Molecules **2010**, *15*, 4382-4400; doi:10.3390/molecules15064382

OPEN ACCESS

molecules

ISSN 1420-3049

www.mdpi.com/journal/molecules

Article

Lead Generation and Optimization Based on Protein-Ligand Complementarity

Koji Ogata [1,†], Tetsu Isomura [2,†], Shinji Kawata [3], Hiroshi Yamashita [3], Hideo Kubodera [3] and Shoshana J. Wodak [1,4,5,*]

[1] Centre for Computational Biology, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada

[2] Advanced Medical Research Laboratories, Mitsubishi Tanabe Pharma Corporation, 1000, Kamoshida-cho, Aoba-ku, Yokohama, Kanagawa 227-0033, Japan

[3] Medicinal Chemistry Research Laboratories, Mitsubishi Tanabe Pharma Corporation, 1000, Kamoshida-cho, Aoba-ku, Yokohama, Kanagawa 227-0033, Japan

[4] Department of Biochemistry, University of Toronto, Toronto, Canada

[5] Department of Molecular Genetics, University of Toronto, Toronto, Canada

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: shoshana@sickkids.ca; Tel.: +1 416 813 6351; Fax: +1 416 813 8339.

Received: 24 May 2010 / Accepted: 7 June 2010 / Published: 17 June 2010

Abstract: This work proposes a computational procedure for structure-based lead generation and optimization, which relies on the complementarity of the protein-ligand interactions. This procedure takes as input the known structure of a protein-ligand complex. Retaining the positions of the ligand heavy atoms in the protein binding site it designs structurally similar compounds considering all possible combinations of atomic species (N, C, O, CH_3, NH, *etc*). Compounds are ranked based on a score which incorporates energetic contributions evaluated using molecular mechanics force fields. This procedure was used to design new inhibitor molecules for three serine/threonine protein kinases (p38 MAP kinase, p42 MAP kinase (ERK2), and c-Jun N-terminal kinase 3 (JNK3)). For each enzyme, the calculations produce a set of potential inhibitors whose scores are in

agreement with IC50 data and Ki values. Furthermore, the native ligands for each protein target, scored within the five top-ranking compounds predicted by our method, one of the top-ranking compounds predicted to inhibit JNK3 was synthesized and his inhibitory activity confirmed against ATP hydrolysis. Our computational procedure is therefore deemed to be a useful tool for generating chemically diverse molecules active against known target proteins.

Keywords: lock-and-key problem; computational structure-based drug design; lead generation; lead optimization

1. Introduction

When a high-resolution structure of a target protein is known, computational structure-based drug design is an efficient and effective methodology for the identification and further optimization of hit compounds in order to generate lead compounds. Several studies [1-3] have reported the successful identification of hit molecules by *in silico* screening of large compound databases using software packages such as DOCK [4], GOLD [5,6] and eHiTS [7]. In many cases, however, identified hits exhibit weak biological activity and poor Adsorption, Distribution, Metabolism, Excretion and Toxicity (ADMET) properties, making them unsuitable scaffolds for further optimization. Consequently, new strategies have to be developed in order to derive molecules with better biological properties from such hits.

Standard hit compound optimization approaches involve the addition, replacement or removal of chemical groups within the hit molecule. However, enhancing the biological activity of the hit often requires a more drastic modification of the core molecular skeleton [8]. *De novo* design [9-13] and scaffold hopping techniques [14-26] are examples of methods that involve such modifications. The general assumption underlying these methods is that compounds with similar geometries will interact in a similar manner with the target protein and therefore, will show similar or improved inhibitory activity. This assumption is based on the lock-and-key model for protein-ligand interactions [27] and most of the methods are based in making changes in the native moiety of the ligand scaffolds and their geometries.

In this paper, we report the application of a computational method for structure-based ligand optimization to three serine/threonine protein kinase systems: p38 MAP kinase, p42 MAP Kinase (Erk2), and c-Jun N-terminal kinase 3 (JNK3). Our method uses as starting point, the atomic coordinates of the protein-ligand complex, determined by X-ray Crystallography, NMR, or modeling techniques [28]. A large number of compound candidates are generated by replacing the atoms of the native ligand with different substituents, keeping fixed the atomic positions of its core skeleton. Each new compound replaces the native ligand in the protein active site complementarity and the new protein-ligand complex is scored in order to rank lead candidates. The top ranking compounds are selected for further analysis. The innovation of our approach lies in the way in which the library of

compound candidates is generated as well as in the rapid identification of the chemical groups (and combinations thereof) that is likely to enhance biological activity. Our method identifies the native ligands among the top five hit compounds for the three serine/threonine Kinases analyzed here. In addition, the five top ranking compounds exhibit significant IC50 levels against ATPase activity [29,30]. In the JNK3 systems, a compound from the 10 top ranking candidates was chemically synthesized and IC50 measurements showed inhibitory activity against ATP hydrolysis. These results suggest that our method can be useful in the identification and generation of lead compounds as drug candidates.

2. Results and Discussion

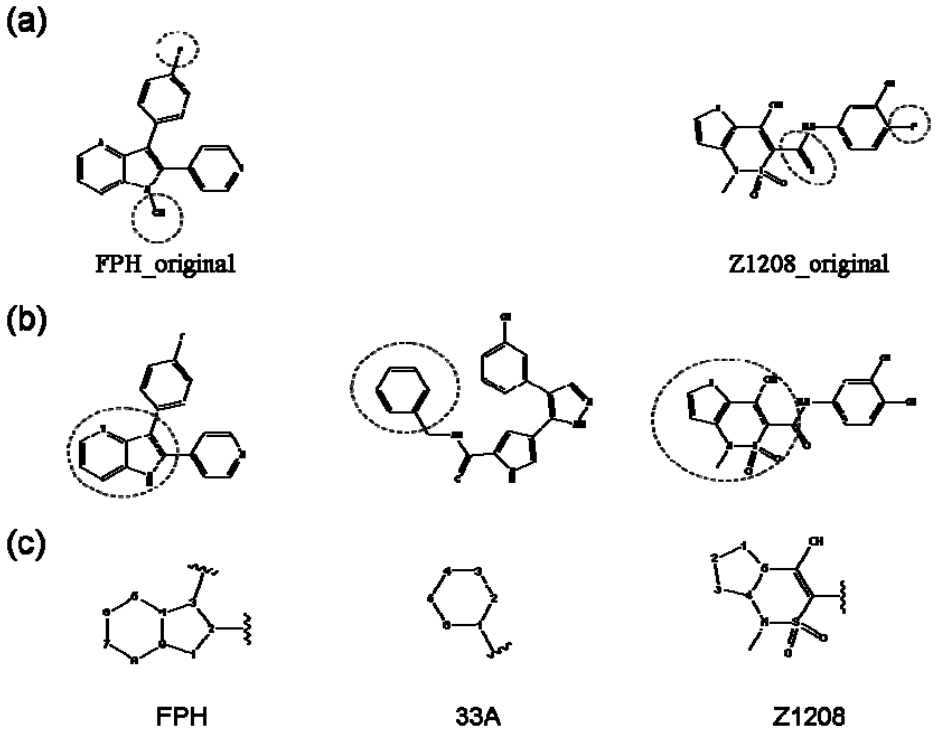
Our approach was tested using three *serine/threonine protein kinases* as targets. The X-ray protein-ligand complex structures used in this study were: p38 MAP kinase/3-(4-fluorophenyl)-2-pyridin-4-yl-1*H*-pyrrolo[3,2-b]pyridine-1-ol (**FPH**) complex, p42 MAP Kinase (Erk2)/*N*-benzyl-4-[4-(3-chlorophenyl)-1*H*-pyrazol-3-yl]-1*H*-pyrole-2-carboxamide (**33A**) complex, and c-Jun N-terminal kinase 3 (JNK3)/*N*-(3,4-dichlorophenyl)-4-hydroxy-1-methyl-2,2-dioxo-1,2-dihydro-2lamda~6~-thieno[3,2-c][1,2]thiazine-3-carboxamide (in house code **Z1208**) complex obtained from the Protein Data Bank (PDB) [31] (see Table 1). In preparing the input structures for our calculations, the FPH and Z1208 were modified as mentioned in Experimental section (Figure 1).

