JOURNAL OF CHEMICAL INFORMATION AND MODELING

LETTER

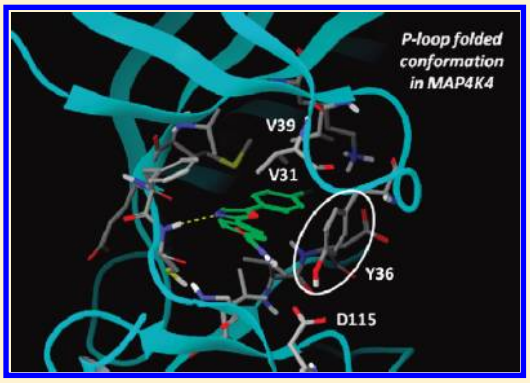
pubs.acs.org/jcim

Understanding the Impact of the P-loop Conformation on Kinase Selectivity

Cristiano R. W. Guimarães,*[,†] Brajesh K. Rai,[‡] Michael J. Munchhof,[†] Shenping Liu,[§] Jian Wang,[†] Samit K. Bhattacharya,[†] and Leonard Buckbinder[∥]

[†]Worldwide Medicinal Chemistry Department, [‡]Computational Sciences Center of Emphasis, [§]Department of Structural Biology and Biophysics, and [∥]CVMED - Target Exploration, Pfizer Global Research and Development, Eastern Point Road, Groton, Connecticut 06340, United States

ABSTRACT:



This work addresses the link between selectivity and an unusual, folded conformation for the P-loop observed initially for MAP4K4 and subsequently for other kinases. Statistical and computational analyses of our crystal structure database demonstrate that inhibitors that induce the P-loop folded conformation tend to be more selective, especially if they take advantage of this specific conformation by interacting more favorably with a conserved Tyr or Phe residue from the P-loop.

INTRODUCTION

Kinases are considered as therapeutic targets in a variety of diseases such as cancer, diabetes, and arthritis. In recent years, many kinase small molecule inhibitors have been developed as potential disease treatments.[1] Despite an overall conserved protein structure, which render inhibitors more likely to cross-inhibit other kinases in the same family or even from other families,[2] kinases also display particularities that can be explored to achieve selectivity.[3] Classical examples are the equilibrium between the in and out conformations for the DFG loop,[4] the helix C orientation,[5] and the size of the selectivity pocket.[6]

In the case of the P-loop, also known as the glycine-rich loop, its conformational variability and the use of this information to build in selectivity for particular kinases have also been suggested.[7] However, a crystal structure obtained internally of the complex between the Mitogen-activated protein kinase 4 (MAP4K4) and a potent and selective inhibitor revealed a folded conformation for the P-loop that, to the best of our knowledge, is extremely unusual. The P-loop, which normally exists in an extended conformation, is a stretch of nine aminoacids that in MAP4K4 is delimited by residues V31 and V39.

MAP4K4, also called HGK or ZC1, is a serine/threonine protein kinase that belongs to the mammalian STE20/MAP4K family.[8] MAP4K4 is overexpressed in human cancers and cancer cell lines and appears to play an important role in cell transformation, invasiveness, adhesion, and migration.[9] Recent studies in adipose tissue, pancreas, muscle, and macrophages suggest that MAP4K4 could also be a potential target for antidiabetic drugs.[10−12] As such, a docking virtual screen using the crystal structure where the MAP4K4 P-loop displays the folded conformation was performed in order to identify additional chemical matter with high selectivity. This exercise led to the discovery of compound **1** and subsequent 2D similarity searching identified compound **2** (Scheme 1), both potent and very selective inhibitors. The heat maps for compounds **1** and **2** in Figure 1 reveal that MAP4K4 is the only kinase in a panel of 36 kinases with significant inhibition at a compound concentration of 1 μM.

Crystal structures were obtained between MAP4K4 and compounds **1** and **2**. The P-loop in those complexes also displays a folded conformation. Although MAP4K4 adopts the folded conformation when bound to compounds **1** and **2**, it should be noted they are not similar in structure to the inhibitor present in the original crystal structure. Figure 2 illustrates the complex for compound **1**; a very similar structure is obtained for compound **2** (Figure 3). It is then plausible that there is a link between the P-loop folded conformation and the high selectivity observed for those inhibitors. This work aims at investigating the uniqueness of the folded conformation among kinases and what drives it, as well as verifying whether compounds complexed to kinases displaying this conformation tend to be more selective.

