266	−3.80	−0.86	−2.94
toluene	298	−4.53	−0.89	−3.64
o-xylene	322	−5.37	−0.90	−4.47
m-xylene	330	−5.23	−0.83	−4.40
p-xylene	331	−5.23	−0.80	−4.43
1,3,5-trimethylbenzene	363	−5.92	−0.90	−5.02
chlorobenzene	290	−4.98	−1.12	−3.86
2-chlorotoluene	318	−5.69	−1.15	−4.54
bromobenzene	294	−5.51	−1.46	−4.05
4-bromotoluene	327	−6.25	−1.39	−4.86
phenol	278	−5.13	−6.61	1.47
p-cresol	311	−5.88	−6.13	0.25
naphthalene	335	−7.03	−2.40	−4.63
1-methylnaphthalene	360	−7.89	−2.44	−5.45
1,3-dimethylnaphthalene	392	−8.50	−2.47	−6.03
pyridine	259	−4.12	−4.69	0.57
2-methylpyridine	292	−4.66	−4.63	−0.03
2,4-dimethylpyridine	325	−5.46	−4.87	−0.59
thiophene	248	−3.84	−1.42	−2.42
2-methylthiophene	283	−4.51	−1.38	−3.13

[a]In Å^2, computed with the OPLS-AA force field using the BOSS program. [b]In kcal/mol, from ref 16.

However, it is clear that the effect on biological activity or binding affinity of replacing a hydrogen atom by a methyl group in an inhibitor is highly context dependent. If there is no room in the binding site and the methyl group would project into the protein, the replacement can eliminate all affinity. The limit that is not obvious is what the maximal gain in affinity could be. Thus, a survey was made for all publications in the *Journal of Medicinal Chemistry* and *Bioorganic Medicinal Chemistry Letters* during 2006−2011 in order to establish the range of effects for single substitutions of a hydrogen atom by a methyl group. More than 2100 cases involving more than 100 different proteins were found with a reported methyl to hydrogen K_i or K_d ratio. The activity ratios were converted to free energy changes via $\Delta\Delta G$ = RT ln K_{Me}/K_H, and the resultant distribution is illustrated in Figure 1. Any methyl to hydrogen change was considered except at the end of an alkyl chain, e.g., propyl to ethyl or ethyl to methyl.

The analysis reveals a roughly Gaussian distribution with a mean of −0.1 kcal/mol, median of 0.0 kcal/mol, and standard deviation of 1.0 kcal/mol. The full range spans more than 10 kcal/mol. It is reasonable to assume that the reports in the journals favor compounds with greater activity so that, if random introductions of methyl groups were made, the resultant distribution would be much more skewed to the right. The key observations from the survey are that on average in reported SAR series introduction of a methyl group is just as likely to hurt as help activity and that it is extremely rare for addition of a methyl group to give a free energy gain greater than 3 kcal/mol; in fact, only 4 of the 2145 cases are in this category. Furthermore, even a 10-fold boost (1.36 kcal/mol) only occurs for 8% of the cases and a 100-fold gain (2.7 kcal/mol) is at the 0.4% level.

To elucidate the factors that can lead to large activity enhancements, four cases with at least 180-fold gains have been analyzed in detail. The selection of the four examples, which

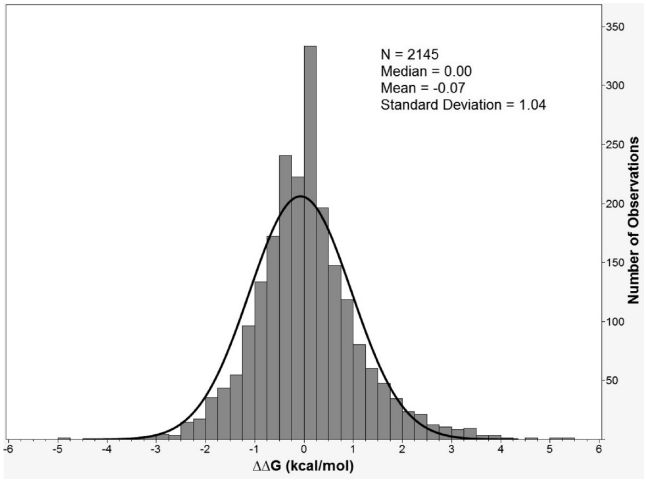


Figure 1. Distribution of free energy changes on activity for substitutions of a hydrogen atom by a methyl group in publications in the *Journal of Medicinal Chemistry* and *Bioorganic Medicinal Chemistry Letters* during 2006−2011.

target p38α MAP kinase,[20] ACK1,[21] PTP1B,[22] and thrombin,[23] was influenced by the availability of crystallographic data for protein−ligand complexes in each series. Three of the four examples feature the replacement of a hydrogen atom on an aromatic ring owing to its common occurrence in SAR studies. Monte Carlo free-energy perturbation (MC/FEP) calculations have been used to model the systems with complete hydration and to obtain computed changes in the free energy of binding for the methylations. In addition, the water placement algorithm JAWS was applied to assess further the proper location of water molecules for the complexes.[17,24,25] The influence of induced ligand strain upon binding was also evaluated by performing a conformational search and conformer focusing analysis[26] for each ligand using a GB/SA implicit solvent model.[27]

METHODS

System Setup. The crystal structures of p38α MAP kinase in complex with a biphenyl amide inhibitor (PDB ID: 3D7Z, 2.1 Å),[20] ACK1 with a pyrazolopyrimidine inhibitor (PDB ID: 3EQR, 2.0 Å),[21] PTP1B with a phenylthiophene inhibitor (PDB ID: 2QBP, 2.5 Å),[22] and thrombin with a pyrrolidinone inhibitor (PDB ID: 2JH0, 1.7 Å)[23] were chosen as starting points. Initial structures for the protein−ligand complexes and free ligands were constructed with the molecular growing program BOMB.[7] All protein residues within ca. 15 Å of any ligand atom were included. A few remote side chains were neutralized in order to maintain overall charge neutrality for each system. The ligand and the protein side chains within 10 Å of any ligand atoms were sampled during the MC simulations. The only constraints were the bond lengths in side chains, and all backbone atoms were frozen after a short conjugate-gradient minimization. The energetics for the systems were evaluated with the OPLS-AA force field for the protein and OPLS/CM1A for the ligand.[28,29] As usual, the CM1A atomic charges were scaled by 1.14 for neutral molecules.

Solvent Setup. The TIP4P water model was used.[30] The unbound ligands and the protein−ligand complexes were solvated in a 25 Å radius water cap with ca. 2000 and 1250 water molecules, respectively. A half-harmonic potential with a force constant of 1.5 kcal/mol Å^2 was applied to water molecules farther than 25 Å from the center of the ligand to prevent evaporation. The initial solvent coordinates were derived from the default hydration protocol[25] in the MCPRO 2.1 program.[31] Because water molecules might be trapped or absent from protein pockets and unable to exchange with the bulk in the course of a MC simulation, initial water distributions for the complexes were

dx.doi.org/10.1021/jm3003697 | *J. Med. Chem.* 2012, 55, 4489−4500

also obtained with the JAWS algorithm.[24] JAWS is typically used to locate expected hydration sites proximal to a binding site. First, possible water positions are found, then their occupancies are determined. In brief, a cubic grid with 1 Å spacing is positioned to cover the binding site. The grid region is defined by overlapping spheres of 5 Å radius centered on solute atoms. A key feature is the use of "θ water molecules" whose interactions with their surroundings are scaled between on ($\theta = 1$) and off ($\theta = 0$). MC simulations are initially performed to locate possible hydration sites by allowing θ water molecules to move on the grid with sampling of their θ values. Grid sites with high θ are clustered to yield the possible hydration sites. A subsequent MC simulation is run with the hydration sites occupied but still with θ sampling to determine the final occupancies. The locations of hydration sites were determined using 10 million (10 M) MC configurations that sample only θ and regular water molecules, followed by 10 M configurations that sample the protein, ligand, θ water molecules, and regular water molecules, and 40 M configurations to estimate the occupancies of the hydration sites.[24,25] When there was significant discrepancy between the initial water placement from the default procedure and JAWS, FEP calculations were performed starting from both arrangements.

MC/FEP Calculations. Differences in free energies of binding were determined from the usual thermodynamic cycle that requires conversion of one ligand to another both free in water and bound to the protein.[2−6] The requisite FEP calculations were executed with MC sampling using MCPRO and 11 windows of simple overlap sampling for each hydrogen/methyl interchange.[32,33] For the unbound ligand, each window consisted of 20 M configurations of equilibration and 30 M configurations for averaging. For the bound calculations, each window covered 10 M configurations of solvent-only equilibration, 20 M configurations of full equilibration, and 30 M configurations of averaging. All MC simulations were run at 298 K. A 9 Å residue-based cutoff was employed for the solvent−solvent, solute−solvent, and intrasolute nonbonded interactions.