Table 1. Proteins and ligands used in this study.

Protein names	PDB entry	Name of ligands	PDB entry	MW
Map kinase P38	1oz1	3-(4-FLUOROPHENYL)-2-PYRIDIN-4-YL-1*H*-PYRROLO-[3,2-B]PYRIDIN-1-OL	FPH [1]	305.306
Map kinase ERK2	2oji	*N*-BENZYL-4-[4-(3-CHLOROPHENYL)-1*H*-PYRAZOL-3-YL]-1*H*-PYRROLE-2-CARBOXAMIDE	33A	376.839
c-Jun N-terminal kinase 3 (JNK3)	2ok1	*N*-BENZYL-4-[4-(3-CHLOROPHENYL)-1*H*-PYRAZOL-3-YL]-1*H*-PYRROLE-2-CARBOXAMIDE	33A	376.839
c-Jun N-terminal kinase 3 (JNK3)	NA	*N*-(3,4-DICHLOROPHENYL)-4-HYDROXY-1-METHYL-2,2-DIOXO-1,2-DIHYDRO-2LAMBDA~6~-THIENO[3,2-c][1,2] THIAZINE-3-CARBOXAMIDE	Z1208 [2]	405.278

[1] The prediction was used for a modified compound in which the OH group in pyrrolo[3,2-b]pyridine ring was replaced by a hydrogen; [2] The tertiary structure was determined by X-ray crystallography at ZOEGENE Corp.

Figure 1. Chemical structures of the native compounds and modifications made to create the input files for our calculations (see text for details). The portion surrounded by the dotted circle was replaced as indicated in the text, and position numbers are displayed for illustrative purposes.



2.1. Redesign of the FPH ligand

For the FPH-p38 MAP kinase complex, 230 candidate molecules were obtained by our calculations after the various filters have been applied. Our results show that the preferred atoms for position 5 in the 6-membered ring (see Figure 1c) were either –O– or –N=, to enable hydrogen bond formation with the Nζ of Lys53 (numbering as in the PDB entry 1oz1) (see Figure 2a). The same hydrogen bond is observed in the original structures of the p38-FPH complex [30]. In order to maintain this hydrogen bond, but also satisfy the bond orders of the 5- and 6- condensed membered rings, specific combinations of the atomic species containing single and double bonds are required. For example, if an –O– group is assigned at position 5 (see Table 2), other positions in the molecule could be occupied by either sp2 atoms (ex. compound 1, 5 and 8), or the combination of two sp3 atoms and six sp2 atoms (ex. compound 3, 4). If an –N= group is assigned at position 5, the other positions in the molecule could be occupied by one sp3 atom and seven sp2 atoms (for example, compounds 2, 6). Detailed analysis of the p38-FPH binding site, revealed that the condensed FPH ring formed hydrophobic interactions with Val38, Leu171, and the aromatic ring of Tyr35 (see Figure 2). These interactions are maintained by some of the compounds described earlier in this paragraph. In summary, our method assigned a hydrogen bond acceptor groups to position 5 and hydrophobic groups to the other positions in the condensed rings thereby satisfying binding features observed in the original crystal structures.

Table 2. The 10 highest scoring FPH substitutions for the FPH-p38 MAN kinase complex.

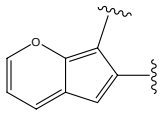
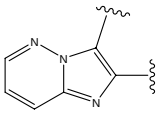
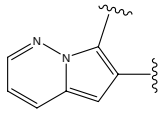
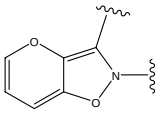
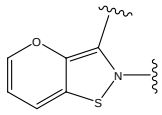
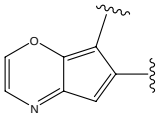
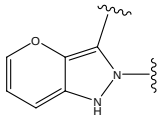
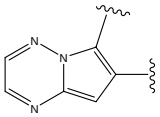
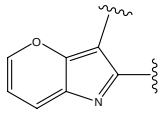
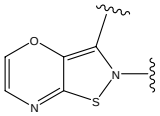
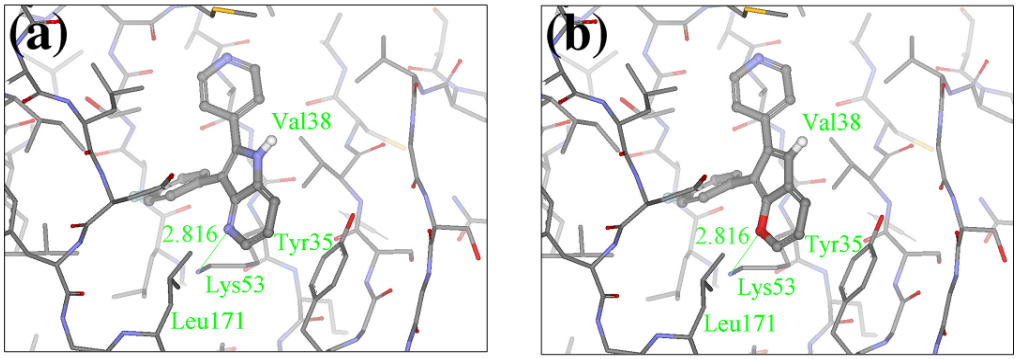
Ranking	Compounds	Score (Kcal/mol)	Ranking	Compounds	Score (Kcal/mol)
1		-38.536 6			-37.916
2		-38.51 7			-37.834
3		-38.474 8			-37.602
4		-38.036 9			-37.577
5		-37.941 10			-37.541

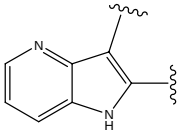
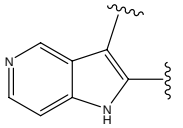
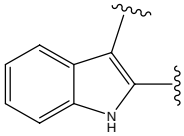
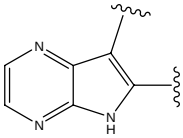
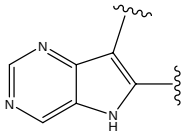
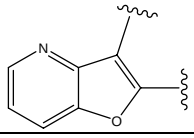
Figure 2. FPH-p38 MAN kinase complex analysis. Binding conformations of (a) native FPH, and (b) the top scoring compound in our calculations.