RESULTS AND DISCUSSION

How unusual is the P-loop folded conformation? In the P-loop folded conformation, the binding site adopts a tunnel-like shape that provides a significant hydrophobic enclosure to the inhibitor (Figure 2). It is important to highlight the extensive $\pi-\pi$ interactions between the MAP4K4 P-loop Y36 residue and the three aromatic rings of compounds **1** and **2** (Figures 2 and 3). The terminal amide and the phenol OH of **1** and **2**, respectively, are the only groups that escape the hydrophobic tunnel, being

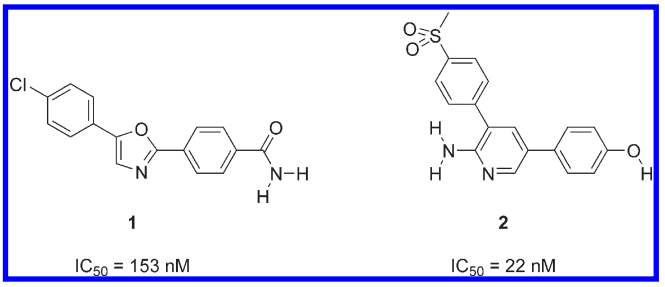
Published: May 13, 2011

dx.doi.org/10.1021/ci200153c | J. Chem. Inf. Model. 2011, 51, 1199–1204

partially exposed to the solvent and in close proximity to the D115 residue; in this geometry, water-mediated hydrogen bonds may be established.

A search in our database of protein−ligand complexes containing 2690 internal and external kinase crystal structures was conducted to verify the occurrence of the folded conformation. Specifically, the complex between MAP4K4 and **1** was used as reference in a Shape similarity search for the stretch of aminoacids that are part of the P-loop. The P-loop folded conformation is incredibly rare. It is seen in only 62 out of 2690 complexes for 7 different kinase targets out of 56 covered in the database (Table 1), an incidence of 12.5%. Of those complexes, 18 are available in the Protein Data Bank (**ABL** − 1M52, 2GQG, 2G2H, 2HZI, 2Z60, 3DK3, 3IK3; **AURA** − 3DJ5, 3DJ6; **cMet** − 3EFK; **FGFR1** − 1FGI, 3C4F, 3JS2; **p38** − 3BX5, 3C5U, 3DS6, 3D7Z, 3E92). Although this unusual conformation is seen for ABL, ACK1, AURA, cMet, FGFR1, MAP4K4, and p38, the extended conformation for the P-loop is still most prevalent in those kinases (Table 1). This suggests that they must have the intrinsic propensity to fold, but that seems to be driven by the ligand. If the ligands are clustered in order to eliminate redundancy in the analysis, there are 40 unique scaffolds that seem to induce the P-loop folded conformation.

Scheme 1. MAP4K4 Inhibitors



What drives the P-loop folded conformation? The first question is whether the intrinsic propensity to adopt the folded conformation derives from the presence of specific aminoacids in