Conformer Focusing Analysis. Free-energy changes for conformer focusing (ΔG_{cf}), which describes the energy cost required for an unbound ligand to achieve its bound conformation in solution, were also evaluated.[26] For simplification, it is assumed that the unbound ligand has multiple conformational states i with degeneracies n_i and that the bound ligand has only one conformational state k.[26] A conformational search was carried out on the ligands using the BOSS program[31] with the OPLS/CM1A force field[29] and GB/SA hydration.[27] The resultant conformer with the lowest root-mean-square deviation to the structure of the bound conformer was identified and compared with the lowest-energy conformer. The ΔG_{cf}

$$\Delta G_{cf} = (\varepsilon_k - \varepsilon_1) + k_B T \ln\left[\sum_i n_i \exp\left(\frac{-(\varepsilon_i - \varepsilon_1)}{k_B T}\right)\right] \quad (1)$$

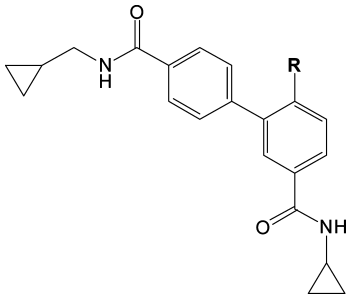
$$\varepsilon_i = E_i + \Delta G_{hyd,i} \quad (2)$$

was then calculated via eq 1 where ε_1 is the energy of the lowest-energy conformer, ε_k is the energy of the conformer closest in structure to the bound conformation, and the second term is the penalty for loss of conformational states. For binding in an aqueous environment, the energy of the unbound states consists of the intrinsic gas-phase energies (E_i) of the conformers and the free energy of hydration ($\Delta G_{hyd,i}$), as evaluated here by the GB/SA method (eq 2).[27]

RESULTS AND DISCUSSION

p38α MAP Kinase. In this case, the introduction of a methyl group in going from inhibitor **1** to **2** in Table 2 was found to improve the activity by at least 200-fold.[20] The corresponding FEP calculation was performed and yielded a change in free energy of binding, $\Delta\Delta G_{bind}$, of -3.3 ± 0.2 kcal/mol. This result is consistent with the difference of at least -3.1 kcal/mol that can be estimated from the K_i values, i.e., $\Delta\Delta G_{exp}$ = RT ln K_{Me}/K_H. Thus, it appears that the calculations are reproducing the experimental observation, so the modeling results can be considered in greater depth to try to understand the origin of the improved K_i.

Table 2. MC/FEP and Conformational Results for Biphenyl Amide Inhibitors of p38α MAP kinase



compd	R	K_i[a] (nM)	$\Delta\Delta G_{exp}$ (kcal/mol)	$\Delta\Delta G_{bind}$ (kcal/mol)	ΔG_{cf} (kcal/mol)
1	H	>2500	0.0	0.0	2.54
2	Me	12	<−3.16	−3.33 ± 0.17	1.91
3	F	460	<−1.00	ND	ND
4	Cl	25	<−2.77	ND	ND
5	OMe	520	<−0.93	ND	ND

[a]From ref 20. ND = not determined.

A typical configuration from the MC simulation for the complex of **2** is displayed in Figure 2. The interior of the binding site is lined by hydrophobic residues. The biphenyl fragment of **2** makes favorable contacts with Tyr35, Val38, Ile84, and Leu167. The amide carbonyl groups form hydrogen bonds with the NH of Met 109 and Asp168. One of the cyclopropyl groups is surrounded by Leu74, Phe169 (behind Leu171), and Leu171, while the other one is mostly exposed to the solvent. The added methyl group extends into a lipophilic pocket that is lined by the side chains of Val38, Ala51, Lys53, Leu104, and Thr106. The methyl group appears to have the optimal size to fill the cavity, as a loss in activity is observed when it is replaced by halogens or larger substituents including

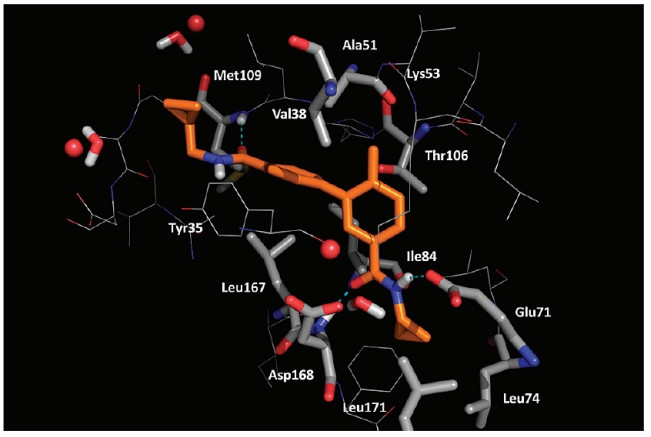


Figure 2. Representative structure from MC/FEP simulations of inhibitor **2** bound to p38α MAP kinase. The red spheres represent the positions of water molecules observed in the crystal structure. Water molecules in this computed image that are proximal to the crystallographic hydration sites are shown in stick rendering. The intermolecular hydrogen bonds are marked as blue dashed lines. For clarity, nonpolar hydrogen atoms and part of the binding site are omitted.

dx.doi.org/10.1021/jm3003697 | J. Med. Chem. 2012, 55, 4489−4500

a methoxy group (**3**–**5** in Table 2).[20] In the crystal structure, there are three water molecules observed within 5 Å of **2**: one water molecule bridges between the backbone carbonyl group of Tyr35 and the carboxylate group of Asp168, and the two other water molecules merge into the bulk solvent.[20] Examination of MC configurations revealed that water molecules are positioned proximal to these hydration sites throughout the simulations. Nevertheless, to further probe the water distribution in the binding pocket, the water placement algorithm JAWS was used on the bound complexes for **1** and **2**. The JAWS results for **2** were found to be in accord with the water positioning from the default procedure and with the crystallographic data.[20] The same water distributions are predicted for the complexes of **1** and **2**, indicating that addition of the methyl group does not displace a water molecule. Thus, it is clear that this case benefits from insertion of the methyl group into an empty hydrophobic pocket. However, the conformational effect of the methyl group also requires consideration. In principle, this effect is included in the computed $\Delta\Delta G_{bind}$; however, owing to limitations in the configurational sampling, the treatment may be incomplete. Consequently, an independent sense of its magnitude is of interest.

It is known that *ortho*-substitution in a biphenyl system shifts the interphenyl dihedral angle distribution from being centered on 30–40° to about 55° and that the torsional barriers are much increased.[34,35] During the course of the MC simulations, these dihedral angle distributions are computed. It should be noted that the simulations for the unbound inhibitors in water are initiated from the structure obtained by extracting the ligand from the complex. When there are significant torsional barriers, the ligand in water typically falls into the nearest conformational well and may stay there for the duration of the simulation. As illustrated in Figure 3, this occurred for both

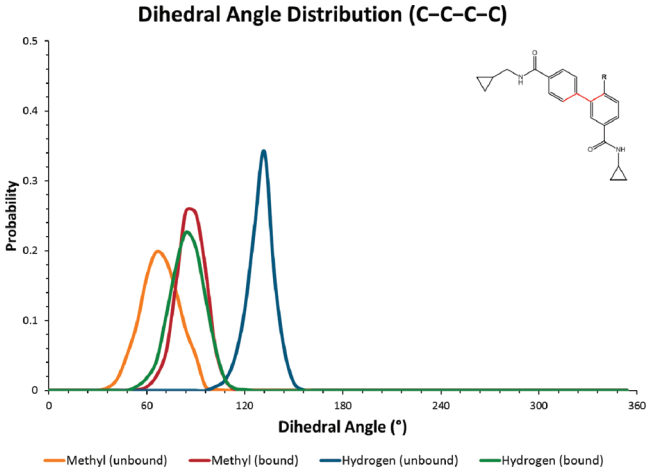


Figure 3. Computed distributions for the biphenyl C–C–C–C dihedral angle (highlighted in red) for the unbound and bound p38α MAPK inhibitors.

unbound **1** (blue curve) and **2** (yellow curve). The distribution for the hydrogen analogue **1** is centered at 130°, which is equivalent to 50°, and the distribution for the methyl analogue **2** is broader and centered at 65°. Thus, the expected shift to a more orthogonal arrangement of the rings does occur with addition of the methyl group. However, in both complexes, the centers of the distributions are further shifted to 85° (Figure 3), as reflected in the orthogonal conformation in Figure 2. The introduction of the *ortho* methyl group produces the added benefit of better preorganization of the unbound structure to be more like the bound one. The benefit should also be apparent and quantified through the conformer focusing analysis.

