JOURNAL OF CHEMICAL INFORMATION AND MODELING

ARTICLE

pubs.acs.org/jcim

Identification of Novel Inhibitors of Tropomyosin-Related Kinase A through the Structure-Based Virtual Screening with Homology-Modeled Protein Structure

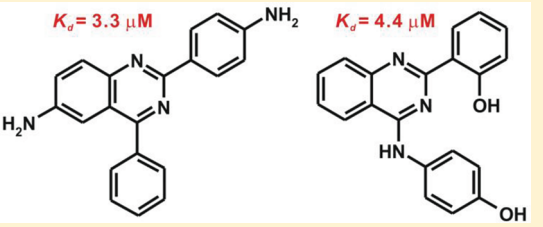
Hwangseo Park,*,† Okyung Chi,† Jinhee Kim,‡ and Sungwoo Hong*,‡

†Department of Bioscience and Biotechnology, Sejong University, 98 Kunja-Dong, Kwangjin-Ku, Seoul 143-747, Korea

‡Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 305-701, Korea

Supporting Information

ABSTRACT: Tropomyosin-related kinase A (TrkA) is a promising target for the development of cancer and pain therapeutics. Here, we report the first successful example of the use of a structure-based virtual screening to identify novel TrkA inhibitors. The accuracy of the virtual screening was improved by introducing an accurate solvation free energy term into the original AutoDock scoring function. We applied a drug design protocol involving homology modeling, docking analysis of a large chemical library, and enzyme inhibition assays to identify six structurally diverse TrkA inhibitors with K_d values ranging from 3 to 40 μM. The significant potencies and good physicochemical properties of these drug candidates strongly support their consideration in a development effort that would involve structure−activity relationship (SAR) studies to optimize the inhibitory activities. We also addressed the structural and energetic features associated with binding of the newly identified inhibitors in the ATP-binding site of TrkA. The results indicate that any structural modifications introduced for the purpose of enhancing the activity of TrkA inhibitors should maximize the attractive interactions within the ATP-binding site and simultaneously minimize the desolvation cost for complexation.



INTRODUCTION

Receptor tyrosine kinases (RTKs) are transmembrane receptors that phosphorylate tyrosine residues in certain specific protein substrates.[1] RTKs include the tropomyosin-related kinases (Trks), which serve as receptors for neurotrophin and play a critical role in the development and maintenance of the central and peripheral nervous systems. This small kinase subfamily includes three homologous isoforms, TrkA, TrkB, and TrkC, which are also known as NTRK1, NTRK2, and NTRK3, respectively.[2] Aside from maintaining and ensuring the survival of neuronal cells, experimental evidence has shown that Trks are involved in malignant transformation, chemotaxis, metastasis, and survival signaling in human cancers including prostate,[3] pancreatic,[4] colon,[5] papillary thyroid,[6] and lung cancers,[7] as well as in breast carcinoma,[8] acute myelogenous leukemia,[9] and neuroblastoma.[10]

Trks are also important mediators of pain. TrkA defects have been shown to prevent patients in various ethnic groups from being able to adequately perceive painful stimuli.[11] The efficacy of certain small-molecule TrkA inhibitors in the treatment of pain has been demonstrated in animal models.[12] TrkA has thus proven to be a promising target for the development of clinical treatments of both cancer and pain.

Structural investigations of TrkA have lagged behind the biochemical and pharmaceutical studies. The three-dimensional structure of TrkA has not yet been reported. A lack of structural information regarding the nature of TrkA-ligand interactions has made it difficult to design potent TrkA inhibitors. Nonetheless, a great deal of effort has been devoted to the discovery of TrkA inhibitors, as described in a recent comprehensive review,[13] with the aim of developing novel anticancer and antinociceptive medicines. These scientific endeavors have led to the identification of various ATP competitive inhibitors including 4-aminopyrazolylpyrimidine,[14] indenopyrrolocarbazole,[15,16] isothiazole,[17] oxindole,[18] and 2-aminothiazol[19] moieties as key structural elements. Most of these inhibitors have been discovered using high-throughput screening of chemical libraries or through structural modification of known inhibitor scaffolds. Thus, a rational design has not been applied so far to the discovery of TrkA inhibitors.

In the present study, we used a computer-aided drug design protocol involving homology modeling of TrkA, a structure-based virtual screening via docking simulations, and an in vitro enzyme assay to identify novel classes of TrkA inhibitors. Virtual screening has not always been successful due to the inaccuracy in the scoring function, which is responsible for a weak correlation of the predicted binding affinities with the experimental results.[20] The virtual screening approach used here is distinct from the others in the implementation of an accurate solvation free energy

Received: August 9, 2011
Published: October 21, 2011

term in the calculation of the binding free energy between TrkA and its putative inhibitor, which would have an effect of increasing the hit rate in the subsequent enzyme assay.[21] To the best of our knowledge, this report is the first example of the successful application of a structure-based virtual screening toward the identification of novel TrkA inhibitors. Docking simulations with the improved binding free energy function provide a valuable tool for predicting the relative potencies of known inhibitors and for enriching chemical libraries with molecules that are likely to have the desired biological activity.

MATERIALS AND METHODS

Homology Modeling of TrkA. Because the 3D structure of TrkA has not yet been reported, we carried out the homology modeling using the X-ray crystal structure of insulin-like growth factor 1 receptor (IGF1R, PDB entry: 3I81).[22] IGF1R provided a template for building a high-quality model structure of TrkA suitable for the virtual screening with docking simulation. The ClustalW program[23] was used to align the sequences of the TrkA and IGF1R kinase domains using the BLOSUM matrices to score the alignments. The parameters GAP OPEN, GAP EXTENTION, and GAP DISTANCE were set to 10, 0.05, and 8, respectively. The opening and extension gap penalties were varied systematically, and the alignment was inspected for violations of the structural integrity within the structurally conserved regions. The best scored sequence alignment was then used to construct a 3D model of the TrkA kinase domain using the latest version of the MODELLER program.[24] During the model building process, we employed an optimization method involving conjugate gradients and molecular dynamics to minimize violations of the spatial restraints. The coordinates for the gap region structures were optimized from a randomized distorted structure that bridged the two anchoring regions, as implemented in the MODELLER program. The loops were modeled using an enumeration algorithm to increase the accuracy of the calculated structures.[25] We then calculated the conformational energies of the predicted TrkA structures using the ProSa 2003 program for evaluation.[26]