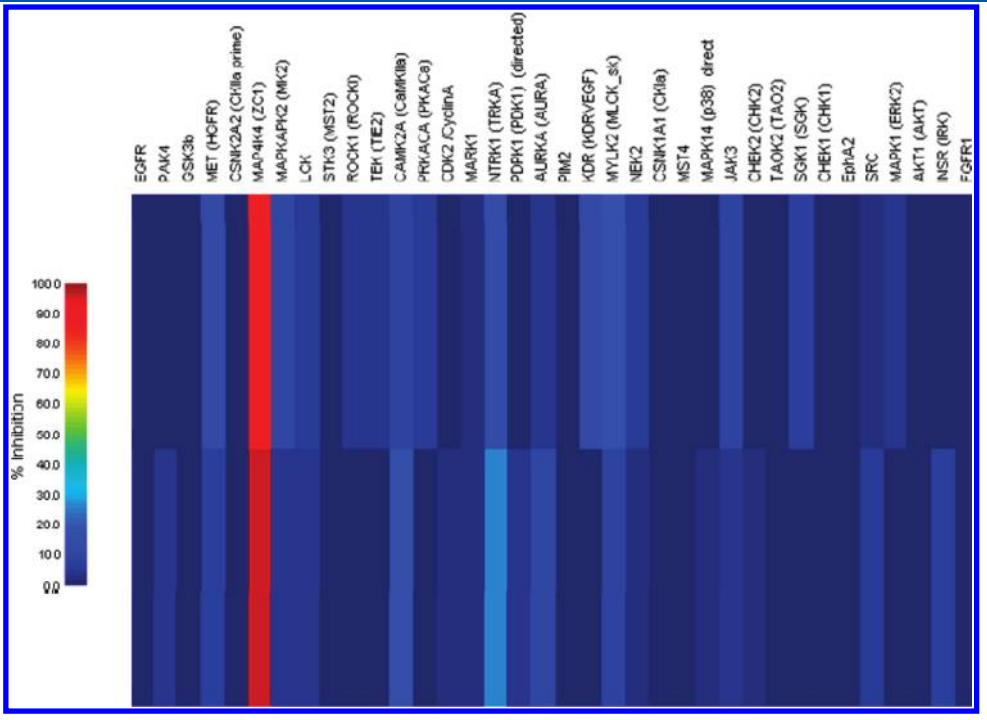


Figure 1. Heat maps obtained from an Invitrogen panel with 36 kinases. Compounds **1** (top) and **2** (bottom) were tested at 1 μM.

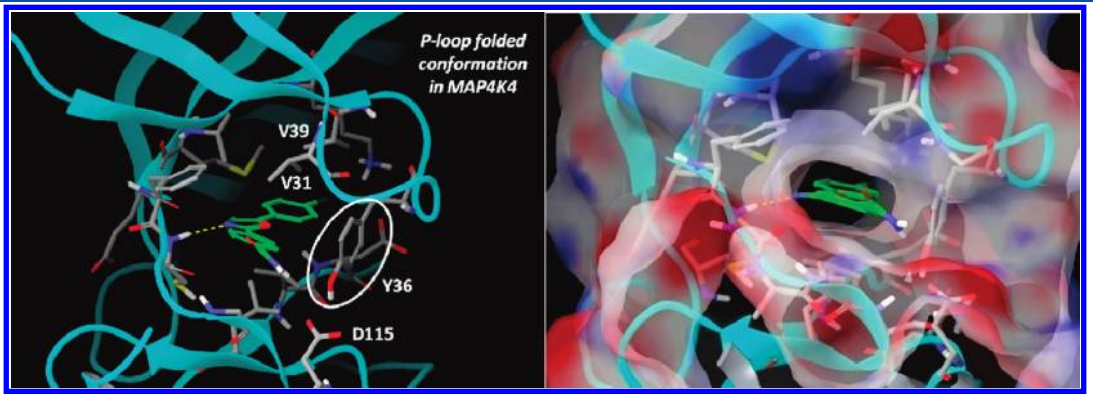


Figure 2. Crystal structure for the complex between MAP4K4 and compound **1**.

dx.doi.org/10.1021/ci200153c |*J. Chem. Inf. Model.* 2011, 51, 1199–1204

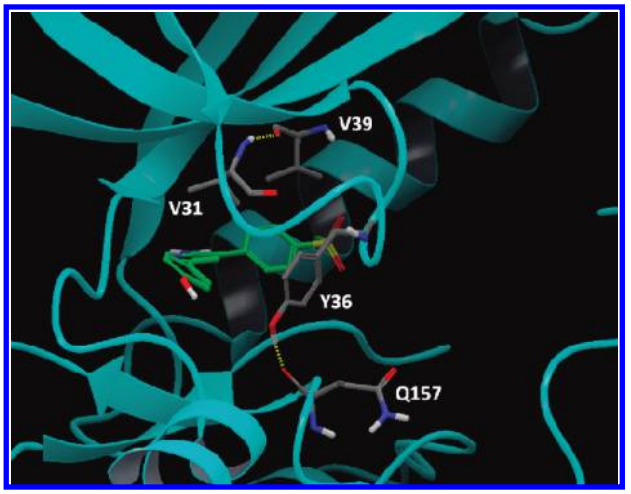


Figure 3. Crystal structure for the complex between MAP4K4 and compound **2**. The hydrogen bond between Y36 and Q157 is highlighted. Most residues are omitted for clarity.