We find that the scores computed by our method for the four compounds 9, 22, 24, and 35, correlate well with the experimental IC50 values (see Table 3). In addition, the suggested top-scoring compound (compound 9), also features the lowest IC50 among the six compounds. In this candidate, the geometry of native FPH (see Figure 1) was modified by replacing a hydroxyl group with a hydrogen atom, as mentioned below (Section 3.2). Having introduced this modification, we might expect that the binding mode of the modified compound would differ somewhat from the native ligand. However, in the current version of our method such alternative binding modes are not considered. Without

experimental information on the binding mode of compound 9, it is therefore difficult to accurately evaluate at this point the actual predictive power of our calculations.

Table 3. Calculated scores *versus* IC_{50} values for FPH substituents.

Compounds	Substitution	Ranking	Score (Kcal/mol)	IC50 (nM)[1]
9		29 -	36.501	6.5
15		153 4	0.118	3100
19a		156 4	0.214	1800
22		56 -	35.568	53
24		66 -	34.414	895
35		34 -	36.299	86

[1] IC_{50} values are from Trejo *et al* [30].

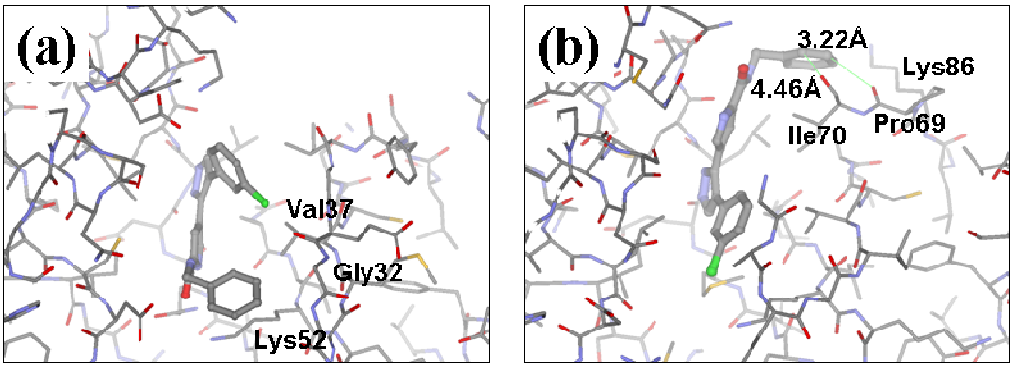
We also find that the scores of low ranking compounds tend to be inconsistent with their IC_{50} values. For example, compounds 15 and 19a, whose ranks are 153 and 156 respectively, have IC_{50} values of 3,100 nM and 1,800 nM (see Table 3). However, despite the discrepancy in these values, our calculations indicate the correct binding trend since the scores of both compounds, are positive, which is indicative of unfavorable binding energies, in agreement with their experimental IC_{50} values. Detailed analysis of the generated structures suggests that this unfavorable score can be explained by the lack of hydrogen bond capability of position 5 in these compounds, as stated above (see Figure 2).

2.2. Redesign of the 33A scaffold to optimize ERK2 binding

The ATP binding sites in ERK2 and JNK3 exhibit different chemical compositions and in particular different ratios of hydrophilic *versus* hydrophobic residues (see Figures 3a,b). Consequently, ligand 33A displays different binding orientations in ERK2 and JNK3, with the chlorobenzene moieties oriented in opposite directions. Analysis of the ERK2 complex structure revealed contacts between positions 3 and 5 of the aromatic ring of the ligand with hydrophobic groups of the protein (Cα in Gly32, Cγ2 in Val37 and Cδ in Lys52) (see Figure 3a). Hydrophobic groups were hence the preferred

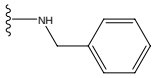
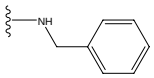
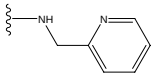
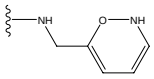
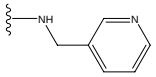
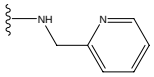
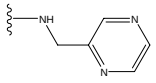
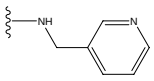
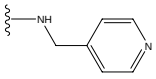
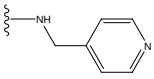
substituents for these two ring positions. Furthermore, position 4 of the aromatic ring showed interactions with the carbonyl oxygen of Ala33, suggesting hydrogen bond donor groups as preferred subtituents for this position.

Figure 3. Structures of the complexes of 33A with (a) ERK2 and (b) JNK3, respectively. Both the proteins and the ligand are displayed using stick models, with the ligand shown using thicker bonds. Comparison of panels (a) and (b) illustrate the difference in orientation of 33A in the two structurally aligned proteins.



Our calculations yielded 23 different substituents for the A33 ring scaffold (Figure 1c), with benzene moiety having the top score for the ligands-ERK2 complexes (see Table 4).

Table 4. Five best scoring substitutions for the complexes ERK2 and JNK3 kinases with A33.

ERK2				JNK3			
Ranking	**Substitutions**	**Score**	**Ki, uM[1]**	**Ranking**	**Substitutions**	**Score**	**Ki, uM[1]**
1		-28.016 0	.086	1		-43.488 0.	55
2		-27.643 ---		2		-42.738 ---	
3		-27.342 0.23		3		-42.557 ---	
4		-26.969 ---		4		-42.117 N	D[2]
5		-25.62 0.16		5		-41.903 N	D[2]

[1] Ki values are from the article by Aronov *et al*. [29]; [2] ND (Not determined). Authors reported that compound 4 was 3-fold less active than 1 and compound 5 was 2-fold less active.

These results are in agreement with the inhibitory activity (lowest K_i value) previously reported [29] and can be explained by the hydrophobic interactions between the aromatic ring and the active site protein residues. The substitution of a –CH residue by –N= decreases the hydrophobic interactions and may explain the lower score values of the other designed compounds. The exception to this rule is the 5th ranking compound where the addition of –N= residue to position 4 in the aromatic ring increased the repulsive energy (decreasing the overall score) due to the proximity of this substituent to the carbonyl oxygen of the residue Ala33. This effect is enhanced by the fact that our software is using a fixed geometry approximation. Two approaches are currently being developed to improve this methodology: (a) consideration of different compound conformations, and (b) relaxation of the protein-ligand complex in order to relieve any residual strain.

2.3. Redesign of the 33A scaffold to optimize JNK3 binding

Analysis of the 3D-structure of the 33A-JNK3 complex revealed that the ring scaffold to be modified interacts with Lys68, Pro69 and Ile70 (see Figure 3b). The ring positions 2 and 3 are close to the carbonyl oxygen atoms of the protein residue Pro69 and Ile70, while the rest of the ring atoms are surrounded by hydrophobic residues. Therefore, positions 2 and 3 were assigned positively charged substituents, whereas hydrophobic substituents were assigned for the remaining ring positions.