With eight rotatable bonds, over 270 unique conformers were found for each inhibitor in GB/SA water. Figure 4 shows the lowest-energy unbound conformer (conformer 1) and the conformer that is most similar to the bound one based on the rmsd of their coordinates. The preference for more orthogonal structures is apparent for **2**. From eq 1, the resultant ΔG_{cf} values are 2.5 and 1.9 kcal/mol for **1** and **2**, respectively. In viewing the structures, the difference of 0.6 kcal/mol is expected to underestimate the conformational benefit for the methyl group. Specifically, for **1**, the conformation most similar to the bound one still has a ca. 35° inter-ring dihedral angle, while 85° is needed for the bound structure. For biphenyl, the best estimate of the torsional barrier is 2.0 kcal/mol from the minimum at 40° to the peak at 90°,[36] while for 2-methylbiphenyl, the energy difference between the minimum at 56° and the peak at 90° is 0.86 kcal/mol using OPLS-AA. Thus, to twist from the minima to ca. 85°, it appears that the conformational benefit for binding derived from the methyl group in **2** is about 1 kcal/mol. The remaining 2–3 kcal/mol in the enhanced activity over **1** can be assigned to placement of the methyl group in the empty hydrophobic pocket between Val38 and Thr106 (Figure 2).

Activated Cdc42Hs-Associated Kinase 1. The SAR data for several pyrazolopyrimidine inhibitors of ACK1 are shown in Table 3.[21] Introduction of 2,6-dimethyl substitution in the anilinyl ring attached to the heterocycle lowers the K_i from >4 μM for the parent compound **6** to 3 nM for **10**; however, an additional methyl group at the 4-position is highly deactivating, yielding a K_i >2.5 μM for **11**. The present MC/FEP calculations were performed for **6**–**11** starting from the 3EQR crystal structure for the complex with **12**. Structures for the complexes with **10** and **11** are illustrated in Figure 5. The binding occurs in the ATP site with characteristic hydrogen bonds between the hinge-region Ala208 and the phenylaminopyrimidine fragment of the inhibitors and with the phenyl ring partly solvent-exposed. On the opposite end of the inhibitor, the 6-methyl group on the 3-anilinyl ring is pointing away from the viewer and the 2-methyl group is in front. Separate FEP results are obtained for 2- and 6-methyl substituents as they do not interconvert during the MC simulations.

The $\Delta\Delta G_{bind}$ results, which used the default hydration procedure, are again in reasonable accord with the experimental data in Table 3. Introduction of the 2-methyl group is computed to be particularly favorable as it projects into a hydrophobic region lined by the side chains of Ile190, Leu259 (not illustrated), and Phe271. Introduction of the 6-methyl group is computed to be less favorable; it does not enter a particularly hydrophobic area and it is in contact with the carbonyl group of Met203. Further methylation at the C4 position is computed to diminish the binding by 4.1 kcal/mol, which is consistent with the observed ca. 1000-fold loss in activity for **11** compared to **10**. The 4-methyl group projects into the sulfur atom of Met203, and it also interferes with the salt bridge between Lys158 and Glu177 (Figure 5b).

When inspecting individual MC configurations, no water molecule was observed for any of the complexes near a potentially significant crystallographic hydration site,[21] which is

dx.doi.org/10.1021/jm3003697 | J. Med. Chem. 2012, 55, 4489–4500

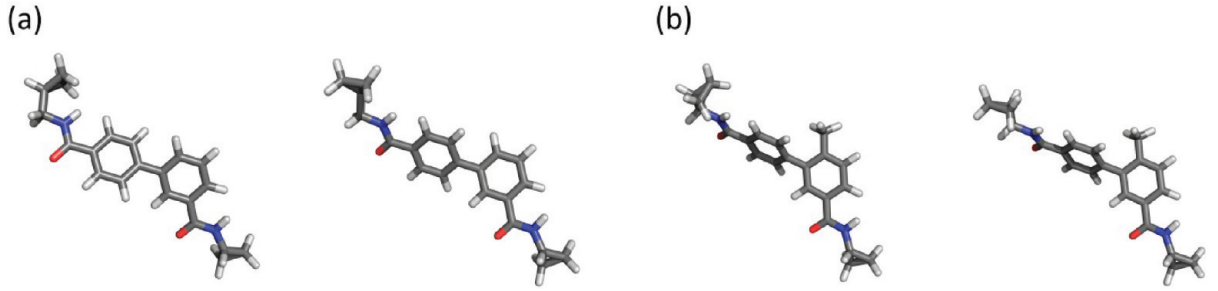
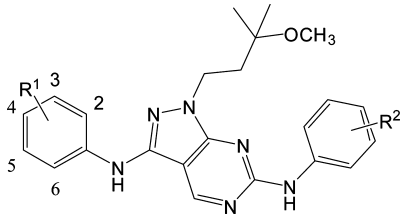


Figure 4. Conformers of (a) **1** and (b) **2**. The left structure is the lowest-energy one from the conformational search with GB/SA hydration and the right conformer is the one most similar to the bound structure. The *ortho* methyl substitution causes the two phenyl rings to be oriented in a more orthogonal manner.

Table 3. MC/FEP and Conformational Results (kcal/mol) for Pyrazolopyrimidine Inhibitors of ACK1



compd	R^1	R^2	K_i[a] (nM)	$\Delta\Delta G_{exp}$	$\Delta\Delta G_{bind}$	$\Delta\Delta G_{bind}$-JAWS	ΔG_{cf}
6	H	H	>4000	0.0	0.0	0.0	2.72
7	2-Me	H	ND	ND	−4.96 ± 0.18	ND	2.51
8	4-Me	H	ND	ND	−0.30 ± 0.17	ND	3.20
9	6-Me	H	ND	ND	−1.48 ± 0.17	ND	3.09
10	2,6-diMe	H	3	<−4.26	−5.66 ± 0.24	−5.50 ± 0.23	2.08
11	2,4,6-triMe	H	>2500	<0.28	−1.57 ± 0.33	−1.24 ± 0.29	1.28
12	2,6-diMe	4-piperazin-1-yl	2	<−4.50	ND	ND	ND

[a]From ref 21. ND = not determined.

indicated by the red sphere near Lys158 in Figure 5. Consequently, a JAWS analysis was applied for the complexes of **6**, **10**, and **11**. It was found that this hydration site should be occupied for all three complexes with a net binding free energy of −0.5 to −1.0 kcal/mol. The water molecule simultaneously forms hydrogen bonds with N2 of the pyrazole ring and the ammonium group of Lys158, and it has a π-type hydrogen bond with the proximal phenyl ring. The FEP calculations for the complexes of **6**, **10**, and **11** were then repeated starting from the JAWS water distribution. However, as shown in Table 3, this had an insignificant effect on the results as the water molecule is present in all initial and final states.

On the conformational side, it can be expected that the introduction of *ortho* methyl groups should again lead to preference for a more perpendicular disposition of the anilinyl phenyl ring and the heterocycle. Indeed, the distributions for the C3−N−C−C dihedral angle for the di-*ortho*-methylated cases are all computed to be centered at 60−80° both for the inhibitors unbound in water and in the complexes. However, for the parent compound **6**, the distribution is centered at 40° for the unbound inhibitor and at 70° in the complex. The situation for the mono-*ortho*-methylated cases is intermediate. Thus, the di-*ortho* methyl groups promote population of the orthogonal conformers, which are well preorganized for binding. The conformer focusing analysis reinforces this view. As shown in Figure 6, the conformational search found that the lowest-energy conformers for **6**−**9** in the GB/SA continuum model of water have the phenyl and heterocyclic rings almost coplanar. However, for **10** and **11**, there is little difference between the lowest-energy conformer and the one closest to the bound structure (Figure 6e,f). Consequently, the computed values of ΔG_{cf} from eq 1 are similar for **6**−**9** in Table 3, while the penalty is ca. 1−1.5 kcal/mol lower for **10** and **11**. The introduction of the two *ortho* methyl groups induces the ring to rotate to be perpendicular relative to the pyrazolopyrimidine ring. For **10**, the improved preorganization coupled with placement of one of the *ortho* methyl groups in the hydrophobic region near Ile190 and Phe271 leads to the more than thousand-fold enhancement of activity relative to **6**. As for ligand **11**, although it is well preorganized for binding, the added *para*-methyl group generates steric conflict, especially with Met203.

Protein Tyrosine Phosphatase 1B. The data in Table 4 for a phenylthiophene series bound to PTP1B show another case with great benefit derived from introduction of an *ortho* methyl group in a biaryl system. A methyl group at the 3-position of the thiophene ring enhances the activity 250-fold.[22] The starting point for the calculations was the 2QBP crystal structure for the bromo analogue **15**. MC/FEP results using the default hydration procedure well reproduce the benefit of the *ortho*-methylation with **14** computed to bind more favorably than **13** by 3.4 ± 0.3 kcal/mol. The two carboxylic acid groups are expected to be deprotonated in aqueous solution and were modeled as such.