Table 1. Kinases Displaying the P-loop Folded Conformation in the Database

target	P-loop folded complexes	total complexes	% of complexes with folded conf.	P5 residue
ABL	9	25	36	Y
ACK1	1	50	2	F
AURA	2	57	4	F
cMet	8	92	9	F
FGFR1	3	9	33	F
MAP4K4	4	12	33	Y
p38	35	331	11	Y

the P-loop. The sequences for ABL, ACK1, AURA, cMet, FGFR1, MAP4K4, and p38 were aligned to 435 other kinases. Table 2 demonstrates that the residues at positions labeled P1, P3, P6, and L2 are conserved for the subset of seven kinase targets. On the other hand, the residues at positions P2, P4, and P7 are very diverse. It is interesting to see that almost any residue can be found at those positions for the 442 kinases analyzed (Table 2), but they tend to be polar for ABL, ACK1, AURA, cMet, FGFR1, MAP4K4, and p38. The residues at P2, P4, and P7 are exposed to the solvent in the folded conformation. As for L1, the observed residues are either L, I, or V out of seven that can be found at that position. Finally, the residue at P5 is either Y or F. In this case, 11 different aminoacids are seen for the group of 442 kinases.

The simple presence of specific aminoacids in the P-loop is unlikely to be the driver for the folding process as 273 out of 442 kinases also have G residues at P1, P3, and P6, F or Y at P5, and V at L2. If the residues at each individual position are limited to only the ones shown in Table 2, there would still be 82 different kinases that fall into that subset. This would represent a much higher occurrence of the folded conformation than what is seen in the crystal structure database. However, it is obvious that a statistical analysis is only as good as the information that is available. In other words, the existence of the P-loop folded conformation cannot be ruled out for other kinase targets not included in the database or present but not bound to the "right" compound.

Table 2. Kinases Displaying the P-loop Folded Conformation in the Database

target	L1	P1	P2	P3	P4	P5	P6	P7	L2
ABL	L	G	G	G	Q	Y	G	E	V
ACK1	L	G	D	G	S	F	G	V	V
AURA	L	G	K	G	K	F	G	N	V
cMet	I	G	R	G	H	F	G	C	V
FGFR1	L	G	E	G	C	F	G	Q	V
MAP4K4	V	G	N	G	T	Y	G	Q	V
p38	V	G	S	G	A	Y	G	S	V
Diversity[a]	7	9	19	4	18	11	7	19	7

[a] Number of different residues found at a given position along the P-loop sequence obtained from 442 kinases.

The above analysis assumes that a residue at a given position along the P-loop is independent of the other ones. It is possible that the folding of the P-loop actually occurs for specific residue combinations and is not based on the simple presence of the observed residues in Table 2. This would definitely reduce the number of kinases with propensity for the P-loop to fold but the limited number of known targets displaying the folded conformation makes the identification of other potential sequences difficult. Extending the analysis to residues not in direct contact with the ligand did not provide any insights for the conformational driver as they are quite diverse among the kinases displaying the folded conformation.