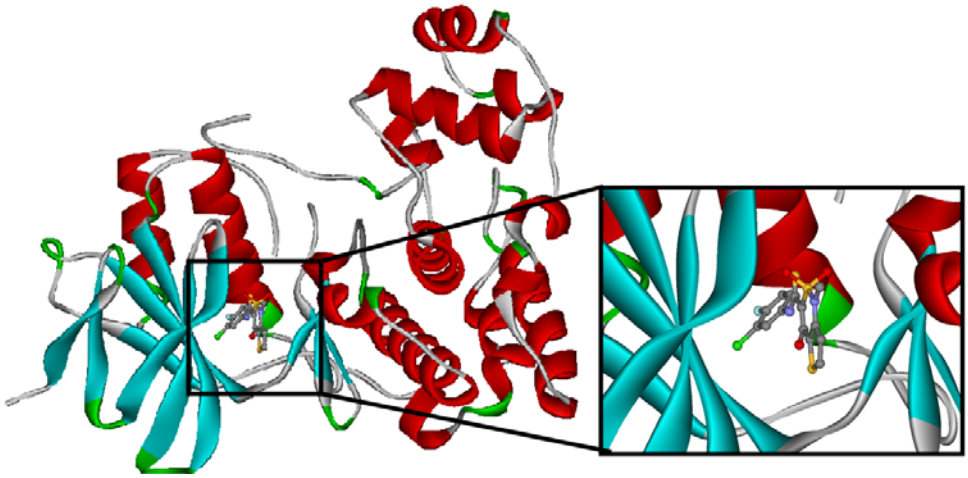
Subject to these constraints, our procedure sampled a total of 23 substitutions for the 6-membered 33A ring (Figure 1c). Here too, our results revealed benzene moiety to be the best ranking compound in agreement with the experimental data[29], with a similar rationale as for the ERK2. Our second "best compound" contains an –O– at ring position 2 and an –NH – group at ring position 3. The somewhat lower score of this compound is due to close polar contact with backbone atoms of the protein. This score is driven by the stabilizing energy from the proximity between ligand –NH- to the carbonyl oxygen O of Pro69 (3.23 Å), and the repulsive energy for the interaction between the ligand –O– at position 2 with the carbonyl oxygen of Ile70 (4.46 Å).

The compounds with 2-, 3- and 4-pyridine moieties were ranked at 3rd, 4th and 5th positions, respectively. With respect to the calculated "best compound", these scores can be rationalized in the same way as in the case of ERK2 case, by a decrease in hydrophobic interactions. However, the authors of the experimental paper [29] found that compounds with 3- and 4-pyridine moieties do not bind to JNK3, which underscores the difficulty in discriminating between compounds that bind from those that do not bind, based on our computed scores alone.

2.4. Redesign of the Z1208 scaffold bound to JNK3.

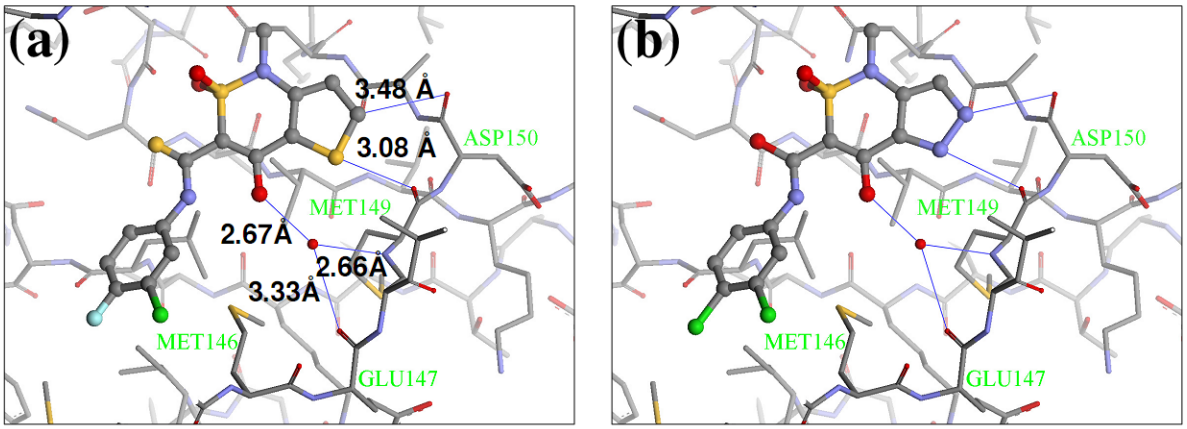
The crystal structure of the native Z1208 with JNK3 complex suggests that the ligand is tightly bound to the ATP binding pocket of the protein (Figure 4).

Figure 4. Binding site of the native Z1208 JNK3 complex.



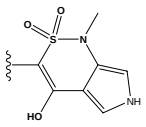
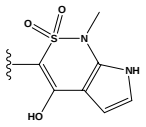
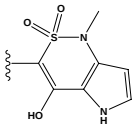
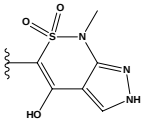
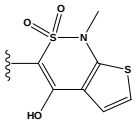
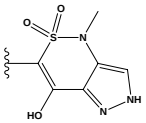
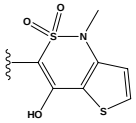
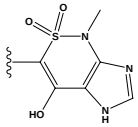
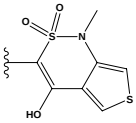
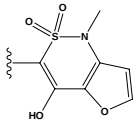
One water molecule appears in the binding site forming hydrogen bonds with the protein backbone (carbonyl oxygen of the Glu147 and the amide nitrogen of Met149) and the hydroxyl group of the condensed ring in the native Z1208 (Figure 5a). Analysis of JNK3-ligand complexes in the PDB show that hydrogen bond interactions between the ligand and the backbone atoms of Glu147 and/or Met149 are common, but the water molecule is not present in all the structures. Therefore for the purpose of this study the water molecule was removed. Furthermore, in the binding site, the 3-chloro-4-fluoro-phenyl moiety of the native Z1208 ligand is positioned near Met146 of JNK3 and surrounded by other hydrophobic residues (Figure 4). According to Scapin *et al.* [32], JNK3 complexes with imidazole-pyrimidines (PDB-codes: 1qmn and 1qmq) that have halogen-phenyl moieties, feature different conformations of the Met146 side chain than in the complex with AMP-PNP (PDB-code 1jnk). We found that the structure of JNK3 protein in the complex with Z1208_original, superimposed well onto those of the JNK3 proteins in 1qmn and 1qmq. In addition the Met146 adopts a similar conformation on all three complexes.

Figure 5. Conformations of (a) native Z1208 and (b) Z1208_8 binding to JNK3.