The computed structure of **14** in complex with PTP1B is displayed in Figure 7. The binding site is shallow, and the top face of the inhibitor is mostly solvent-exposed. The thiophene ring is sandwiched between Tyr46 and Phe182, and the two

dx.doi.org/10.1021/jm3003697 | J. Med. Chem. 2012, 55, 4489−4500

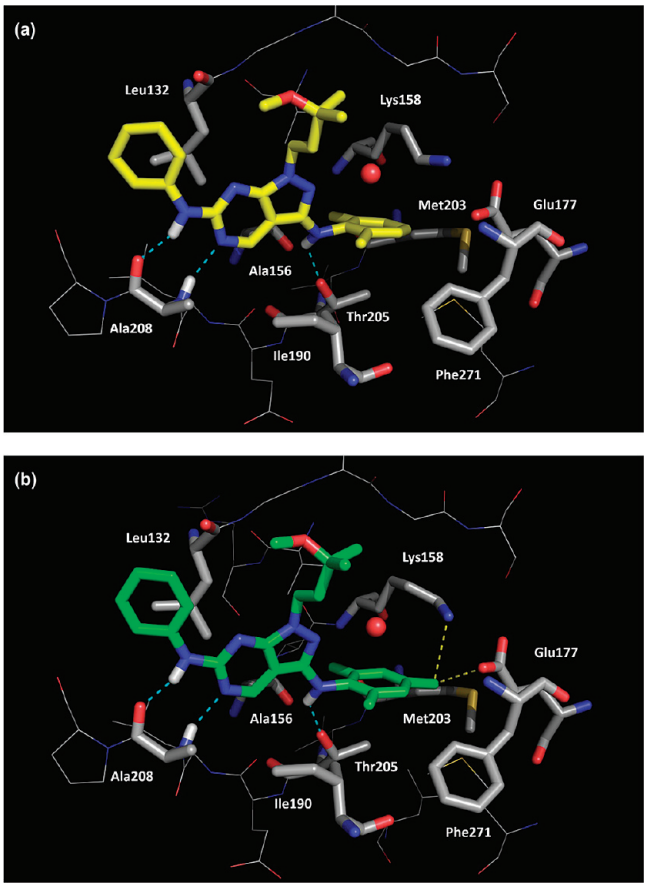


Figure 5. Representative configurations from the MC simulations for inhibitor **10** (a) and **11** (b) bound to ACK1. The red sphere indicates the position of a water molecule that is observed in the crystal structure of **12** complexed with ACK1. The 4-methyl group of **11** introduces steric clashes in the region of Met203, Lys158, and Glu177.

carboxylate groups extend toward the catalytic site and form hydrogen bonds with Lys120 and Arg221. The methyl group on the thiophene ring is inserted into a well-formed hydrophobic pocket lined by the side chain of Ile219 and C_β and C_γ of Gln262. Although the activity boost is well reproduced by the MC/FEP results, there is a discrepancy in the positioning of water molecules in the computations with the default procedure and hydration sites in the crystal structure for the complex of **15**.[22] There are four well-buried water molecules, indicated by red spheres in Figure 7b, in the crystal structure that are absent from the computed structure (Figure 7a). A hydration site of particular concern is below the thiophene ring, 3.1 Å from one of the carboxylate oxygens and 3.2 Å from the bromine of **15**.

Thus, JAWS analyses were performed for the complexes of **13** and **14**. Because the top-side of the binding site is solvent-exposed, the grid was positioned to cover the lower region between the ligand and protein. For the complex of **14**, the calculations predict one water molecule to be adjacent to the two carboxylates and four other water molecules to be proximal to the 4-amino piperidine and benzylsulfamoyl group (green spheres Figure 7b). The hydration site below the thiophene ring becomes disfavored when the methyl group is added to the thiophene ring; this hydration site has absolute binding free energies of −3.2 and +2.4 kcal/mol when **13** and **14** are bound, respectively, i.e., the water molecule is present in the crystal structure for the complex of the bromo analogue **15** and in the computed structure for the parent **13**, but it is computed to be absent for the methyl analogue **14**. The FEP calculations were then repeated using the solvent distribution generated from the JAWS analysis for **14**. As shown in Table 4, this had no impact on the computed $\Delta\Delta G_{bind}$.

In addition, some conformational benefit is expected for the methyl group in **14**. The C−C−C−C torsion angle for the biaryl junction in the crystal structure for **15** with PTP1B is 35°.[22] The computed distributions for the unbound **14** and the bound **13** and **14** are also centered in the 30−40° range; however, the MC simulations for the unbound **13** sampled in the −25° to 25° region. The same patterns are observed in crystal structures for planar five-membered heterocycles with a phenyl group at C2.[35] The unsubstituted examples prefer coplanar structures, while a substituent at C3 causes a ca. 35° twist. With the OPLS/CM1A force field, 2-phenylthiophene is computed to be coplanar and 3-methyl,2-phenylthiophene has a 30.4° biaryl dihedral angle. Consistently, the conformer focusing penalty in Table 4 for **14** is computed to be 0.6 kcal/mol less than that for **13**. The presence of the methyl group in **14** promotes the 30−40° twisted conformation that properly positions the methyl group in the hydrophobic pocket near Ile219. The absolute ΔG_{cf} values are relatively large at 4−5 kcal/mol owing in part to the large number of possible conformers for these inhibitors with 10 rotatable bonds; ca. 400 conformers were found for both **13** and **14**.

Thrombin. The final example considered in detail is for thrombin inhibitors **16** and **17** in Table 5. Unlike the cases discussed above, the methyl substitution is on an acyclic double bond. The experimental data show that addition of a 1-methyl group to the vinylsulfonamide linker lowers the K_i from 367 nM to 2 nM.[23] The MC/FEP calculations for the conversion of **16** to **17** gave a consistent result, favoring the binding of **17** by 3.7 ± 0.3 kcal/mol.

The computed structure of **17** bound to thrombin is shown in Figure 8. The sulfonyl group and the C4−C5 edge of the pyrrolidinone ring are solvent exposed, and the chlorothiophene fragment is well buried in a channel that terminates with Tyr228. The morpholine ring is also well-buried between the side chains of His57, Tyr60A, and Trp60D. The pyrrolidinone oxygen forms a hydrogen bond with the backbone NH of Gly216. The additional methyl group on the vinyl linker makes hydrophobic contacts with the hydrocarbon chain of Glu192; in the crystal structure, the methyl carbon atom is 3.80 and 4.70 Å away from the C_α and C_β of Glu192, respectively. The region cannot be described as a hydrophibic pocket. The image is more of the methyl group being inserted along the hydrophobic surface formed by the C_α and C_β of Glu192. Using the default hydration procedure in MCPRO, water molecules are found proximal to four of the seven crystallographic hydration sites near the ligand,[23] as shown in Figure 8. Most notably, the two relatively isolated water sites between Glu192 and the morpholine ring of **17** are not occupied in the MC simulations. A JAWS analysis was performed starting from the crystal structure, and the computed picture did not change; the hydration sites between Glu192 and the morpholine ring are not computed to be favorable. However, it should be noted that in the crystal structure, the carboxylate group of Glu192 is fully solvent exposed (extended toward the viewer in Figure 8), while in the MC simulations, the carboxylate group reorients to make closer contact with His57 and Ser195. There is also significant variation in the location of water molecules in crystal