Although it is difficult to pinpoint which P-loop sequences are more prone to fold, and consequently determine what additional kinases would have this property, it is clear that the aromatic residues F or Y at P5 are essential. They establish extensive $\pi-\pi$ and/or hydrophobic interactions with the inhibitors in the 62 crystal structures where the P-loop is folded, suggesting this interaction plays an important role on the stabilization of this specific conformation. In the case of $\pi-\pi$ interactions, the T-shape orientation and variations are the most prevalent, but the parallel orientation is also observed. The P-loop is quite ordered in the 62 complexes with B-factors for backbone and side-chain atoms ranging from 10 to 50 Å^2, with side-chains displaying somewhat larger values. One should note that the simple presence of F and Y at P5 does not guarantee a folded conformation, even for ABL, ACK1, AURA, cMet, FGFR1, MAP4K4, and p38, because the extended conformation is the most prevalent for them (Table 1). On the same topic, a closer look at Table 1 suggests that the P-loop of the kinases with a Y residue at P5 might have a higher tendency to fold. To investigate that, the intermolecular interactions between the ligand and the residue at P5 were calculated for the 62 kinases in the database displaying the P-loop folded conformation. Figure 4 shows that the complexes with Y at P5 tolerate less favorable interactions between the aromatic residue and the ligand, whereas the lower limit for the interaction with F at P5 seems to be −3.5 kcal/mol. This suggests that the folding process for the latter might be harder and require additional/more favorable interactions with the ligand. The lower energy threshold to achieve the folded conformation for the P-loop in the case of the Y residue could be due to an additional stabilization provided by a hydrogen bond between its side-chain OH group and a backbone carbonyl, as illustrated by the interaction between Y36 and Q157 in the complex between MAP4K4 and compound **2** (Figure 3).

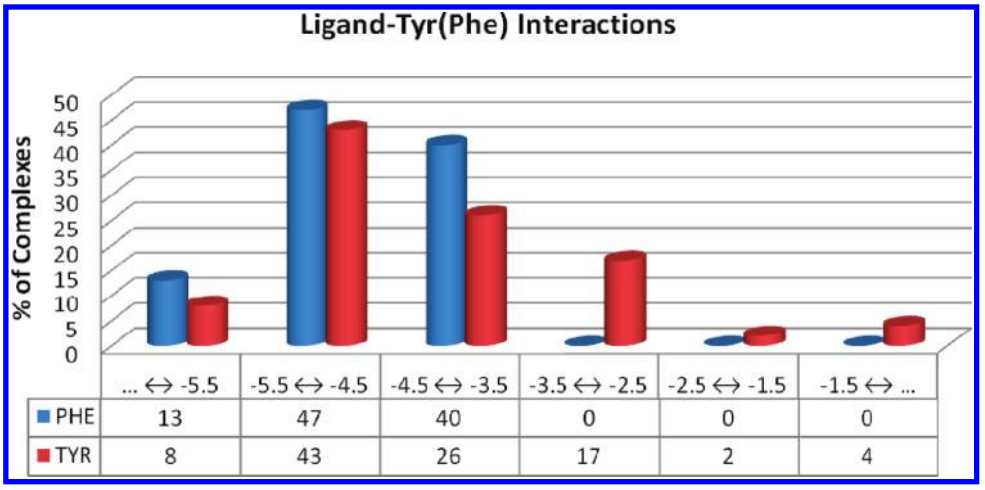


Figure 4. Distribution of complexes with respect to ligand−P5 residue intermolecular interactions for the 62 complexes that display the P-loop folded conformation. The distributions for complexes bearing F (blue bars) and Y (red bars) residues at P5 are plotted individually. Interaction energies are binned every 1 kcal/mol.

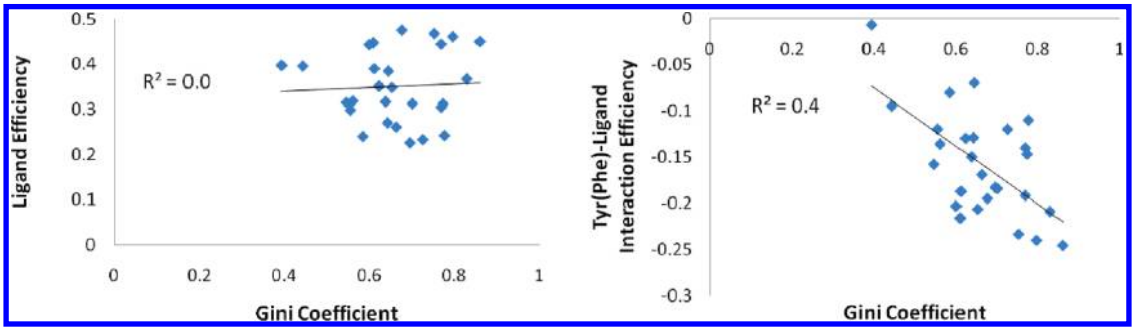


Figure 5. Gini coefficient versus ligand efficiency (left) and Tyr(Phe)−ligand interaction efficiency (right).