Virtual Screening of TrkA Inhibitors. We used the AutoDock program[27] in the structure-based virtual screening of TrkA inhibitors because the outperformance of its scoring function over those of the other ones had been shown in several target proteins.[28] The atomic coordinates of TrkA obtained from the homology modeling were used as the receptor model in the virtual screening with docking simulations. Special attention was paid to the assignment of the protonation states of the ionizable Asp, Glu, His, and Lys residues in the homology-modeled structure of TrkA. The Asp and Glu side chains were assumed to be neutral if one of their carboxylate oxygens pointed toward a hydrogen-bond accepting group including the backbone aminocarbonyl oxygen at a distance within 3.5 Å, which is generally accepted as the distance limit for a hydrogen bond of moderate strength.[29] Similarly, the lysine side chains were assumed to be protonated unless an NZ atom was proximal to a hydrogen-bond donating group. The same procedure was applied toward determining the protonation states of the ND and NE atoms in the His residues. After determining the protonation state of each protein atom, we carried out 200 cycles of energy minimization with the AMBER program to remove the bad steric contacts.

A chemical library of organic molecules to be screened for inhibitory activity was prepared from the chemical database distributed by InterBioScreen Ltd. (http://www.ibscreen.com), which comprises about 460,000 synthetic and natural organic compounds. Prior to conducting the virtual screening with docking simulations, the library was filtered according to Lipinski's "Rule of Five" to adopt only the compounds with the physicochemical properties of potential drug candidates[30] and without reactive functional group(s). To remove the structural redundancies in the chemical library, structurally similar compounds with a Tanimoto coefficient exceeding 0.85 were clustered into a single representative molecule. Molecular similarities were measured using the fingerprints of each molecule, generated using the Daylight software as an ASCII string of 1's and 0s. In this way, a docking library consisting of ~240,000 compounds was constructed. The CORINA program was used to obtain the 3D atomic coordinates of the prefiltered compounds, followed by the assignment of atomic charges according to the Gasteiger−Marsilli method[31] to be consistent with those of TrkA. Docking simulation of each compound in the docking library was then performed to calculate the binding mode and binding free energy in the ATP binding site of TrkA.

The docking simulation of each compound was conducted using an improved empirical AutoDock scoring function, in which a new solvation model for organic molecules was introduced. This modified scoring function can be expressed as follows

$$\Delta G_{bind}^{aq} = W_{vdW}\sum_{i=1}\sum_{j=1}\left(\frac{A_{ij}}{r_{ij}^{12}}-\frac{B_{ij}}{r_{ij}^{6}}\right) + W_{hbond}\sum_{i=1}\sum_{j=1}E(t)\left(\frac{C_{ij}}{r_{ij}^{12}}-\frac{D_{ij}}{r_{ij}^{10}}\right) + W_{elec}\sum_{i=1}\sum_{j=1}\frac{q_i q_j}{\varepsilon(r_{ij})r_{ij}} + W_{tor}N_{tor} + W_{sol}\sum_{i=1}S_i(Occ_i^{\max}-\sum_{j>i}V_j e^{-r_{ij}^2/2\sigma^2}) \quad (1)$$

Here, W_{vdW}, W_{hbond}, W_{elec}, W_{tor}, and W_{sol} are the weighting factors of van der Waals, hydrogen bond, electrostatic interactions, torsional term, and desolvation energy of the putative inhibitors, respectively. r_{ij} represents the interatomic distance, and A_{ij}, B_{ij}, C_{ij}, and D_{ij} are related to the depth of the potential energy well and the equilibrium separation between the protein and ligand atoms. The hydrogen-bond term includes an additional weighting factor, $E(t)$, representing the angle-dependent directionality. A cubic equation approach was employed to obtain the dielectric constant, $\varepsilon(r_{ij})$, which was required to compute the interatomic electrostatic interactions between TrkA and a ligand molecule.[32] In the entropic term, N_{tor} is the number of rotatable bonds in a ligand. In the desolvation term, S_i and V_i are the solvation parameter and the fragmental volume of atom i,[33] respectively, whereas Occ_i^{max} indicates the maximum atomic occupancy. The molecular solvation free energy term in eq 1 included the atomic parameters developed by Kang et al.[34] because those of the atoms other than carbon were unavailable in the current version of AutoDock. This modification of the solvation free energy term is expected to increase the accuracy of the virtual screening because the underestimation of the ligand solvation energy often leads to the overestimation of the binding affinities of ligands with many polar groups.[21]

The docking simulation of a compound in the docking library started with the calculation of 3D grids of interaction energy for

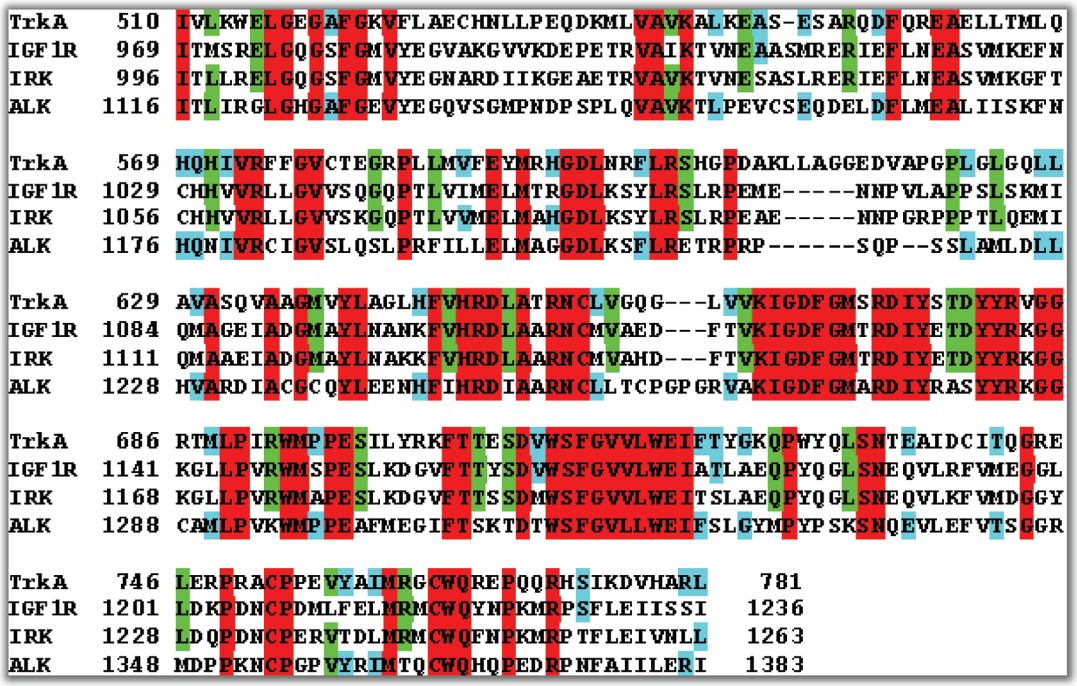


Figure 1. Multiple sequence alignment of the TrkA kinase domain with three structural homologues.