dx.doi.org/10.1021/jm3003697 | J. Med. Chem. 2012, 55, 4489−4500

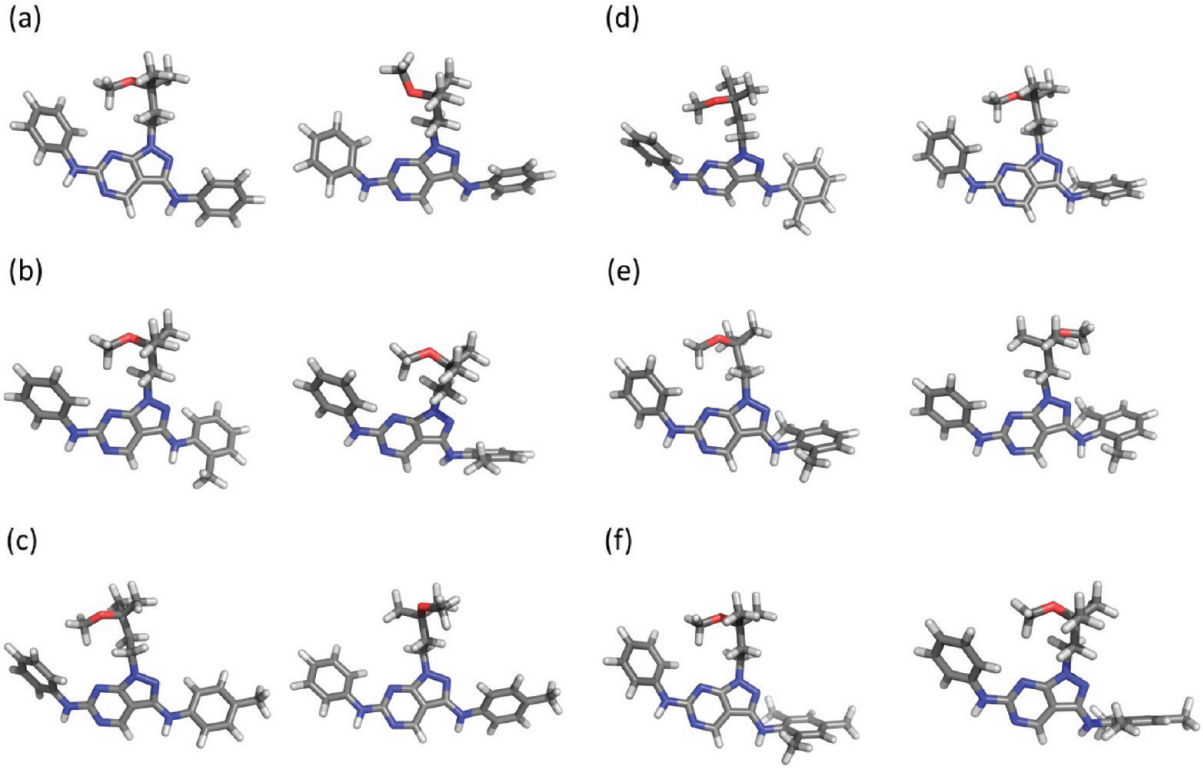
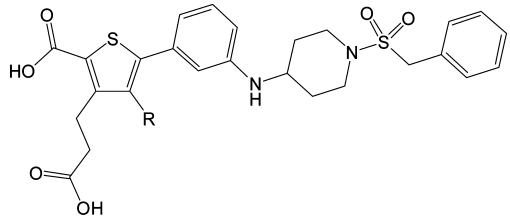


Figure 6. Conformers of (a) **6**, (b) the C2-Me analogue **7**, (c) the C4-Me analogue **8**, (d) the C6-Me analogue **9**, (e) the 2,6-diMe analogue **10**, and (f) the 2,4,6-triMe analogue **11**. The left structure is the lowest-energy one from the conformational search with GB/SA hydration, and the right conformer is the one most similar to the bound structure.

Table 4. MC/FEP and Conformational Results for Phenylthiophene Inhibitors of PTP1B



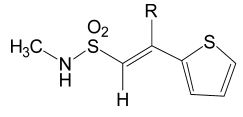
compd	R	K_i[a] (nM)	$\Delta\Delta G_{exp}$ (kcal/mol)	$\Delta\Delta G_{bind}$ (kcal/mol)	$\Delta\Delta G_{bind}$-JAWS (kcal/mol)	ΔG_{cf} (kcal/mol)
13	H	2000	0.0	0.0	0.0	4.81
14	Me	8	−3.27	−3.40 ± 0.26	−3.42 ± 0.26	4.23
15	Br	4	−3.68	ND	ND	ND

[a]From ref 22. ND = not determined.

structures for other analogues of **16**, with the number of hydrations sites ranging from two to seven.[23] Thus, there is some uncertainty in the hydration and positioning of the carboxylate group of Glu192; however, this is not expected to have impact on the effect of the added methyl group in **17** as its contacts with the C_α and C_β of Glu192 are invariant.

Results of B3LYP/6-31G(d) calculations have suggested that conformational effects for the vinylsulfonamide linker play an important role in the activity enhancement for **17** over **16**.[37] Model systems, *N*-methylethenesulfonamide and *N*-methyl-2-propene-1-sulfonamide, were considered in the gas phase. The additional methyl group was found to reduce the penalty to achieve the bound conformation for the vinylsulfonamide fragment by 1.7 kcal/mol. However, with a more complete model, the present force field results do not reveal the beneficial effect of the added methyl group unless the thiophene ring is included in the model.

Specifically, as illustrated in Figure 9, with OPLS/CM1A the lowest-energy conformer for **18** has a C═C−S−N angle of 43°, which is not far from the value of 57° for the conformation of **17** in the crystal structure for the complex with thrombin.



18, R = H
19, R = CH_3

However, the vinyl group and thiophene ring are coplanar, while in the crystal structure the S−C−C═C dihedral angle is 153°.[23] If the C═C−S−N dihedral angle is forced to 57°, the energy goes up by 1.0 kcal/mol, and if the S−C−C═C dihedral angle is then forced to 153°, the total strain is 2.7 kcal/mol. Now, for the methylated analogue **19**, the lowest-energy

dx.doi.org/10.1021/jm3003697 | J. Med. Chem. 2012, 55, 4489−4500

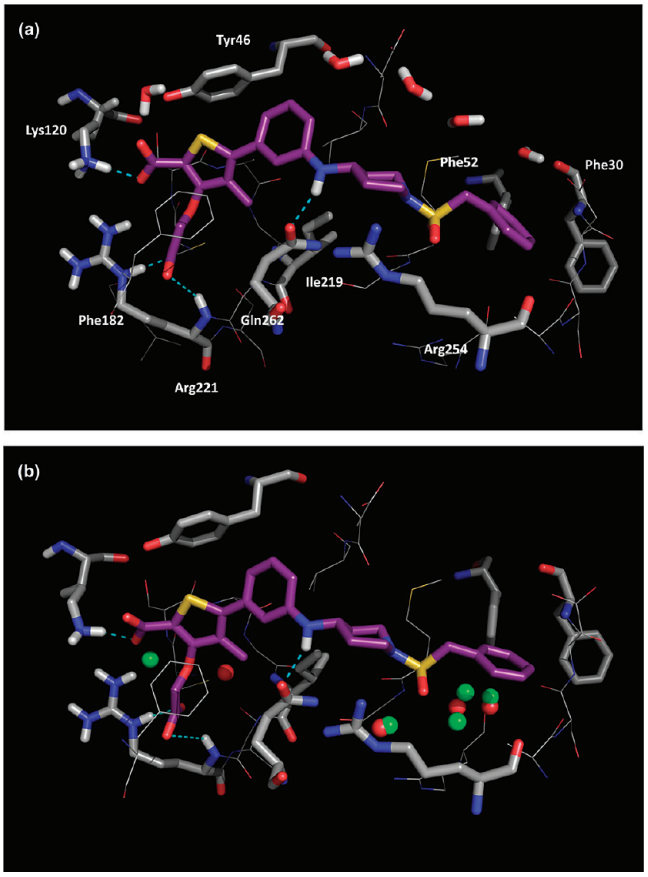
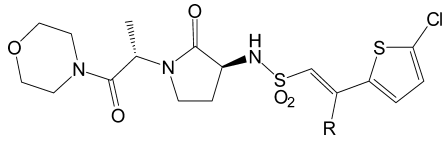


Figure 7. Representative structures for **14** bound to PTP1B from the MC/FEP calculations. (a) Some water molecules are illustrated in sticks that arose from the default hydration procedure. (b) The initial solvent distribution obtained using the water placement algorithm JAWS included water molecules positioned at the sites of the green spheres, which are proximal to hydration sites observed in the crystal structure for **15** (red spheres).

Table 5. MC/FEP and Conformational Results for Pyrrolidinone Inhibitors of Thrombin



compd	R	K_i[a] (nM)	$\Delta\Delta G_{exp}$ (kcal/mol)	$\Delta\Delta G_{bind}$ (kcal/mol)	ΔG_{cf} (kcal/mol)
16	H	367	0.0	0.0	7.39
17	Me	2	−3.09	−3.65 ± 0.33	6.09

[a]From ref 23.

conformer is significantly different, with a C=C−S−N dihedral angle of 183°, while the third conformer is very similar to the bound conformation, with dihedral angles of 60° (C=C−S−N) and 156° (S−C−C=C). The computed energy difference between the lowest-energy and third conformers is 1.44 kcal/mol. If the C=C−S−N dihedral angle for the lowest-energy conformer of **19** is changed to the crystallographic value of 57°, the energy goes up by 1.45 kcal/mol, basically yielding the third conformer. Although this is more than the 1.0 kcal/mol for **18**, changing the S−C−C=C dihedral angle for **19** to the crystallographic 153° only costs an additional 0.03 kcal/mol,

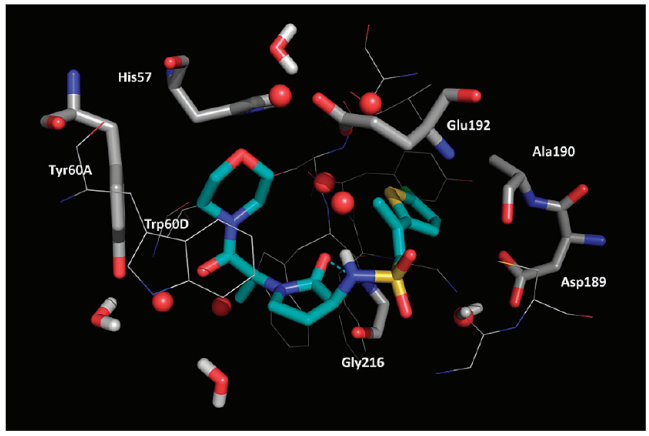


Figure 8. Representative configuration of **17** bound to thrombin from the MC/FEP calculations. Red spheres represent the positions of water molecules observed in the 2JH0 crystal structure in which **17** is also the ligand. Selected water molecules from the MC configuration are shown in stick rendering.

yielding a total strain of 1.48 kcal/mol. Thus, with these model calculations the induced strain to achieve the bound conformation is less for the methylated **19** than **18** by 2.7 − 1.5 = 1.2 kcal/mol in the gas phase, which becomes 0.5 kcal/mol when the GB/SA hydration is included.