Finally, one should note that an inhibitor that is potent against a P-loop folded kinase does not have to be active against the other kinases with that property, even if a Y residue is present at P5. Although those kinases achieve the folded conformation more easily, potency and selectivity are both derived from a multitude of effects. Compound **2**, for example, is unable to inhibit AURA, cMet, FGFR1, and p38 at 1 μM as shown by the panel in Figure 1. It is also selective against ABL and ACK1, which are part of another kinase panel routinely used. While AURA, ACK1, cMet, and FGFR1 display a F residue at P5, ABL and p38, like MAP4K4, have a Y residue instead.

Are compounds more selective when the P-loop displays the folded conformation? The Gini coefficient (G),[13] which ranges from 0 (no selectivity) to 1 (maximum selectivity), was employed here as a measure of selectivity. It simplifies the analysis of the inhibition data obtained for a large number of kinases; the information hidden within hundreds of K_i, IC_{50}, or single-point data is difficult to interpret, particularly if selectivity of several compounds needs to be compared. The G values for compounds **1** and **2** are 0.6. As a comparison, the values for the nonselective Staurosporine[14] and the very selective Lapatinib,[15] a potent dual inhibitor of epidermal growth factor receptor (EGFR, ErbB-1) and ErbB-2, are 0.07 and 0.97, respectively.

Out of the 62 ligands bound to kinases displaying the folded conformation for the P-loop, 27 also have available single-point inhibition data at 1 μM from kinase panels. It is interesting to see that most of them are selective; only 6 ligands have G values below 0.6, an arbitrary cutoff between high and low selectivity based on the G values for compounds **1** and **2**. In other words, although inhibitors that induce the P-loop folded conformation tend to be very selective, it seems that is not enough to guarantee selectivity in some cases.

Besides the panel data, IC_{50} values are also available for the 27 compound subset upon inhibition of their respective P-loop folded kinases. The breakdown is 3, 1, 7, 4, and 12 IC_{50} values against ABL, AURA, cMet, MAP4K4, and p38, respectively. They are all sub μM potent inhibitors. One could argue that their high selectivity has nothing to do with the P-loop conformation, but it is rather a function of the ligand efficiency coming from other binding contributions. However, Figure 5 shows that there is no correlation between general ligand efficiency and the Gini coefficient. An interpretation for this is that the features of the compound that make it more or less efficient against their respective P-loop folded kinases may also affect the potency (efficiency) against other kinases, e.g., ligand desolvation penalty, ligand flexibility that impact strain and entropy penalty, and hydrogen bonds to the hinge. These are all somewhat nonspecific binding contributions. On the other hand, if the Gini coefficient is plotted against the ligand interaction efficiency with the residue at P5, defined as the interaction energy between that residue and the ligand divided by the number of heavy atoms of the ligand, a correlation emerges (Figure 5). In other words, kinase inhibitors that induce the folding of the P-loop get increasingly selective if they take

advantage of the folded conformation by interacting more efficiently with the Y or F residue at P5.

In summary, the P-loop folded conformation is very unusual and was found in only 62 complexes belonging to 7 kinase targets (ABL, ACK1, AURA, FGFR1, cMet, MAP4K4, and p38) out of 2690 internal and external kinase crystal structures contained in our database. The fact that the extended conformation is still the most prevalent in those kinases indicates that, while they must have the intrinsic propensity to fold, it seems to be driven by the ligand. However, a scaffold capable of inducing the folded conformation might not necessarily achieve that for all kinases with this property, especially when a F residue is present at the P-loop P5 position; they seem to have a higher energy threshold than the Y residue for the folding to occur. It is also clear from our analysis that inducing the folded conformation only results in greater selectivity in kinase panels if the inhibitors interact efficiently with the F or Y residue at P5. A potential advantage of the P-loop folded conformation, besides selectivity, is that it leads to very potent inhibitors at low molecular weight. This contrasts, for example, with the chemical matter typically observed in the case of kinases that adopt the DFG-out conformation. Finally, work is in progress to determine what drives the folding of the P-loop. The results should help predicting other kinases with this property and guide the design of selective kinase inhibitors.