all possible atom types to cope with the whole chemical library. These uniquely defined potential grids for TrkA were then used in common for docking simulations of all compounds under consideration for virtual screening. For the center of the common grids, we used the center of mass coordinates of the IGF1R inhibitor[22] whose binding mode in the ATP binding site of TrkA had also been predicted in the homology modeling together with the structure of TrkA itself. The calculated grid maps were of dimension 61 × 61 × 61 points with the spacing of 0.375 Å, yielding a receptor model that includes the atoms within 22.9 Å of the grid center. These grid maps are actually sufficient to cope with the entire part of the kinase domain of TrkA. For each compound in the docking library, 10 docking runs were performed with the initial population of 50 individuals. Maximum number of generations and energy evaluation were set to 27,000 and 2.5 × 10^5, respectively. Docking simulations with the improved AutoDock scoring function were then carried out in the ATP-binding site of TrkA to score and rank the compounds in the docking library according to the binding affinity for TrkA.

RESULTS AND DISCUSSION

Homology Modeling of the TrkA Catalytic Domain. The peptide sequence of human TrkA comprising 796 amino acid residues was retrieved from the UniProtKB protein knowledgebase (http://www.uniprot.org, accession number: P04629). Domain analysis of the full amino acid sequence indicated that residues 510−781 corresponded to the cytoplasmic kinase domain, and the remaining N-terminal and central regions constituted the extracellular receptor and transmembrane domains, respectively. Although the 3D structure of the TrkA kinase domain has not yet been described, we found three structural homologues of TrkA through the database search at NCBI BLAST (http://blast.ncbi.nlm.nih.gov): IGF1R,[22] insulin receptor kinase (IRK),[35] and anaplastic lymphoma kinase (ALK).[36] These three candidates for the structural template are also homologous to each other with the sequence identities exceeding 45% and structural root-mean-square deviations (rmsd's) of backbone atoms falling within 1.0 Å. As shown in Figure 1, the sequence identity between TrkA and the three templates amounts to 40−42%. These high sequence identities indicate that the high-quality structures of TrkA can be generated because the rmsd between the experimental and the homology-modeled structures has been known to fall within 1.5 Å when the sequence identity between the template and the target ranges 40−50%.[37] Therefore, we constructed the three structural models for the kinase domain of TrkA through the homology modeling with the three structural templates.

Three structural models of the TrkA kinase domain obtained from homology modeling were evaluated using the ProSa 2003 program[26] by examining whether the interactions between each residue and the rest of the protein structure could be maintained favorable. This program calculates the knowledge-based mean fields to judge the quality of a protein fold, and it has been widely used to estimate the stability of a protein conformation. Compared in Figure 2 are the ProSa 2003 energy profiles of the three homology-modeled TrkA structures obtained using the three structural templates: IGF1R, IRK, and ALK. We note that although the ProSa energy remains negative in a large part of protein, it becomes positive at residues 150−200 in the homology-modeled structures of TrkA with 3D structures of IRK and ALK. On the other hand, the ProSa energies appear to be maintained negative for most amino acid residues in the structural model obtained with 3D structure of IGF1R. On the basis of these energetic features, the homology-modeled structure of the kinase domain of TrkA using the X-ray crystal structure of IGF1R as the template was selected to be used in the subsequent docking simulations for virtual screening of TrkA inhibitors.

Virtual Screening and in Vitro Enzyme Assay. Of the 240,000 compounds subject to the virtual screening with the

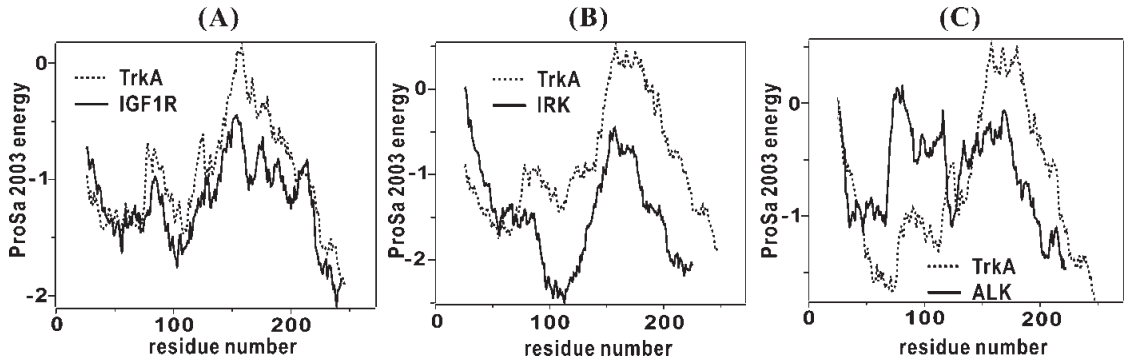


Figure 2. Comparative view of the ProSa energy profiles of the homology-modeled TrkA kinase domain and those of (A) IGF1R, (B) IRK, and (C) ALK. For convenience, the amino acids are renumbered from 1 instead of retaining the original numbers.