Our calculations generated 64 designed compounds. The top 10 ranking compounds are displayed in Table 5. Compounds with the best and the second best scores contain one –NH– group in different positions of the 5-membered ring. The addition of a second substituent in the same ring of the molecule (–NH– and –N=), produced only a slightly higher score (compound 8 in Table 5). Visual inspection of the modeled ligand-protein active site reveals that oxygen atoms in Met149 and Asp150 make contacts in the 5-membered ring, and those atoms are engaged in repulsive interactions when the negatively charged –N= groups are introduced (Figure 5a). Hence replacements which combine sulfur or carbon atoms with the –NH– group is preferable to those of two –N= groups.

Table 5. Ten best scoring substitutions of Z1208 bund to JNK3 and their IC50 values.

Ranking	Substitutions	Score	IC50 (uM)	Ranking	Substitutions	Score	IC50 (uM)
1		-37.376 N	A	6		-35.815 N	A
2		-36.783 N	A	7		-35.783 N	A
3		-36.6 N	A	8		-35.569	
4		-36.285		9		-35.189 N	A
5		-36.145 N	A	10		-35.175 N	A

We see that the native compound is ranked 4th in this list, but its score is only slightly worse than the predicted best-scoring ligand. Unfortunately, the chemical synthesis of all three best ranking compounds was very difficult with current techniques, demonstrating that the filters used in the pre-calculation steps were not sufficient to remove all the unsynthesizable compounds and hence further improvements of the various filters used in our procedure are required. Nevertheless, it is quite encouraging that the native ligand was the top ranking synthesizable compound.

2.5. Discussion

Additional validation of our method by synthesis of one molecule

An additional validation method step was carried out by synthesizing in the laboratory one of the top-ranking newly designed inhibitors. The chosen compound was Z1208-8, ranked 8th in Table 5, and its inhibitory activity was measured against ATP hydrolysis by JNK3. This compounds was selected because it was easier to synthesize compared to other higher ranked compounds in the same list.

The biological assay showed that Z1208-8 was able to inhibit ATP binding to JNK3 with an IC_{50} 62.9 μM, representing a six fold poorer activity than the native ligand (Z1208, 9.6 μM). The analysis of the 3D structure of the X1208-8/JNK3 complex indicated that the two major contributions to the overall energy score are: 1) the NH group in the 5-membered ring forming a hydrogen bond with the carbonyl oxygen of Asp150, which would favor binding, and 2) a close contact of the –N= group in the 5-membered ring with the backbone oxygen of Met149 (Figure 5b), which would disfavor binding. The second contribution is higher in Z1208-8 higher than in the native compounds, suggesting an explanation to the higher IC_{50} value observed in the new inhibitor.

Our method is thus capable of designing new compounds with inhibitor activity against the target enzyme, but their activity can be lower than that of the starting hit compound. We expect however, that several cycles of rational design using our method and experimental analysis of the top ranking candidates should lead not only to compounds with different chemical properties from those of the original molecule, but also to those with a strong or stronger inhibitory activity.

This paper has shown that our computational method is effective in a scaffold-based redesign of kinase inhibitors FPH, 33A and Z1208. For a defined scaffold and keeping fixed the "geometry" of its core skeleton, our method was capable of sampling a large number of different substituents providing a set of compounds with potential inhibitory activity against the protein targets. Compounds were ranked based on an energy function, and in all the cases native inhibitors were identified in the 5 top ranking compounds. Validation of our method was performed by comparing the scores of designed molecules with empirical data on their inhibitory activity (IC50 and Ki values). In the JNK3 inhibitor design study, one of the top ranked compounds was synthesized and its inhibitory activity was confirmed experimentally.

Future developments will address the following outstanding limitations of our method: 1) more effective filters to remove compounds difficult or impossible to synthesize, 2) improve the scoring function to enhance compound ranking accuracy, and 3) take into account protein and ligand conformational flexibility and different ligand poses in the protein active site. With the introduction of these improvements, our computational approach holds the promise of becoming a useful tool for lead optimization.

3. Experimental

3.1. Lead optimization procedure

The lead/hit optimization procedure used in this study was previously reported by Ogata *et al* [33] and is only briefly summarized here. The first step consists in extracting the atomic coordinates of the ligand's heavy atoms (referred to as 'geometry') from a high resolution structure of the protein-ligand complex. The geometry is then divided into fragments which are grouped into three partial structures types: *rings*, *linkers* (defined as the fragments that connect rings), and *terminals* (defined as other types of fragments). In addition, all the atoms in the geometry are classified according to their bond order types (sp3, sp2, etc) and atomic species (CH_3, CH_2, CH, NH_2, NH etc). For example, consider the geometry X···Y··Z, in which X, Y, and Z represent the atoms in the geometry and '···' is a generic representation of the bonds connecting the atoms. Replacing Y with a "=CH–" generates a chemically incomplete compound X=CH–Z for which "=" and "–" indicate a double and a single bond, respectively. Then, X should be assigned to an atomic species linked through a double bond to Y (ex. O= or CH2=). Similarly, Z should be assigned to an atomic species capable of linking to Y through a single bond (ex. $–CH_3$ or $–NH_2$). By assembling all possible combinations of these atomic species, four compounds are obtained: $O=CH–CH_3$, $O=CH–NH_2$, $CH_2=CH–CH_3$, and $CH_2=CH–NH_2$. For the work presented in this paper, eighteen atomic species were used (see Table 6).

Table 6. Atomic Chemotypes used in this study.

Atom groups	Bond type	No. of bonds	No. of hydrogens	Atom groups	Bond type	No. of bonds	No. of hydrogens
s—CH_3	p3	4	3	—OH	sp3	2	1
s—CH_2—	p3	4	2	—O—	all	2	0
—CH(—)—	sp3	4	1	=O	all	2	0
—C(—)—	sp3	4	0				
s=CH_2	p2	4	2	—Cl	sp3	1	0
s=CH—	p2	4	1				
=C<	sp2	4	0				
s —NH_2	p3	3	2				
—NH—	all	3	1	—S—	all	2	0
=N—	sp2	3	0	=S<	all	4	0
N(—)<	all	3	0	S(=)(=)<	all	6	0

After assigning all possible combinations of atomic species to the native ligand's core coordinates, all the partial structures are considered and bond order requirements are satisfied thereby generating the hit compound database. Compounds in the newly generated database have similar core geometries as the native ligand and each atomic position satisfies different chemically meaningful combinations.

In a second step, compounds from the newly generated database are subjected to two filters. The Rishton non leadlikeness filter (to remove undesirable functional groups, see Figure 6) [34] and the Lipinski's rule of five (compounds with more than five hydrogen-bond donors, more than 10 hydrogen-bond acceptors, molecular mass greater than 500 Da, logP values greater than 5, or more than 10 rotatable bonds are not desirable for orally active drugs) [35]. From the remaining list of compounds, molecules with ring(s) and condensed ring structures were selected because known hits for the three target kinase proteins contain such structures. These molecules were treated as the final list of lead candidates and ranked based on a scoring function, *Score,* evaluated for the protein-ligand complex. The scoring function comprises four empirical energy terms:

$$Score = E_v + E_e + E_h + E_s$$

where, E_v is the van deer Waals interaction energy, E_e the electrostatic interaction energy, E_h the hydrogen bond energy, and E_s the solvation energy.