EXPERIMENTAL SECTION

The internal and external kinase crystal structures in our database were analyzed to identify the occurrence of the P-loop folded conformations. An important feature of this database is that all structures from a given family are prealigned based on rigid body superposition of key, conserved residues from that family. As a result, all kinase structures are aligned in a common reference frame, making the comparison of their structures, in particular of their P-loop conformations, relatively straightforward.

Using the P-loop conformation of MAP4K4 in a complex with compound **1** as the query, pairwise comparisons were performed to identify crystal structures with similar P-loop conformations. An internally developed Python script was used to automatically identify the P-loop residues and to extract the corresponding mainchain heavy atoms and coordinates (side chain and hydrogens are not used). Similarity of a given pair of P-loop conformations is calculated using the OEShape toolkit from OpenEye.[16] As described elsewhere,[17] this software implements a method for assessing similarity between a pair of molecules by representing their shape as atom-centered Gaussian functions. Using this representation, an in-place volume overlap is calculated between the reference and the database P-loop conformations, and a Tanimoto coefficient, which provides a normalized measure of similarity that ranges between 0 (no overlap) and 1 (identical P-loop conformations), is obtained.

As noted earlier, because the kinase structures used in this analysis were prealigned, an in-place comparison of the shapes of the P-loop conformations was sufficient to assess their similarity. In addition, while this software provides the ability to compute the overlap of both shape and chemical features, in the current analysis only shape similarity was calculated as the goal was to identify the folded conformations of this loop. Finally, 2690 database crystal structures were ranked in decreasing order of similarity using the Tanimoto coefficient. On the basis of a Tanimoto cutoff of 0.6 and subsequent visual inspection of the structures, 62 kinase structures with folded P-loop conformations were identified.

The 62 kinase complexes displaying the P-loop folded conformation were submitted to a series of restrained, partial energy minimizations. To compute the interaction energies after the refinement step, all residues were deleted except for the ligand and the residue at P5. The interaction energies were calculated using the Embrace procedure in Macromodel.[18] The OPLS_2005 force field was used.[19]

AUTHOR INFORMATION

Corresponding Author

*Phone: (860) 686-2915; e-mail: cristiano.guimaraes@pfizer.com.

ACKNOWLEDGMENT

The authors thank Alan Mathiowetz for helpful discussions.

REFERENCES

(1) Sciabola, S.; Stanton, R. V.; Wittkopp, S.; Wildman, S.; Moshinsky, D.; Potluri, S.; Xi, H. Predicting kinase selectivity profiles using free-Wilson QSAR analysis. *J. Chem. Inf. Model.* **2008**, *48*, 1851–1867.

(2) Luo, Y. Selectivity assessment of kinase inhibitors: Strategies and challenges. *Curr. Opin. Mol. Ther.* **2005**, *7*, 251–255.

(3) Stout, T. J.; Foster, P. G.; Matthews, D. J. High-throughput structural biology in drug discovery: Protein kinases. *Curr. Pharm. Des.* **2004**, *10*, 1069–1082.

(4) Dietricha, J.; Hulmea, C.; Hurleya, L. H. The design, synthesis, and evaluation of 8 hybrid DFG-out allosteric kinase inhibitors: A structural analysis of the binding interactions of Gleevec®, Nexavar®, and BIRB-796. *Bioorg. Med. Chem.* **2010**, *18*, 5738–5748.

(5) Noble, M. E. M; Endicott, J. A.; Johnson, L. N. Protein kinase inhibitors: Insights into drug design from structure. *Science* **2004**, *303*, 1800–1804.