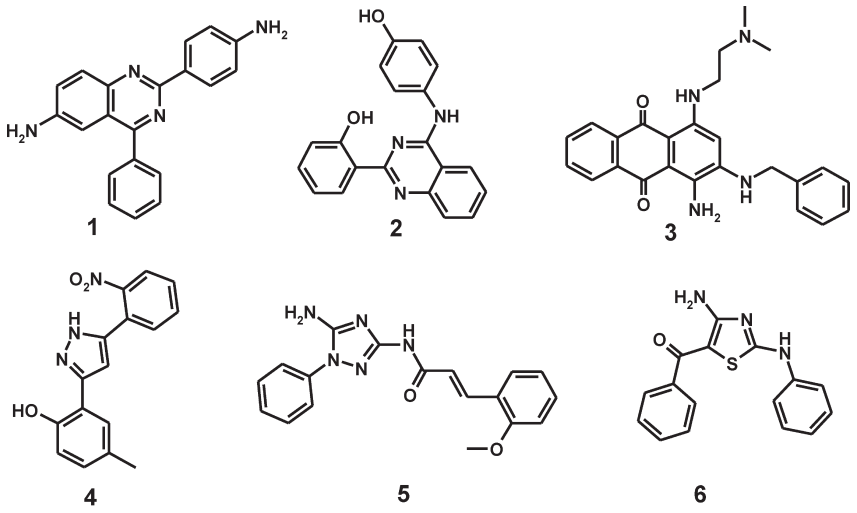


Figure 3. Chemical structures of the newly identified TrkA inhibitors.

Table 1. Inhibitory Activity, Calculated Binding Free Energy in the Gas Phase (ΔG_b^{gas}), Solvation Free Energy (ΔG^{sol}), and Binding Free Energy in Solution (ΔG_b^{aq}) for the TrkA Inhibitors[a]

compound	ΔG_b^{gas}	ΔG^{sol}	ΔG_b^{aq}	POC (%)	K_d (μM)	ΔG_b^{exp}
1	−20.82	−8.70	−12.12	0.55	3.3	−7.45
2	−22.94	−10.08	−12.86	0.15	4.4	−7.28
3	−27.34	−17.22	−10.12	16	35	−6.06
4	−24.09	−14.19	−9.90	16	40	−5.98
5	−23.12	−13.85	−9.27	19	37	−6.03
6	−19.31	−8.40	−10.91	20	40	−5.98

[a] The ΔG_b^{aq} values indicate the final binding free energies for the best binding configurations. The experimental ΔG_b (ΔG_b^{exp}) values calculated from the corresponding K_d values are also given. All energy values are given in kcal/mol.

docking simulations, 200 top-scored compounds were selected as virtual hits. One hundred eighty-nine of them were available from the compound supplier and were tested over TrkA at the concentration of 50 μM in a high-throughput binding assay (KINOMEscan, Ambit Biosciences).[38] Eighteen compounds were discovered as hits with a percent of control (POC) value ≤50, and six of these with the POC value of ≤20 were selected for the determination of K_d values. These full K_d determinations for the six compounds were conducted to validate the initial hits selected from POC measurements through the competition binding assay at the points of 11 inhibitor concentrations prepared with 3-fold serial dilutions. K_d values were then calculated with the Hill equation using the standard dose−response curves for the six initial hits. The K_d value was measured in duplicate for each inhibitor, and the average values of two were reported. The chemical structures and inhibitory activities of the newly identified inhibitors are shown in Figure 3 and Table 1, respectively. Weak TrkA inhibitors with POC values exceeding 20 are shown in the Supporting Information. We note that the most potent inhibitors **1** and **2** possess the quinazoline group in common in the middle of molecular structure. This two-membered ring seems to serve as an effective mimic for the adenine group of ATP because **1** and **2** are found to be potent inhibitors of TrkA with the associated K_d values of 3.3 and 4.4 μM, respectively. Compounds **3**−**6** are structurally diverse and have similar inhibitory activity with K_d values ranging from 35 to 40 μM. To the best of our knowledge, none of these six compounds has been reported as TrkA inhibitors so far. In addition, no additional biological activity was found for the six inhibitors at least in the two most popular chemical databases, ChEMBL and PubChem. Because these inhibitors are structurally diverse and their physicochemical properties have proved to satisfy the "Rule of Five"

for a drug candidate, they deserve consideration for further development by structure−activity relationship (SAR) studies to optimize the inhibitory activities to develop new anticancer and antinociceptive drugs.

We now discuss the energetic features associated with binding of the newly identified inhibitors to the ATP binding site of TrkA. Table 1 lists the calculated binding free energies and solvation free energies of the six identified inhibitors. Actually, these inhibitors were selected through the virtual screening with the scoring function in which the solvation free energy of a putative inhibitor had been implemented as the fifth term in eq 1. Keeping it in mind that the binding free energy of a protein−ligand complex in aqueous solution (ΔG_b^{aq}) can be approximated as the difference between that in the gas phase (ΔG_b^{gas}) and the solvation free energy of the ligand (ΔG^{sol}),[21] we computed the two energy components separately to estimate their relative contributions to ΔG_b^{aq}. When the two energy components for the six inhibitors are compared, it follows immediately that ΔG^{sol} values correspond to 40−60% of ΔG_b^{gas} and can even be larger in magnitude than ΔG_b^{aq}. This result confirms that the solvation term should be one of the most significant energy components in the scoring function.

The predicted binding free energies calculated for **1**−**6** agreed better with the relative inhibitory activities when the scoring function lacking the ligand solvation term (ΔG_b^{gas}) was modified by subtracting ΔG^{sol}. This yielded a physically more reasonable free energy (ΔG_b^{aq}). For example, the values of ΔG_b^{gas} for the most potent inhibitors (**1** and **2**) were less favorable than those of the weaker inhibitors (**3**−**5**). This indicates that the former would have a lower binding affinity for TrkA than the latter, which is in contrast to the experimental results for their inhibitory activities. On the other hand, **1** and **2** could be predicted to be more potent inhibitors than **3**−**5** when the ΔG_b^{aq} values were used to estimate the binding affinities for TrkA, implying that the solvation free energy term should be included in the scoring function to explain their relative inhibitory activities. Indeed, the ΔG^{sol} values become more negative in going from **1** and **2** to **3**−**5**, which means that the former compounds should be less stabilized in aqueous solution than the latter ones. As a consequence of such a decrease in desolvation cost for complexation in the ATP-binding site of TrkA, the ΔG_b^{aq} values of **1** and **2** become more favorable than those of **3**−**5**. **1** and **2** could thus be identified as the most potent TrkA inhibitors in the virtual screening by virtue of the implementation of the solvation free energy term for a putative inhibitor. Similarly, the low desolvation cost for **6** can be attributed to the experimental finding that its K_d value is similar to those of **3**−**5**, because the ΔG_b^{gas} value of the former is found to be less favorable than the latter by 3.8−8.0 kcal/mol. These results further exemplify the superiority of the new scoring function with the solvation free energy term to the previous one. The significant contributions of ΔG^{sol} to ΔG_b^{aq} found in this study also indicate that in order to enhance the potency of a TrkA inhibitor with a structural modification, the resulting increase in the strength of enzyme−inhibitor interactions should be sufficient to surmount the increased stabilization in aqueous solution.