E_v and E_e were obtained using the AMBER force field with the GAFF parameter set. [36] E_h was defined as:

$$E_b(r_{ij}) = \begin{cases} \exp(-(r_{ij} - r_H)^2) & 1.0 \le r_{ij} \le 3.0 \\ 0 & else \end{cases},$$

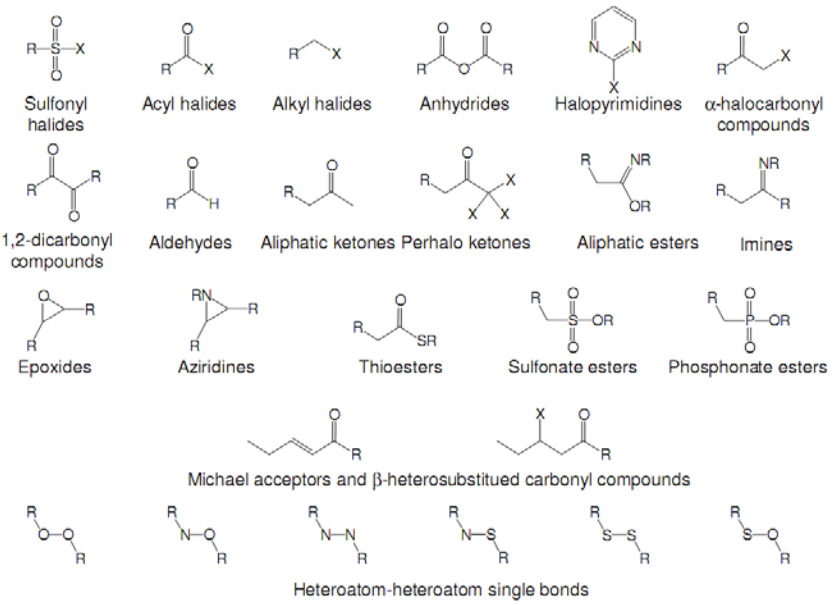
where r_H is the distance between the hydrogen and the heavy atom (H···X, set at 2.0 Å for this study).

E_s was computed for the bound and unbound states of the ligand, protein and the complex as:

$$E_s = \sum_i \sigma_i A_i ,$$

where A_i and σ_i are the solvent-accessible surface area and the proportionality factor for the solvent-accessible surface area of atom *i*, respectively[37]. The free energy of $G^{protein}$ and $G^{compound}$ were calculated in the same manner. This method has been designed to provide a ranked list of compounds with better "drug-type" properties (more stable, druglikeness and synthesizable compounds) than other approaches.

Figure 6. Nonleadlikness filter. The substituent types were extracted from Rishton work [34]. The electrophilic functional groups shown here are the most common protein-reactive covalent-acting false positives in biochemical assays. Compounds with substituents shown in this figure were removed from our results.



3.2. Application to serine/threonine protein kinases

Our approach was tested using three *serine/threonine protein kinases* as targets. The X-ray protein-ligand complex structures used in this study were: p38 MAP kinase/ 3-(4-fluorophenyl)-2-pyridin-4-yl-1*H*-pyrrolo[3,2-b]pyridine-1-ol (**FPH**) complex, p42 MAP Kinase (Erk2)/*N*-benzyl-4-[4-(3-chlorophenyl)-1*H*-pyrazol-3-yl]-1*H*-pyrole-2-carboxamide (**33A**) complex, and c-Jun N-terminal kinase 3 (JNK3)/*N*-(3,4-dichlorophenyl)-4-hydroxy-1-methyl-2,2-dioxo-1,2-dihydro-2lamda~6~-thieno[3,2-c][1,2]thiazine-3-carboxamide (in house code **Z1208**) complex obtained from the Protein Data Bank (PDB) [31] (see Table 2). The three structures display different ligand binding modes, and featured differences in the electrostatic potentials at the ATP-binding site [29,30,38-40]. In addition, inhibitory activity against the target proteins has been reported for series of compounds. These compounds were derived by small modifications (changing or adding substituents) of the native ligand structures and atom types [29,30,38]. We used this data to validate the results of our calculations, which involves potential compound candidates with larger structural and chemical differences than the original authors considered in their study.

In preparing the input structures for our calculations, the following steps were performed (see Figure 1): 1) all water molecules were removed from the original complexes' PDB files; 2) In FPH, the hydroxyl group attached to the 5- and 6-membered condensed ring was replaced by a hydrogen atom because the modified compound has a larger number of similar compounds with experimentally

demonstrated inhibitory activity than the native compound; 3) In FPH and Z1208, all the fluorine atoms were replaced by chlorine atoms as this replacement made the chemical synthesis easier and; 4) the thioamide group in Z1208 was replaced by an amide group for the same reason.

In JNK3, two ligands were used to create the input structure: the native Z1208 and a derivative of Z1208 that acts as an ATP hydrolysis inhibitor. The experimental data used for this analysis were the in-house X-ray crystal structure (2.1 Å resolution and R-factor= 24.6%, see Figure 4) and the IC50 values of 22.8 μM and 9.6 μM for native Z1208 and Z1208-derived ATP hydrolysis inhibitor respectively.

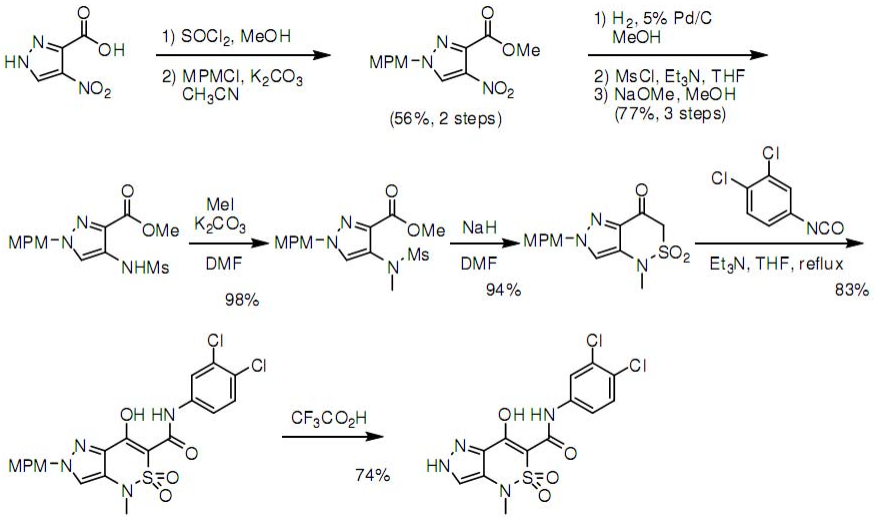
3.3. Inhibition assay

To measure the inhibitory activity of the Z1208, we used the following assay system: adenosine triphosphate (ATP), phosphoenolpyruvate (PEP), nicotinamide adenine dinucleotide (NADH), and a solution mixture of pyruvate kinase and L-lactate dehydrogenase (PK-L-LDH) were purchased from Roche Diagnostics. Other reagents were purchased from Sigma-Aldrich. JNK3 was expressed and purified by the method of Xie *et al.* [41]. After a purification step, JNK3 was activated by GST-fused MKK7 and further purified against a glutathione-fixed column. Inhibitory activity was estimated by detecting the inhibition of ATP hydrolysis reaction monitored by the coupled reaction of NADH oxidation; a slightly modified method of Xie *et al.* Experimental conditions were: 100 nM JNK3 in 50 mM Hepes, pH 7.6, 10 mM $MgCl_2$, 1 mM NADH, 90 mg/mL PK, 30 mg/mL L-LDH, 2 mM PEP, 200 mM ATP, and each concentration of compound under 1% DMSO. The conversion of NADH was measured by kinetic monitoring with SpectraMax 190 (Molecular Devices).