Qualitatively, the same conclusion is reached as before;[37] however, here a key factor is that the added methyl group in **19** causes the vinyl and thiophene fragments to twist out-of-plane by ca. 30°. Similar to the PTP1B case, in the MC simulations for the complexes of both **16** and **17**, the distributions for the S−C−C=C dihedral angle cover about a 50° range and are centered at 140−150°, i.e., with a 30−40° out-of-plane twist. The same is true for the unbound **17**, while the unbound **16** samples on either side of coplanar geometries (S−C−C=C = 180°). Finally, when the full conformer focusing analysis was done for **16** and **17**, the advantage for the methylated compound turned out to be 1.3 kcal/mol (Table 5).

This case further illustrates the complexity of such conformational analyses that can arise from the choices of model system, computational methods, and inclusion of solvation.[27] In the end, it is likely that the ca. 3 kcal/mol enhancement of the activity for the methylated **17** over **16** is coming about evenly from the hydrophobic contacts for the methyl group and better conformational preorganization for binding.

Summary and Additional Cases. A pattern has emerged from these four cases that show particularly large activity gains for introduction of a methyl group. Namely, the added methyl group leads to both increased burial of hydrophobic surface area and to improved structural preorganization that diminishes the conformer focusing penalty. The conformational benefit arises from positioning the methyl group such that it causes out-of-plane torsion of a proximal ring to better match the preferred binding-site geometry. The placement is typically *ortho* or its equivalent on an aryl ring as in **2**, **10**, and **14** or at the 1-position of an attached group as in **17**. This analysis is supported by considering the 10 cases with the largest activity enhancements that were found in the present literature search. The cases are detailed in the Supporting Information and show K_H/K_{Me} ratios ranging from 63 to 3428. Seven of the 10 examples conform to the noted pattern. Among the cases that we have examined, the p38 kinase one stands out as showing

dx.doi.org/10.1021/jm3003697 | J. Med. Chem. 2012, 55, 4489−4500

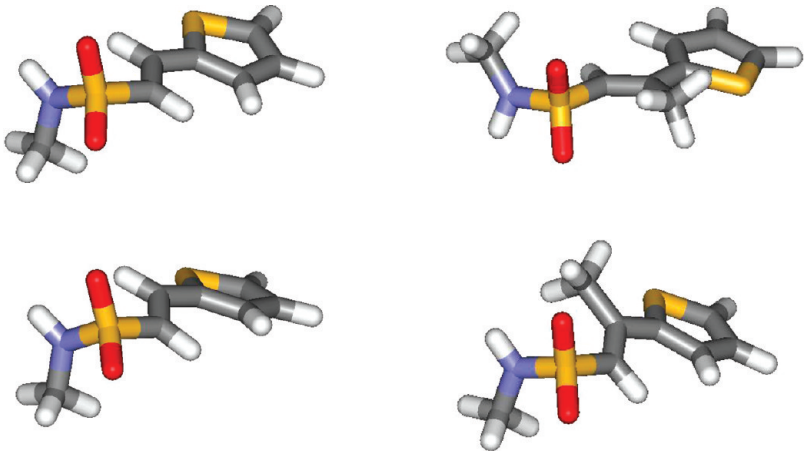


Figure 9. (left) The lowest-energy conformer of **18** (top) and the conformer distorted to the complex geometry (bottom). (right) The lowest-energy conformer of **19** (top) and the third lowest-energy conformer. Results using the OPLS/CM1A force field in the gas phase.

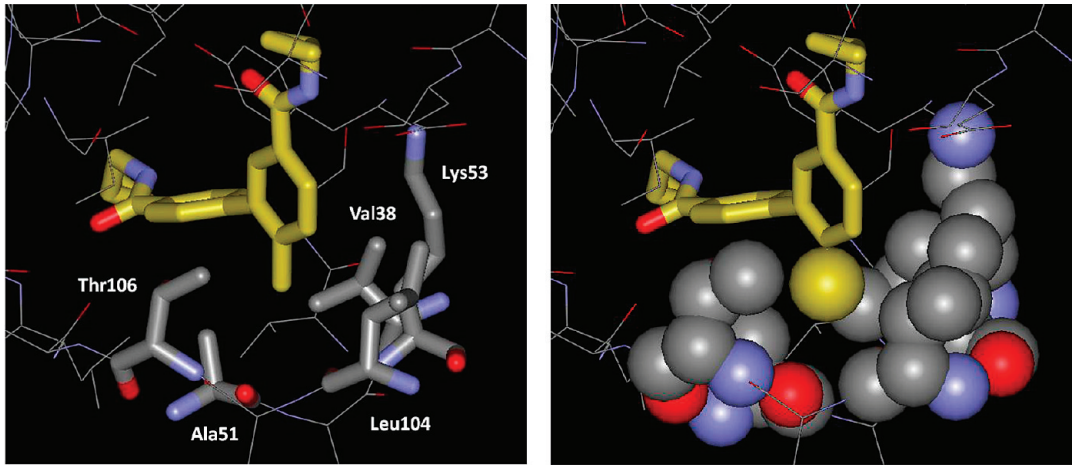
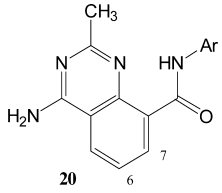


Figure 10. Renderings from the 3D7Z crystal structure[20] of the complex of p38α MAP kinase with **2**. The residues forming the hydrophobic pocket are highlighted. The space-filling rendering on the right includes the methyl group from **2** as a gold sphere.

the best formed hydrophobic pocket and the clear benefit of the torsional change induced by the added methyl group. This is further illustrated in Figure 10.

For balance, it is also worth noting an example in which an induced conformational change diminishes activity. In studies of **20** as inhibitors of phosphatidylinositol-3-kinase (PI3K),



20

introduction of a methyl group at C7 was found to be highly deactivating, e.g., by a factor of 631 in K_i values when Ar = 7-aza-indol-5-yl.[38] The 3PRZ crystal structure for a close analogue of **20** with hydrogen on C7 provides elucidation (Figure 11). The PI3K binding site is slot-like and suited for binding with planar inhibitors. The planarity of **20** is deliberately aided by inclusion of the intramolecular hydrogen bond with the secondary amide group.[38] A 7-methyl group disrupts the planarity by introducing a 20−80° twist to the N−C−C8−C9 dihedral angle according to various force field and quantum mechanical calculations. Unbound in water, out-of-plane torsion of the amide fragment of 7-methyl-**20** would also allow better hydration of the carbonyl group. Furthermore, C6 and C7 are fully solvent exposed in the 3PRZ crystal structure.