The importance of ligand solvation effects on the accuracy of virtual screening was further assessed by comparing the calculated and experimental ΔG_b (ΔG_b^{exp}) values for the six compounds for which K_d values were determined. The squared correlation coefficient between ΔG_b^{aq} and ΔG_b^{exp} was calculated to be 0.79, as compared to 0.08 in the comparison between ΔG_b^{gas} and ΔG_b^{exp} values. This indicates that the scoring function in the virtual screening can make sense only when the effects of ligand solvation are taken into account.

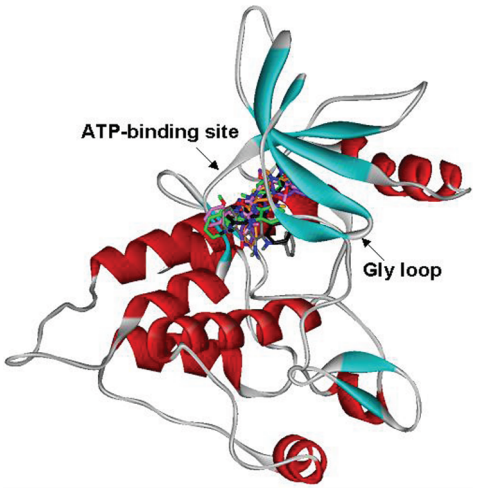


Figure 4. Comparative view of the binding modes **1**−**6** in the ATP binding site of TrkA. Carbon atoms of **1**, **2**, **3**, **4**, **5**, and **6** are indicated in black, green, gray, purple, pink, and orange, respectively.

It is difficult to calculate the entropic and enthalpic contributions to binding free energy separately in docking simulations because the entropic term is merged into the enthalpic function to produce the scoring function. However, the extent of entropic contribution to binding free energy can be estimated in a qualitative manner because the entropic term is assumed to be proportional to the number of rotatable bonds (N_{tor}) in the current form of scoring function. We note in this regard that N_{tor}'s of **1**, **2**, and **4** are two or three, while those of **3**, **5**, and **6** range from four to eight. This indicates that the former inhibitors would undergo less loss of entropy than the latter ones upon binding in the ATP-binding site of TrkA. On the other hand, the latter inhibitors should be energetically more stabilized than the former ones in the ATP-binding site to overcome the more entropic penalty for complexation.

Structural insights into the inhibitory mechanisms of the identified TrkA inhibitors were gained through the comparative investigations of their binding modes. Figure 4 shows the lowest-energy conformations of **1**−**6** in the ATP binding site of TrkA calculated using the modified AutoDock program. The results of these docking simulations were self-consistent in the sense that the inhibitors were positioned in similar configurations with comparable interactions with the amino acid residues in the ATP binding site. For example, the hydrogen-bonding groups on the inhibitors appeared to point toward the backbone groups at the ATP binding site with their hydrophobic groups being situated in proximity to the Gly loop. The possibility of allosteric TrkA inhibition was examined by conducting the docking simulations using the grid maps for the receptor model that included the entire TrkA structure. A binding configuration in which an inhibitor resided outside the ATP binding site was not observed for any of the six new inhibitors. These results support the possibility that the newly identified inhibitors may impair the TrkA catalytic activity by specifically binding to the ATP binding site.

We next turn to the identification of the detailed interactions responsible for stabilizing the identified TrkA inhibitors in the

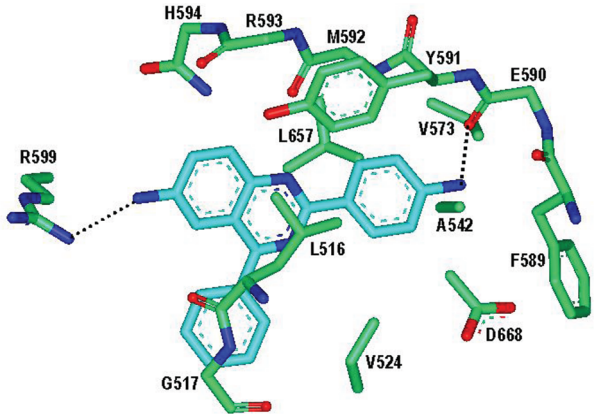


Figure 5. Calculated binding mode of **1** in the ATP-binding site of TrkA. Carbon atoms of the protein and the ligand are indicated in green and cyan, respectively. Each dotted line indicates a hydrogen bond.

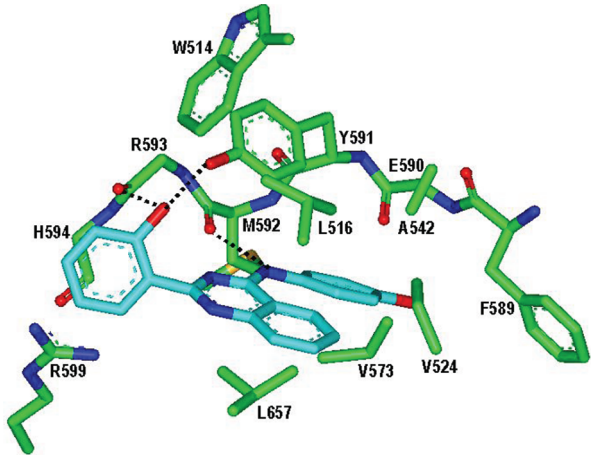


Figure 6. Calculated binding mode of **2** in the ATP-binding site of TrkA. Carbon atoms of the protein and the ligand are indicated in green and cyan, respectively. Each dotted line indicates a hydrogen bond.