3.4. Synthesis of Z1208-8

Z1208-8 synthetic path is displayed in Scheme 1 [42-45].

Scheme 1. Steps of the chemical synthesis of Z1208-8.



Acknowledgements

We thank H. Kitano, F. Ohno, H. Nakano and H. Kano of Mitsubishi Tanabe Pharma Corporation, S. Kondo and S. Sugio of MCRC Corporation, and J. Gonzalez-Outeirino and V. Mallikarjunaradhya of Centre for Computational Biology in The Hospital for Sick Children for scientific discussions. We furthermore thank R. Nakanishi, C. Okumura, and S. Shimomura of Mitsubishi Tanabe Pharma Corporation for technical assistance. We are grateful to the staff of the Centre for Computational Biology for help with the High Performance computer systems. SJW is Canada Research Chair in Bioinformatics and Computational Biology, and acknowledges the Canada Fund for Innovation and the Canada Institute for Health Research for support.

References

1. Yamazaki, K.; Kusunose, N.; Fujita, K.; Sato, H.; Asano, S.; Dan, A.; Kanaoka, M. Identification of phosphodiesterase-1 and 5 dual inhibitors by a ligand-based virtual screening optimized for lead evolution. *Bioorg. Med. Chem. Lett.* **2006,** *16*, 1371-1379.
2. Yan, S.; Appleby, T.; Larson, G.; Wu, J.Z.; Hamatake, R. K.; Hong, Z.; Yao, N. Thiazolone-acylsulfonamides as novel HCV NS5B polymerase allosteric inhibitors: convergence of structure-based drug design and X-ray crystallographic study. *Bioorg. Med. Chem. Lett*. **2007,** *17*, 1991-1995.
3. Carosati, E.; Mannhold, R.; Wahl, P.; Hansen, J.B.; Fremming, T.; Zamora, I.; Cianchetta, G.; Baroni, M. Virtual screening for novel openers of pancreatic K(ATP) channels. *J. Med. Chem.* **2007,** *50*, 2117-2126.
4. Meng, E.C.; Gschwend, D.A.; Blaney, J.M.; Kuntz, I.D. Orientational sampling and rigid-body minimization in molecular docking. *Proteins* **1993,** *17*, 266-278.
5. Jones, G.; Willett, P. Docking small-molecule ligands into active sites. *Curr. Opin. Biotechnol.* **1995,** *6*, 652-656.
6. Jones, G.; Willett, P.; Glen, R.C.; Leach, A.R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997,** *267*, 727-748.
7. Zsoldos, Z.; Reid, D.; Simon, A.; Sadjad, S.B.; Johnson, A.P. eHiTS: A new fast, exhaustive flexible ligand docking system. *J. Mol. Graph. Model.* **2007,** *26*, 198-212.
8. Zhao, H. Scaffold selection and scaffold hopping in lead generation: a medicinal chemistry perspective. *Drug Discov. Today* **2007,** *12*, 149-155.
9. Schnecke, V.; Bostrom, J. Computational chemistry-driven decision making in lead generation. *Drug Discov. Today* **2006,** *11*, 43-50.
10. Bohm, H.J. LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. *J. Comput. Aided Mol. Des.* **1992,** *6*, 593-606.
11. DeWitte, R.S.; Ishchenko, A.V.; Shakhnovich, E.I. SMoG: de Novo design method based on simple, fast, and accurate free energy estimates. 2. Case studies in molecular design. *J. Am. Chem. Soc.* **1997,** *119*, 4608-4617.

12. Gehlhaar, D.K .; Moerder, K.E.; Zichi, D.; Sh erman, C.J.; Ogd en, R.C.; Freer, S.T. De no vo design of enzyme inhibitors by Monte Carlo ligand generation. *J. Med. Chem.* **1995,** *38*, 466-472.
13. Nishibata, Y.; Itai, A. Confirmation of usefulness of a structur e const ruction program based o n three-dimensional recept or structure for rational lea d gene ration. *J. Med. C hem.* **1993,** *36*, 2921-2928.
14. Ahlstrom, M .M.; Ridd erstrom, M. ; Lu thman, K.; Zam ora, I. Virtu al scr eening and s caffold hopping based on GRID molecular interaction fields. *J. Chem. Inf. Model* **2005,** *45*, 1313-1323.
15. Bergmann, R.; L inusson, A .; Zamora, I. S HOP: Scaffold HOPping by GRID -Based S imilarity Searches. *J. Med. Chem.* **2007,** *50*, 2708-2717.
16. Schneider, G.; N eidhart, W. ; G iller, T.; Sch mid, G. "S caffold-Hopping" b y Topological Pharmacophore Search: A Contribution to Virtual Screening. *Angew. Chem. Int. Ed. Engl.* **1999,** *38*, 2894-2896.
17. Abolmaali, S.F.; Ostermann, C.; Zell, A. The Compressed Feature Matrix--a novel descriptor for adaptive similarity search. *J. Mol. Model.* **2003,** *9*, 66-75.
18. Renner, S .; Schneide r, G. S caffold-hopping potential o f l igand-based si milarity concept s. *ChemMedChem* **2006,** *1*, 181-185.
19. Barker, E.J.; Buttar, D.; Cosgrove, D.A.; Gardiner, E.J.; Kitts, P.; Willett, P.; Gillet, V.J. Scaffold hopping using clique detection applied to reduced graphs. *J. Chem. Inf. Model.* **2006,** *46*, 503-511.
20. Naerum, L.; Norskov-La uritsen, L. ; Olesen, P.H. Scaffold hopping and optim ization t owards libraries of glycogen synthase kinase-3 inhibitors. *Bioorg. Med. Chem. Lett.* **2002,** *12*, 1525-1528.
21. Lloyd, D.G.; Buenemann, C.L.; Todorov, N.P.; Manallack, D.T.; Dean, P.M. Scaffold hopping in de novo design. Ligand generation in the absence of receptor information. *J. Med. Chem.* **2004,** *47*, 493-496.
22. Nair, P .C.; Sobhia, M.E. Fingerprint Direct ed Scaffold Hopping fo r Identificati on of CCR2 Antagonists. *J. Chem. Inf. Model.* **2008,** *48*, 1891-1902.
23. Mason, J.S.; Morize, I.; Menard, P.R.; Cheney, D.L.; Hulme, C.; Labaudiniere, R.F. New 4-point pharmacophore m ethod for m olecular similarity and di versity applications: ov erview of the method and appli cations, including a novel ap proach to the de sign of co mbinatorial libraries containing privileged substructures. *J. Med. Chem.* **1999,** *42*, 3251-64.
24. Andrews, K.M.; Cramer, R.D. Toward general methods of targeted library design: topomer shape similarity searching with diverse structures as queries. *J. Med. Chem.* **2000,** *43*, 1723-1740.
25. Jenkins, J.L.; Glick, M.; D avies, J.W. A 3D similarity method for sca ffold hopping from known drugs or natural ligands to new chemotypes. *J. Med. Chem.* **2004,** *47*, 6144-6159.
26. Bohl, M. ; D unbar, J.; Gifford, E. M.; Heritag e, T.; W ild, D .J.; Willett, P .; Wilton, D .J. Scaffold Searching: Automated Identification of Similar Ring Sy stems for the Design of Combinatorial Libraries. *Quant. Struct.-Act. Relat.* **2002,** *21*, 590-597.
27. Jorgensen, W.L. Rusting of the lock and key model for protein-ligand binding. *Science* **1991,** *254*, 954-955.
28. Grant, M.A . Protein str ucture pred iction in struct ure-based ligand d esign and virt ual screening. *Comb. Chem. High Throughput Screening* **2009,** *12*, 940-960.