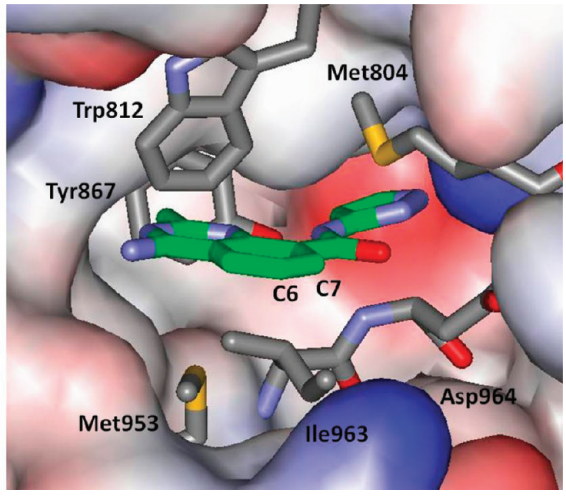


Figure 11. Rendering from the 3PRZ crystal structure[38] for a complex of **20** (Ar = 3-pyrazolyl) with PI3 kinase. The slot-like geometry of the binding site favors planar ligands.

dx.doi.org/10.1021/jm3003697 | J. Med. Chem. 2012, 55, 4489−4500

Consistently, methylation at C6 provides an example of no effect, e.g., the K_i values for the 4-pyrazolyl analogue of the inhibitor in Figure 11 and its 6-methyl derivative are 6 and 5 nM.[38] Similarly, for **20** with Ar = 2-methyl-pyridin-5-yl, the K_i values for the 6-H and 6-CH_3 analogues are both 0.5 nM.[38]

In our own experimental efforts to discover anti-HIV agents, activity data have been obtained for ca. 15 methyl/hydrogen pairs, although these are EC_{50} results from a cell-based assay using infected human T-cells. The examples for which there was the greatest improvement in activity are shown with structures **21**−**30**.[13,39,40] For **21**/**22** and **25**/**26**, there is roughly a factor of 10 gain with the methylated analogue, while for **23**/**24** and **27**/**29**, it is a factor of 5. These compounds are non-nucleoside inhibitors of HIV-1 reverse transcriptase, which bind in an allosteric site about 10 Å from the active site. The activity data are consistent with the computed structures for the complexes, e.g., Figure 12.[13,39,40] For **22**, the methyl group

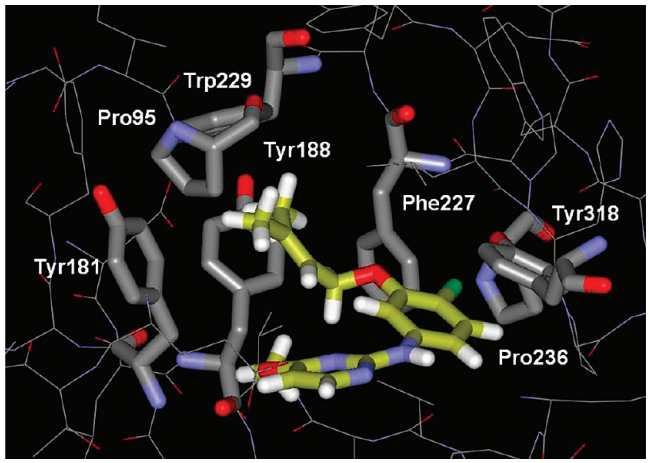
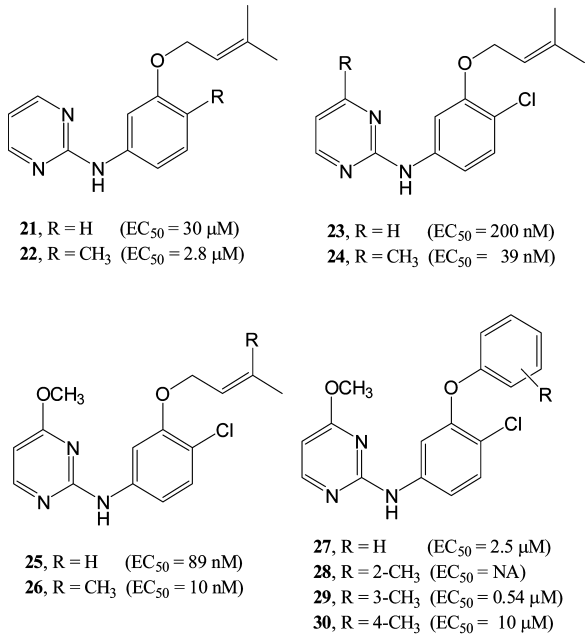


Figure 12. Computed structure of **26** bound to HIV-1 reverse transcriptase.

enters the hydrophobic channel between Tyr318 and Pro236. For **24**, the methyl group should be positioned in the hydrophobic region near the C_β atoms of Tyr181 and Tyr188. For **26**, the added methyl group is expected to help fill the hydrophobic space between Tyr188 and Trp229. The greater benefit from the methyl group in **22** likely reflects a conformational component from reduction of the conformational freedom of the adjacent dimethylallyoxy group. For **27**−**30**, the latter group is replaced by a phenoxy substituent. In this case, *ortho*- and *para*-methylation lead to less active compounds owing to steric clashes with Tyr188 and Trp229, while for the *meta*-analogue **29**, the added methyl group is well placed near the conflux of the side chains of Tyr181, Tyr188, and Trp229. Overall, our experiences are consistent with Figure 1; it is rare to find a factor of 10 improvement in activity for addition of a methyl group.

CONCLUSION

The present work was carried out to address the range and origins of effects of replacing a hydrogen atom by a methyl group in small molecules that regulate biological activity. Several key points emerged. A literature survey revealed that in reported SAR studies, a 10-fold or greater activity boost occurs 8% of the time, while 100-fold increases in activity are extremely rare, occurring with a 0.4% frequency. If one observes



31

an activity gain greater than this for a hydrogen to methyl change, further validation would be prudent. From the in-depth analyses of the four examples, it appears that to reach the 10-fold level, placement of a methyl group in a hydrophobic environment may be adequate; however, to go beyond that, the methyl group also generally needs to induce a propitious conformational change. This is typically achieved by *ortho* methylation in biaryl systems or by branching at an atom attached to a ring. Among the considered examples, the greatest reduction in the conformational penalty for binding was 1.3 kcal/mol for **17**. Thus, in lead optimization, exploration of such analogues should be pursued routinely, although there is no guarantee of a payoff as in the cases of 7-methyl-**20** and **28**.

On the technical side, the MC/FEP calculations did well in all cases in reproducing the observed activity changes for the replacement of hydrogen atoms by methyl groups. Although these are simple perturbations, it was found that initial placement of water molecules in simulations using explicit solvation needs continued careful consideration.[17,24,25,41] The estimation of the conformer focusing penalty (eq 1) could also be improved in view of its use of the approximate GB/SA method and a single conformation for the protein−ligand complex. Advances in these areas as well as for the force fields and sampling methods are still needed to provide enhanced reliability in the guidance for lead optimization from computer simulations.

ASSOCIATED CONTENT

Supporting Information

The raw data for Figure 1 are provided along with detailed information on the 10 cases showing the greatest benefit for activity upon introduction of a methyl group. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: 203-432-6278. Fax: 203-432-6299. E-mail: william.jorgensen@yale.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Gratitude is expressed to the National Institutes of Health (AI44616 and GM32136) and the National Foundation for Cancer Research for support.

■ ABBREVIATIONS USED

p38α MAP kinase, p38α mitogen-activated protein kinase; ACK1, activated Cdc42Hs-associated kinase 1; PI3K, phosphatidylinositol-3-kinase; PTP1B, protein tyrosine phosphatase 1B; OPLS, optimized potentials for liquid simulations; OPLS-AA, OPLS all-atom model; CM1A, charge model 1A; GB/SA, generalized Born−surface area hydration model