ATP-binding site. The calculated binding mode of **1** in the ATP-binding site of TrkA is shown in Figure 5. It is noted that one of the terminal anilinic groups of **1** donates a hydrogen bond to the backbone aminocarbonyl oxygen of Glu590. The other anilinic group also appears to form a hydrogen bond with the side-chain guanidium ion of Arg599. These two hydrogen bonds seem to be the most significant binding forces to stabilize **1** in the ATP-binding site because no stronger interactions involving the positive and negative ions were observed in the calculated TrkA-**1** complex. The inhibitor **1** can be further stabilized in the ATP-binding site of TrkA by the hydrophobic interactions of its nonpolar aromatic rings with the side chains of Leu516, Val524, Ala542, Val573, Phe589, Tyr591, and Leu657. Judging from the overall structural features derived from docking simulations, the micromolar inhibitory activity of **1** can be attributed to the establishment of the multiple hydrogen bonds and hydrophobic interactions in a simultaneous fashion in the ATP-binding site of TrkA. Considering the low molecular weight of ~330, **1** is expected to serve as a good inhibitor scaffold from which much more potent inhibitors can be derivatized.

Figure 6 shows the calculated binding mode of **2** in the ATP-binding site of TrkA. We note that one of the terminal phenolic

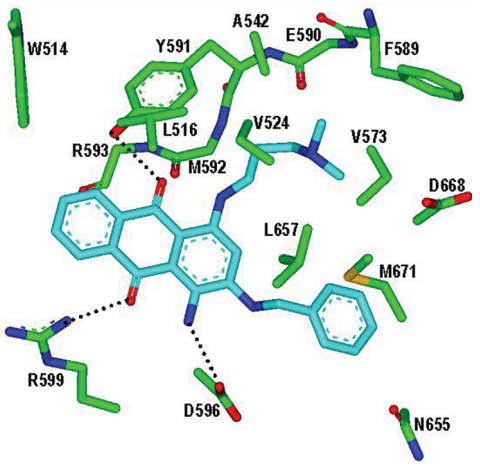


Figure 7. Calculated binding mode of **3** in the ATP-binding site of TrkA. Carbon atoms of the protein and the ligand are indicated in green and cyan, respectively. Each dotted line indicates a hydrogen bond.

groups of **2** receives and donates a hydrogen bond from the side chain of Tyr591 and to the backbone aminocarbonyl oxygen of Arg593, respectively. A stable hydrogen bond is also established between the backbone aminocarbonyl oxygen of Met592 and the −NH− group of the inhibitor bridging the phenyl and pyrimidine rings. The pattern for the establishment of hydrogen bonds in the TrkA-**2** complex is thus quite different from that in the TrkA-**1** complex. It is also noteworthy that the number of hydrogen bonds increases from two in TrkA-**1** to three in the TrkA-**2** complex, which would have the effect of increasing the inhibitory activity for the target protein. This strengthening of the hydrogen bond interactions can be invoked to explain the decrease in ΔG_b^{gas} by 2.12 kcal/mol in going from **1** to **2** (Table 1). On the other hand, the hydrophobic interactions in the TrkA-**2** complex appear to be established in a similar way to those in the TrkA-**1** complex: its aromatic rings form the van der Waals contact with the side chains of Leu516, Val524, Ala542, Val573, Phe589, Tyr591, and Leu657. Despite the increased hydrogen bond stabilization in the TrkA-**2** complex, the inhibitory activity of **2** would be limited similar to that of **1** due to the higher desolvation cost for **2** than **1** by 1.38 kcal/mol. Nonetheless, **2** is also expected to serve as a good inhibitor scaffold for the discovery of anticancer and antinociceptive drugs due to the micromolar inhibitory activity and low molecular weight.

Judging from the ΔG_b^{gas} values shown in Table 1, compound **3** seems to be bound in the ATP-binding site of TrkA in the strongest form. As can be seen in Figure 7, the two carbonyl groups of **3** play a role of hydrogen bond receptor with respect to the side-chain phenolic and guanidium ion groups of Tyr591 and Arg599. An additional hydrogen bond is also established between the anilinic moiety of the inhibitor and the side-chain carboxylate ion of Asp596. Although the number of hydrogen bonds in the TrkA-**3** complex is the same as that in the TrkA-**2** complex, the hydrogen-bond stabilization seems to be greater in the former than in the latter due to the involvement of the two ionized residues (Arg599 and Asp596) in the hydrogen bond interactions as compared to none in the TrkA-**2** complex. The hydrophobic side chains of Leu516, Val524, Ala542, Val573, Phe589, Tyr591, and Leu657 that have been shown to stabilize **1** and **2** are also observed at the interface in the TrkA-**3** complex. Beside these common groups, the side chains of Trp514 and Met671

dx.doi.org/10.1021/ci200378s |J. Chem. Inf. Model. 2011, 51, 2986–2993

also appear to establish van der Waals contacts with the nonpolar groups of **3**, indicating that hydrophobic interactions in the TrkA-**3** complex should also be established in a stronger form than in the TrkA-**1** and TrkA-**2** complexes. The strengthening of both hydrogen bond and hydrophobic interactions in the ATP binding site of TrkA can thus be invoked to explain the most negative ΔG_b^{gas} value of **3**. Nonetheless, the inhibitory activity of **3** seems to be limited with the K_d value of 35 μM due to the high desolvation cost of more than 17 kcal/mol. This high solvation free energy of **3** should be attributed to the presence of the tertiary amine group at the end of the molecular structure that should be assumed to be positively charged at physiological pH. Thus, structural and energetic features found in this study confirm that in order to optimize the activity of a TrkA inhibitor, the structural modifications should be made in such a way to maximize the attractive interactions in the ATP-binding site and to minimize the desolvation cost for complexation in a simultaneous fashion.

CONCLUSIONS

We have identified six novel TrkA inhibitors by applying a computer-aided drug design protocol involving the homology modeling and the structure-based virtual screening with docking simulations under consideration of the effects of ligand solvation in the scoring function. These inhibitors have desirable physicochemical properties as a drug candidate and reveal a moderate potency with K_d values ranging from 3 to 40 μM. Therefore, each of the newly discovered inhibitors deserves consideration for further development by SAR studies to optimize the inhibitory activities. The energetic features associated with complexation of TrkA and its inhibitors indicate that the solvation free energy term should be necessary in the scoring function to explain the relative inhibitory activities of the identified inhibitors. Such a significant contribution of the ligand solvation effects to the binding free energy also indicates that the increased stabilization in solution due to structural changes should be overcome by an even stronger enzyme−inhibitor interaction in order to enhance the inhibitory activity. Detailed binding mode analyses with docking simulations show that the inhibitors can be stabilized in ATP-binding site of TrkA by the simultaneous establishment of multiple hydrogen bonds and van der Waals contacts.