(6) Probst, G. D.; Bowers, S.; Sealy, J. M.; Truonga, A. P.; Homa, R. K.; Galemmo, R. A., Jr.; Konradi, A. W.; Shama, H. L.; Quincy, D. A.; Pan, H.; Yao, N.; Lin, M.; Tóth, G.; Artis, D. R.; Zmolek, W.; Wong, K.; Qin, A.; Lorentzen, C.; Nakamura, D. F.; Quinn, K. P.; Sauer, J.-M.; Powell, K.; Ruslim, L.; Wright, S.; Chereau, D.; Ren, Z.; Anderson, J. P.; Bard, F.; Yednock, T. A.; Griswold-Prenner, I. Highly selective c-Jun N-terminal kinase (JNK) 2 and 3 inhibitors with in vitro CNS-like pharmacokinetic properties prevent neurodegeneration. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 315–319.

(7) Patel, R. Y.; Doerksen, R. J. Protein kinase-inhibitor database: Structural variability of and inhibitor interactions with the protein kinase P-loop. *J. Prot. Res.* **2010**, *9*, 4433–4442.

(8) Yao, Z.; Zhou, G.; Wang, X. S.; Brown, A.; Diener, K.; Gan, H.; Tan, T. H. A novel human STE20-related protein kinase, HGK, that specifically activates the c-Jun N-terminal kinase signaling pathway. *J. Biol. Chem.* **1999**, *274*, 2118–2225.

(9) Liang, J. J.; Wang, H.; Rashid, A.; Tan, T.-H.; Hwang, R. F.; Hamilton, S. R.; Abbruzzese, J. L.; Evans, D. B.; Wang, H. Expression of MAP4K4 is associated with worse prognosis in patients with stage II pancreatic ductal adenocarcinoma. *Clin. Cancer Res.* **2008**, *14*, 7043–7049.

(10) Tang, X.; Guilherme, A.; Chakladar, A.; Powelka, A. M.; Konda, S.; Virbasius, J. V.; Nicoloro, S. M. C.; Straubhaar, J.; Czech, M. P. An RNA interference-based screen identifies MAP4K4/NIK as a negative regulator of PPARγ, adipogenesis, and insulin-responsive hexose transport. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 2087–2092.

(11) Aouadi, M.; Tesz, G. J.; Nicoloro, S. M.; Wang, M.; Chouinard, M.; Soto, E.; Ostroff, G. R.; Czech, M. P. Orally delivered siRNA targeting macrophage Map4k4 suppresses systemic inflammation. *Nature* **2009**, *458*, 1180–1184.

(12) Bouzakri, K.; Zierath, J. R. MAP4K4 gene silencing in human skeletal muscle prevents tumor necrosis factor-α-induced insulin resistance. *J. Biol. Chem.* **2007**, *282*, 7783–7789.

(13) Graczyk, P. P. Gini coefficient: A new way to express selectivity of kinase inhibitors against a family of kinases. *J. Med. Chem.* **2007**, *50*, 5773–5779.

(14) Karaman, M. W.; Herrgard, S.; Treiber, D. K.; Gallant, P.; Atteridge, C. E.; Campbell, B. T.; Chan, K. W.; Ciceri, P.; Davis, M. I.; Edeen, P. T.; Faraoni, R.; Floyd, M.; Hunt, J. P.; Lockhart, D. J.; Milanov, Z. V.; Morrison, M. J.; Pallares, G.; Patel, H. K.; Pritchard, S.; Wodicka, L. M.; Zarrinkar, P. P. A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **2008**, *26*, 127–132.

(15) Wood, E. R.; Truesdale, A. T.; McDonald, O. B.; Yuan, D.; Hassell, A.; Dickerson, S. H.; Ellis, B.; Pennisi, C.; Horne, E.; Lackey, K.; Alligood, K. J.; Rusnak, D. W.; Gilmer, T. M.; Shewchuk, L. A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): Relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. *Cancer Res.* **2004**, *64*, 6652–6659.

(16) *OEShape Toolkit*; OpenEye Scientific Software: Santa Fe, NM.

(17) Rush, T. S., III; Grant, J. A.; Mosyak, L.; Nicholls, A. A shape-based 3-D scaffold hopping method and its application to a bacterial protein−protein interaction. *J. Med. Chem.* **2005**, *48*, 1489–1495.

(18) *MacroModel*, version 9.0; Schrödinger, LLC: New York.

(19) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. J. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.