■ REFERENCES

(1) Barreiro, E. J.; Kümmerle, A. E.; Fraga, C. A. M. The Methylation Effect in Medicinal Chemistry. *Chem. Rev.* **2011**, *111*, 5215−5246.
(2) Jorgensen, W. L. Free Energy Calculations, a Breakthrough for Modeling Organic Chemistry in Solution. *Acc. Chem. Res.* **1989**, *22*, 184−189.
(3) Kollman, P. Free Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chem. Rev.* **1993**, *93*, 2395−2417.
(4) Lamb, M. L.; Jorgensen, W. L. Computational Approaches to Molecular Recognition. *Curr. Opin. Chem. Biol.* **1997**, *1*, 449−457.
(5) Simonson, T.; Georgios, A.; Karplus, M. Free Energy Simulations Come of Age: Protein−Ligand Recognition. *Acc. Chem. Res.* **2002**, *35*, 430−437.
(6) Michel, J; Verdonk, M. L.; Essex, J. W. Protein−Ligand Complexes: Computation of the Relative Free Energy of Different Scaffolds and Binding Modes. *J. Chem. Theory Comput.* **2007**, *3*, 1645−1655.
(7) Jorgensen, W. L. Efficient Drug Discovery and Optimization. *Acc. Chem. Res.* **2009**, *42*, 724−733.
(8) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. Alchemical free energy methods for drug discovery: progress and challenges. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150−160.
(9) Erion, M. D.; Dang, Q.; Reddy, M. R.; Kasibhatla, S. R.; Huang, J.; Lipscomb, W. N.; van Poelje, P. D. Structure-Guided Design of AMP Mimics That Inhibit Fructose-1,6-bisphosphatase with High Affinity and Specificity. *J. Am. Chem. Soc.* **2007**, *129*, 15480−15490.
(10) Steinbrecher, T.; Hrenn, A.; Dormann, K.; Merfort, I.; Labahn, A. Bornyl (3,4,5-trihydroxy)-cinnamate—an optimized human neutrophil elastase inhibitor designed by free energy calculations. *Bioorg. Med. Chem.* **2008**, *16*, 2385−2390.
(11) (a) Leung, S. S. F.; Tirado-Rives, J.; Jorgensen, W. L. Vancomycin Analogs: Seeking Improved Binding of D-Ala−D-Ala and D-Ala−D-Lac Peptides by Side-Chain and Backbone Modifications. *Bioorg. Med. Chem.* **2009**, *17*, 5874−5886. (b) Leung, S. S. F.; Tirado-Rives, J.; Jorgensen, W. L. Vancomycin Resistance: Modeling Backbone Variants with D-Ala−D-Ala and D-Ala−D-Lac Peptides. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 1236−1239.
(12) Leung, C. S.; Zeevaart, J. G.; Domaoal, R. A.; Bollini, M.; Thakur, V. V.; Spasov, K. A.; Anderson, K. S.; Jorgensen, W. L. Eastern Extension of Azoles as Non-Nucleoside Inhibitors of HIV-1 Reverse Transcriptase; Cyano Group Alternatives. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 2485−2488.
(13) Jorgensen, W. L.; Bollini, M.; Thakur, V. V.; Domaoal, R. A.; Spasov, K. A.; Anderson, K. A. Efficient Discovery of Potent Anti-HIV Agents Targeting the Tyr181Cys Variant of HIV Reverse Transcriptase. *J. Am. Chem. Soc.* **2011**, *133*, 15686−15696.
(14) Bollini, M.; Domaoal, R. A.; Thakur, V. V.; Gallardo-Macias, R.; Spasov, K. A.; Anderson, K. A.; Jorgensen, W. L. Computationally-Guided Optimization of a Docking Hit to Yield Catechol Diethers as Potent Anti-HIV Agents. *J. Med. Chem.* **2011**, *54*, 8582−8591.
(15) (a) Jorgensen, W. L.; Gao, J.; Ravimohan, C. Monte Carlo Simulations of Alkanes in Water: Hydration Numbers and the Hydrophobic Effect. *J. Phys. Chem.* **1985**, *89*, 3470−3473. (b) Li, I. T. S.; Walker, G. C. Signature of Hydrophobic Hydration in a Single Polymer. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 16527−16532.
(16) Abraham, M. H.; Whiting, G. S. Thermodynamics of Solute Transfer from Water to Hexadecane. *J. Chem. Soc., Perkin Trans. 2* **1990**, 291−300.
(17) Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. J. Energetics of Displacing Water Molecules from Protein Binding Sites: Consequences for Lead Optimization. *J. Am. Chem. Soc.* **2009**, *131*, 15403−15411.
(18) Bissantz, C.; Kuhn, B.; Stahl, M. A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem.* **2010**, *53*, 5061−5064.
(19) Andrews, P. R.; Craik, D. J.; Martin, J. L. Functional Group Contributions to Drug−Receptor Interactions. *J. Med. Chem.* **1984**, *27*, 1648−1657.
(20) Angell, R.; Aston, N. M.; Bamborough, P.; Buckton, J. B.; Cockerill, S.; deBoeck, S. J.; Edwards, C. D.; Holmes, D. S.; Jones, K. L.; Laine, D. I.; Patel, S.; Smee, P. A.; Smith, K. J.; Somers, D. O.; Walker, A. L. Biphenyl Amide p38α Kinase Inhibitors 3: Improvement of Cellular and in Vivo Activity. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4428−4432.
(21) Kopecky, D. J.; Hao, X.; Chen, Y.; Fu, J.; Jiao, X. Y.; Jaen, J. C.; Cardozo, M. G.; Liu, J.; Wang, Z.; Walker, N. P. C; Wesche, H.; Li, S.; Farrelly, E.; Xiao, S.-H.; Kayser, F. Identification and Optimization of N^3,N^6-diaryl-1*H*-pyrazolo[3,4-*d*] pyrimidine-3,6-diamines as a Novel Class of ACK1 Inhibitors. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 6352−6356.
(22) Wilson, D. P.; Wan, Z.; Xu, W.; Kinincich, S. J.; Follows, B. C.; Joseph-McCarthy, D.; Foreman, K.; Moretto, A.; Wu, J.; Zhu, M.; Binnun, E.; Zhang, Y.; Tam, M.; Erbe, D. V.; Tobin, J.; Xu, X.; Leung, L.; Shilling, A.; Tam, S. Y.; Mansour, T. S.; Lee, J. Structure-Based Optimization of Protein Tyrosine Phosphatase 1B Inhibitors: From the Active Site to the Second Phosphotyrosine Binding Site. *J. Med. Chem.* **2007**, *50*, 4681−4698.
(23) Young, R. J.; Brown, D.; Burns-Kurtis, C. L.; Chan, C.; Convery, M. A.; Hubbard, J. A.; Kelly, H. A.; Pateman, A. J.; Patikis, A.; Senger, S.; Shah, G. P.; Toomey, J. R.; Watson, N. S.; Zhou, P. Selective and Dual Action Orally Active Inhibitors of Thrombin and Factor Xa. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 2927−2930.
(24) Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. Prediction of the Water Content in Protein Binding Sites. *J. Phys. Chem. B* **2009**, *113*, 13337−13346.
(25) Luccarelli, J.; Michel., J.; Tirado-Rives, J.; Jorgensen, W. L. Effects of Water Placement on Predictions of Binding Affinities of p38α MAP Kinase. *J. Comp. Theory Comput.* **2010**, *6*, 3850−3856.
(26) Tirado-Rives, J.; Jorgensen, W. L. Contribution of Conformer Focusing on the Uncertainty in Predicting Free Energies for Protein−Ligand Binding. *J. Med. Chem.* **2006**, *49*, 5880−5884.
(27) Jorgensen, W. L.; Ulmschneider, J. P.; Tirado-Rives, J. Free Energies of Hydration from a Generalized Born Model and An All-Atom Force Field. *J. Phys. Chem. B* **2004**, *108*, 16264−16270.
(28) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *188*, 11225−11236.
(29) Jorgensen, W. L.; Tirado-Rives, J. Potential Energy Functions for Atomic-Level Simulations of Water and Organic and Biomolecular Systems. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6665−6670.
(30) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(31) Jorgensen, W. L.; Tirado-Rives, J. Molecular Modeling of Organic and Biomolecular Systems Using BOSS and MCPRO. *J. Comput. Chem.* **2005**, *26*, 1689−1700.
(32) Lu, N.; Kofke, D. A.; Woolf, T. B. Improving the Efficiency and Reliability of Free Energy Perturbation Calculations Using Overlap Sampling Methods. *J. Comput. Chem.* **2004**, *25*, 28−39.

(33) Jorgensen, W. L.; Thomas, L. T. Perspective on Free-Energy Perturbation Calculations for Chemical Equilibria. *J. Chem. Theory Comput.* **2008**, *4*, 869−876.
(34) Tamm, C. *Stereochemistry*; Elsevier: Amsterdam, 1982; pp 17−18.
(35) Brameld, K. A.; Kuhn, B.; Reuter, D. C.; Stahl, M. Small Molecule Conformational Preferences Derived from Crystal structure Data. A Medicinal Chemistry Focused Analysis. *J. Chem. Inf. Model.* **2008**, *48*, 1−24.
(36) Johansson, M. P.; Olsen, J. Torsional Barriers and Equilibrium Angle of Biphenyl: Reconciling Theory and Experiment. *J. Chem. Theory Comput.* **2008**, *4*, 1460−1471.
(37) Senger, S.; Chan, C.; Convery, M. A.; Hubbard, J. A.; Shah, G. P.; Watson, N. S.; Young, R. J. Sulfonamide-Related Conformational Effects and Their Importance in Structure-Based Design. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 2931−2934.
(38) Liu, K. K. C.; Huang, X.; Bagrodia, S.; Chen, J. H.; Greasley, S.; Cheng, H.; Sun, S.; Knighton, D.; Rodgers, C.; Rafidi, K.; Zou, A.; Xio, J.; Yan, S. Quinazolines with Intramolecular Hydrogen Bonding Scaffold (iMHBS) as PI3K/mTOR Dual Inhibitors. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 1270−1274.
(39) Ruiz-Caro, J.; Basavapathruni, A.; Kim, J. T.; Wang, L.; Bailey, C. M.; Anderson, K. S.; Hamilton, A. D.; Jorgensen, W. L. Optimization of Diarylamines as Non-Nucleoside Inhibitors of HIV-1 Reverse Transcriptase. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 668−671.
(40) Thakur, V. V.; Kim, J. T.; Hamilton, A. D.; Bailey, C. M.; Domaoal, R. A.; Wang, L.; Anderson, K. S.; Jorgensen, W. L. Optimization of Pyrimidinyl- and Triazinyl-amines as Non-Nucleoside Inhibitors of HIV-1 Reverse Transcriptase. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 5664−5667.
(41) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* **2008**, *130*, 2817−2831.