ASSOCIATED CONTENT

Supporting Information. Chemical structures and docking simulation results of the TrkA inhibitors with POC values lower than 50. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: +82-2-3408-3766, +82-42-350-2811. Fax: +82-2-3408-4334, +82-42-350-2812. E-mail: hspark@sejong.ac.kr (H.P.); hongorg@kaist.ac.kr (S.H.).

ACKNOWLEDGMENT

This work was supported by Grant No. K11061 from the Korea Institute of Oriental Medicine (KIOM) and through the National Research Foundation of Korea (NRF-2011-0003609, NRF-2011-0016436 and 2011-0020322).

REFERENCES

(1) Gschwind, A.; Fischer, O. M.; Ullrich, A. The discovery of receptor tyrosine kinases: targets for cancer therapy. *Nat. Rev. Cancer* **2004**, *4*, 361–370.

(2) Patapoutian, A.; Reichardt, L. F. Trk receptors: mediators of neurotrophin action. *Curr. Opin. Neurobiol.* **2001**, *11*, 272–280.

(3) Papatsoris, A. G.; Liolotsa, D.; Deliveliotis, C. Manipulation of the nerve growth factor network in prostate cancer. *Expert Opin. Invest. Drugs* **2007**, *16*, 303–309.

(4) Sclabas, G. M.; Fujioka, S.; Schmidt, C.; Li, Z.; Frederick, W. A. I.; Yang, W.; Yokoi, K.; Evans, D. B.; Abbruzzese, J. L.; Hess, K. R.; Zhang, W.; Fidler, I. J.; Chiao, P. J. Overexpression of tropomyosin-related kinase B in metastatic human pancreatic cancer cells. *Clin. Cancer Res.* **2005**, *11*, 440–449.

(5) Bardelli, A.; Parsons, D. W.; Silliman, N.; Ptak, J.; Szabo, S.; Saha, S.; Markowitz, S.; Willson, J. K. V.; Parmigiani, G.; Kinzler, K. W.; Vogelstein, B.; Velculescu, V. E. Mutational analysis of the tyrosine kinome in colorectal cancers. *Science* **2003**, *300*, 949.

(6) Nakagawara, A. Trk receptor tyrosine kinases: a bridge between cancer and neural development. *Cancer Lett.* **2001**, *169*, 107–114.

(7) Davies, H.; Hunter, C.; Smith, R.; Stephens, P.; Greenman, C.; Bignell, C.; Teague, J.; Butler, A.; Edkins, S.; Stevens, C.; Parker, A.; O'Meara, S.; Avis, T.; Barthorpe, S.; Brackenbury, L.; Buck, G.; Clements, J.; Cole, J.; Dicks, E.; Edwards, K.; Forbes, S.; Gorton, M.; Gray, K.; Halliday, K.; Harrison, R.; Hills, K.; Hinton, J.; Jones, D.; Kosmidou, V.; Laman, R.; Lugg, R.; Menzies, A.; Perry, J.; Petty, R.; Raine, K.; Shepherd, R.; Small, A.; Solomon, H.; Stephens, Y.; Tofts, C.; Varian, J.; Webb, A.; West, S.; Widaa, S.; Yates, A.; Brasseur, F.; Cooper, C. S.; Flanagan, A. M.; Green, A.; Knowles, M.; Leung, S. Y.; Looijenga, L. H. J.; Malkowicz, B.; Pierotti, M. A.; Teh, B. T.; Yuen, S. T.; Lakhani, S. R.; Easton, D. F.; Weber, B. L.; Goldstraw, P.; Nicholson, A. G.; Wooster, R.; Stratton, M. R.; Futreal, P. A. Somatic mutations of the protein kinase gene family in human lung cancer. *Cancer Res.* **2005**, *65*, 7591–7595.

(8) Lagadec, C.; Meignan, S.; Adriaenssens, E.; Foveau, B.; Vanhecke, E.; Romon, R.; Toillon, R. A.; Oxombre, B.; Hondermarck, H.; Le Bourhis, X. TrkA overexpression enhances growth and metastasis of breast cancer cells. *Oncogene* **2009**, *28*, 1960–1970.

(9) Tognon, C.; Knezevich, S. R.; Huntsman, D.; Roskelley, C. D.; Melnyk, N.; Mathers, J. A.; Becker, L.; Carneiro, F.; MacPherson, N.; Horsman, D.; Poremba, C.; Sorensen, P. H. B Expression of the ETV6-NTRK3 gene fusion as a primary event in human secretory breast carcinoma. *Cancer Cell* **2002**, *2*, 367–376http://www.cell.com/cancer-cell/abstract/.

(10) Brodeur, G. M. Neuroblastoma: biological insights into a clinical enigma. *Nat. Rev. Cancer* **2003**, *3*, 203–216.

(11) Indo, Y.; Tsuruta, M.; Hayashida, Y.; Karim, M. A.; Ohta, K.; Kawano, T.; Mitsubuchi, H.; Tonoki, H.; Awaya, Y.; Matsuda, I. Mutations in the TrkA/NGF receptor gene in patients with congenital insensitivity to pain with anhidrosis. *Nat. Genet.* **1996**, *13*, 485–488.

(12) Read, S. J.; Dray, A. Osteoarthritic pain: a review of current, theoretical and emerging therapeutics. *Expert Opin. Invest. Drugs* **2008**, *17*, 619–640.

(13) Wang, T.; Yu, D.; Lamb, M. L. Trk kinase inhibitors as new treatments for cancer and pain. *Expert Opin. Ther. Patents* **2009**, *19*, 305–319.

(14) Wang, T; Lamb, M. L.; Scott, D. A.; Wang, H.; Block, M. H.; Lyne, P. D.; Lee, J. W.; Davies, A. M.; Zhang, H.-J.; Zhu, Y.; Gu, F.; Han, Y.; Wang, B.; Mohr, P. J.; Kaus, R. J.; Josey, J. A.; Hoffmann, E.; Thress, K.; MacIntyre, T.; Wang, H.; Omer, C. A.; Yu, D. Identification of 4-aminopyrazolylpyrimidines as potent inhibitors of Trk kinases. *J. Med. Chem.* **2008**, *51*, 4672–4684.

(15) Gingrich, D. E.; Yang, S. X.; Gessner, G. W.; Angeles, T. S.; Hudkins, R. L. Synthesis, modeling, and in vitro activity of (3′*S*)-*epi*-K-252a analogues. Elucidating the stereochemical requirements of the 3′-sugar alcohol on TrkA tyrosine kinase activity. *J. Med. Chem.* **2005**, *48*, 3776–3783.