29. Aronov, A.M.; Baker, C.; Bemis, G.W.; Cao, J.; Chen, G.; Ford, P.J.; Germann, U.A.; Green, J.; Hale, M.R.; Jacobs, M.; Janetka, J.W.; Maltais, F.; Martinez-Botella, G.; Namchuk, M.N.; Straub, J.; Tang, Q .; X ie, X. Flipped out: structure-guided design of selectiv e p yrazolylpyrrole ERK inhibitors. *J. Med. Chem.* **2007,** *50*, 1280-1287.
30. Trejo, A .; Arzeno, H.; Browner, M.; Ch anda, S.; Che ng, S.; Comer, D. D.; Dalrymple, S .A.; Dunten, P.; Lafargu e, J.; Lovejoy, B.; Freire-Moar, J.; Lim, J.; McIntosh, J.; M iller, J .; Papp, E.; Reuter, D.; Roberts, R.; Sanpablo, F.; Saunders, J.; Song, K.; Villasenor, A.; Warren, S.D.; Welch, M.; Weller, P.; Whiteley, P.E.; Zeng, L.; Goldstein, D.M. Design and synthesis of 4-azaindoles as inhibitors of p38 MAP kinase. *J. Med. Chem.* **2003,** *46*, 4702-13.
31. Berman, H.M.; Westbrook, J.; F eng, Z.; Gilli land, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000,** *28*, 235-42.
32. Scapin, G.; Pate l, S.B. ; Lisnock, J .; B ecker, J .W.; LoGrasso, P.V . The structure of JNK3 in complex with small molecule inh ibitors: structural basis for p otency and s electivity. *Chem. Biol.* **2003,** *10*, 705-712.
33. Ogata, K.; Isomura, T.; Yamashita, H.; Kubodera, H., A Quantitative Approach to the Estimation of Chemical Space from a Given Geometry by the Combination of Atomic Species. *QSAR Comb. Sci.* **2007,** *26*, 596-607.
34. Rishton, G.M. No nleadlikeness and le adlikeness in b iochemical scr eening. *Drug Discov. Today* **2003,** *8*, 86-96.
35. Lipinski, C.A. Drug-like properties and the causes of poor solubili ty and poor per meability. *J. Pharmacol. Toxicol. Methods* **2000,** *44*, 235-249.
36. Wang, J.; Wolf, R.M.; Caldw ell, J.W.; Koll man, P.A.; Case, D.A. Development and testing of a general amber force field. *J. Comput .Chem.* **2004,** *25*, 1157-1174.
37. Ooi, T.; Oobatake, M.; Nemethy, G.; Scheraga, H.A. Accessible surface areas as a measure of the thermodynamic par ameters of hydration of pe ptides. *Proc. Natl. Acad. Sci. USA* **1987,** *84*, 3086-3090.
38. Peifer, C. ; Kinkel, K.; Abadleh, M .; S chollmeyer, D.; Laufer, S . From f ive- to six- membered rings: 3,4-diarylquinolinone as lead for novel p38MAP kinase inhibitors. *J. Med. Chem.* **2007,** *50*, 1213-21.
39. Kulkarni, R.G.; Sr ivani, P .; Achaiah, G.; Sastry , G.N. Strate gies to design py razolyl ur ea derivatives for p38 ki nase inh ibition: a molecular modeling study. *J. Comput. Ai ded Mol. Des.* **2007,** *21*, 155-66.
40. Gaillard, P.; Jeanclaude-Etter, I.; A rdissone, V.; Arkinstall, S.; Cambet, Y.; Camps, M.; Chabert, C.; Church, D.; Cirillo, R.; G retener, D.; Halazy, S.; Nichols, A.; Szyndralewiez, C.; Vitte, P.A.; Gotteland, J.P. Design and synthesis of the first generation of novel potent, selective, and *in vivo* active (ben zothiazol-2-yl)acetonitrile inhi bitors of the c-Jun N-ter minal kin ase. *J. M ed. Chem.* **2005,** *48*, 4596-4607.
41. Xie, X.; Gu, Y.; Fox, T.; Coll, J.T.; Fleming, M.A.; Markland, W.; Caron, P.R.; Wilson, K.P.; Su, M.S. Cry stal st ructure o f JNK 3: a kinase i mplicated in n euronal apoptosis. *Structure* **1998,** *6*, 983-991.

42. Lombardino, J.G. Preparation of Some 4-Hydroxyl-1-Methyl-1h-2.1-Benzothiazine-3-Carboxanilide 2,2-Dioxides. *J. Heterocycl. Chem.* **1972,** *9*, 315-317.
43. Coppo, F.T.; Fawzi, M.M. Novel heterocycles. Synthesis of 2,3-dihydro-6-methyl-2-phenyl-4H,6H-pyrano[3,2-c][2,1]benzothiazin-4-one 5,5-dioxide and related compounds. *J. Heterocycl. Chem.* **1998,** *35*, 983-987.
44. Coppo, F. T.; Fawzi, M.M. Synthesis of 1-methyl-7-(trifluoromethyl)-1H-pyrido[2,3-c][1,2]thiazin-4(3H)-one 2,2-dioxide. *J.Heterocycl. Chem.* **1998,** *35*, 499-501.
45. Coppola, G.M.; Hardtmann, G.E. Novel Heterocycles .4. Synthesis of the Pyrido[2,3-C]-1,2-Thiazine Ring-System. *J. Heterocycl.Chem.* **1979,** *16*, 1361-1363.

Sample Availability: Samples of the compounds are available from the authors.