(16) Tripathy, R.; Angeles, T. S.; Yang, S. X.; Mallamo, J. P. TrkA kinase inhibitors from a library of modified and isosteric Staurosporine aglycone. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3551–3555.

(17) Lippa, B.; Morris, J.; Corbett, M.; Kwan, T. A.; Noe, M. C.; Snow, S. L.; Gant, T. G.; Mangiaracina, M.; Coffey, H. A.; Foster, B.; Knauth, E. A.; Wessel, M. D. Discovery of novel isothiazole inhibitors of the TrkA kinase: Structure−activity relationship, computer modeling, optimization, and identification of highly potent antagonists. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3444–3448.

(18) Wood, E. R.; Kuyper, L.; Petrov, K. G.; Hunter, R. N., III; Harris, P. A.; Lackey, K. Discovery and in vitro evaluation of potent TrkA kinase inhibitors: oxindole and aza-oxindoles. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 953–957.

(19) Kim, S. H.; Tokarski, J. S.; Leavitt, K. J.; Fink, B. E.; Salvati, M. E.; Moquin, R.; Obermeier, M. T.; Trainor, G. L.; Vite, G. G.; Stadnick, L. K.; Lippy, J. S.; You, D.; Lorenzi, M. V.; Chen, P. Identification of 2-amino-5-(thioaryl)thiazoles as inhibitors of nerve growth factor receptor TrkA. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 634–639.

(20) Warren, G. L.; Andrews, C. W.; Capelli, A. M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. A critical assessment of docking programs and scoring functions. *J. Med. Chem.* **2006**, *49*, 5912–5931.

(21) Shoichet, B. K.; Leach, A. R.; Kuntz, I. D. Ligand solvation in molecular docking. *Proteins* **1999**, *34*, 4–16.

(22) Wittman, M. D.; Carboni, J. M.; Yang, Z.; Lee, F. Y.; Antman, M.; Attar, R.; Balimane, P.; Chang, C.; Chen, C.; Discenza, L.; Frennesson, D.; Gottardis, M. M.; Greer, A.; Hurlburt, W.; Johnson, W.; Langley, D. R.; Li, A.; Li, J.; Liu, P.; Mastalerz, H.; Mathur, A.; Menard, K.; Patel, K.; Sack, J.; Sang, X.; Saulnier, M.; Smith, D.; Stefanski, K.; Trainor, G.; Velaparthi, U.; Zhang, G.; Zimmermann, K.; Vyas, D. M. Discovery of a 2,4-disubstituted pyrrolo[1,2-*f*][1,2,4]triazine inhibitor (BMS-754807) of insulin-like growth factor receptor (IGF-1R) kinase in clinical development. *J. Med. Chem.* **2009**, *52*, 7360–7363.

(23) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **1994**, *22*, 4673–4680.

(24) Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815.

(25) Fiser, A.; Do, R. K.; Sali, A. G. Modeling of loops in protein structures. *Protein Sci.* **2000**, *9*, 1753–1773.

(26) Sippl, M. J. Recognition of errors in three-dimensional structures of proteins. *Proteins* **1993**, *17*, 355–362.

(27) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639–1662.

(28) Park, H.; Lee, J.; Lee, S. Critical assessment of the automated AutoDock as a new docking tool for virtual screening. *Proteins* **2006**, *65*, 549–554.

(29) Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press: Oxford, 1997; pp 79−97.

(30) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings Adv. *Drug Delivery Rev.* **1997**, *23*, 3–25.

(31) Gasteiger, J.; Marsili, M. Iterative partial equalization of orbital electronegativity a rapid access to atomic charges. *Tetrahedron* **1980**, *36*, 3219–3228.

(32) Park, H.; Jeon, J. H. Cubic equation governing the outer-region dielectric constant of globular proteins. *Phys. Rev. E* **2007**, *75*, 021916.

(33) Stouten, P. F. W.; Frömmel, C.; Nakamura, H.; Sander, C. An effective solvation term based on atomic occupancies for use in protein simulations. *Mol. Simul.* **1993**, *10*, 97–120.

(34) Kang, H.; Choi, H.; Park, H. Prediction of molecular solvation free energy based on the optimization of atomic solvation parameters with genetic algorithm. *J. Chem. Inf. Model.* **2007**, *47*, 509–514.

(35) Chamberlain, S. D.; Wilson, J. W.; Deanda, F.; Patnaik, S.; Redman, A. M.; Yang, B.; Shewchuk, L.; Sabbatini, P.; Leesnitzer, M. A.; Groy, A.; Atkins, C.; Gerding, R.; Hassell, A. M.; Lei, H.; Mook, R. A.; Moorthy, G.; Rowand, J. L.; Stevens, K. L.; Kumar, R.; Shotwell, J. B. Discovery of 4,6-bis-anilino-1*H*-pyrrolo[2,3-*d*]pyrimidines: potent inhibitors of the IGF-1R receptor tyrosine kinase. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 469–473.

(36) Bossi, R. T.; Saccardo, M. B.; Ardini, E.; Menichincheri, M.; Rusconi, L.; Magnaghi, P.; Orsini, P.; Avanzi, N.; Borgia, A. L.; Nesi, M.; Bandiera, T.; Fogliatto, G.; Bertrand, J. A. Crystal structures of anaplastic lymphoma kinase in complex with ATP competitive inhibitors. *Biochemistry* **2010**, *49*, 6813–6825.

(37) Baker, D.; Sali., A. Protein structure prediction and structural genomics. *Science* **2001**, *294*, 93–96.

(38) Fabian, M. A.; Biggs, W. H., III; Treiber, D. K.; Atteridge, C. E.; Azimioara, M. D.; Bendetti, M. G.; Carter, T. A.; Ciceri, P.; Edeen, P. T.; Floyd, M.; Ford, J. M.; Galvin, M.; Gerlach, J. L.; Grotzfeld, R. M.; Herrgard, S.; Insko, D. E.; Insko, M. A.; Lai, A. G.; Lelias, J.-M.; Mehta, S. A.; Milanov, Z. V.; Velasco, A. M.; Wodicka, L. M.; Patel, H. K.; Zarrinkar, P. P.; Lockhart, D. J. A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat. Biotechnol.* **2005**, *23*, 329–336.