

# Sequence, Structure, and Active Site Analyses of p38 MAP Kinase: Exploiting DFG-out Conformation as a Strategy to Design New Type II Leads

Preethi Badrinarayanan and G. Narahari Sastry\*

Molecular Modeling Group, Organic Chemical Sciences, Indian Institute of Chemical Technology,  
Tarnaka, Hyderabad- 500 607, India

Received September 1, 2010

A new knowledge, structure, and sequence based strategy involving the effective exploitation of the DFG-out conformation is delineated. A comprehensive analysis of the structure, sequence, cocrystals, and active sites of p38 MAP kinase crystal structures present in Protein Data Bank (PDB) and the FDA approved MAP kinase drugs has been done, and the information is used for the design of type II leads. The 98 crystal structures, 138 cocrystals, and 31 FDA drugs comprise of 7 different sequences of 2 organisms viz., *Homo sapiens* and *Mus musculus* differing in sequence length, constituting both homo- and heterochains. Multiple sequence alignment with ClustalW showed >95% sequence similarity with highly conserved domains and a high propensity for mutations in the activation loop. The bound ligands were extracted, and their interactions with DFG in and out conformations were studied. These cocrystals and FDA drugs were fragmented on the basis of their binding interactions and their affinity to ATP and allosteric sites. The fragment library thus generated contains 106 fragments with overlapping drug fragments. A blue print constituting three main parts viz., head (ATP region), linker (DFG region), and tail (allosteric region) has thus been formulated and used to design 64 type II p38 MAP kinase inhibitors. The above strategy has been employed to design potent type II p38 MAP kinase inhibitors, which are shown to be very promising.

## 1. INTRODUCTION

In recent years, kinases have emerged as preferred drug targets triggering the heightened activity for kinase inhibitor drugs. The 518 kinases encoded in the human genome play pivotal roles as popular and most haunted drug targets due to their involvement in signal transduction pathways, which are wired through a phosphotransfer cascade thereby eliciting a multitude of physiological response.<sup>1</sup> Protein kinases share a similar tertiary structure, ATP binding site, and catalyze analogous reactions, i.e., the phosphorylation of proteins. Nevertheless, each kinase has unique structural details, for example, protein–protein interaction segments, specific dynamic properties, and allosteric binding site. This uniqueness of a protein kinase can be exploited pharmaceutically to create specificity.<sup>2–7</sup>

The type I kinase inhibitors that target the ATP binding site of kinase in its active conformation are usually effective, but their low specificity and selectivity for the desired target due to high similarity in the ATP binding site of kinases make them unsuccessful in many instances. Therefore it is imperative to explore the structural requirements to achieve selectivity. One promising approach is to design type II “dual site binders” which target both ATP and allosteric sites. The allosteric site is the hydrophobic pocket adjacent to the ATP binding site, accessible only in the inactive conformation of kinase. This region being less similar promises increased selectivity and specificity. ATP site of the kinase is revealed when it is in the active DFG-in conformation, a conformation otherwise conductive to phosphotransfer, whereas in the

inactive form, i.e., in the DFG-out conformation, a hydrophobic pocket is opened up due to the conformational rearrangement of the activation loop.<sup>8</sup> The functional state of kinases including MAP kinases is characterized by the conformation assumed by the DFG-loop which is present in most kinases. The transition of the DFG-loop is triggered by the shift of the F169 side chain by 10 Å from its usual buried location (DFG-in conformation) to a location that sterically interferes with ATP binding (DFG-out conformation). This transition is a unique feature of the kinase family. The former conformation can be ascribed to both activated and nonactivated kinases, whereas the latter is found only as a nonactive form. Therefore, analysis of sequence and structure of the human kinase is of outstanding interest to identify differences in the sequence and the structure between the inactive and active kinase conformations. The targeted inhibition of protein kinases thus represents a key therapeutic strategy for the treatment of a diverse range of pathological conditions.<sup>9,10</sup>

The DFG-out binding mode was discovered by serendipity in combination with structure–activity relationship (SAR)-guided medicinal chemistry.<sup>11</sup> The studies with the development of the eight distinct crystallographically proven type II inhibitors in the public domain: imatinib,<sup>12</sup> BIRB796,<sup>13</sup> sorafenib,<sup>14</sup> AAL993,<sup>15</sup> diaryl urea,<sup>13</sup> indole amide,<sup>16</sup> anilinoquinazoline,<sup>17</sup> and 4-aminopyrimidinoquinazoline<sup>18</sup> have demonstrated the importance of the DFG-out binding mode in kinases in general and p38 MAP kinase in particular. A number of sequence analysis studies have dealt with the similarity and the differences in the ATP binding site of kinases as well as the conserved regions. However, comprehensive studies considering sequence, structure, and active

\* Corresponding author: E-mail: gnsastry@gmail.com. Telephone: +91-40-2719 3016.

site aspects of one particular kinase with extra precision on the conformation of DFG-loop are rather scarce. Therefore, in the current project we focus on the detailed analysis of active site, DFG-loop, along with emphasis on the sequence and the structural analysis of p38 MAP kinase involved in the proliferation of inflammatory diseases.<sup>19–21</sup>

The p38 MAP kinase crystal structures of *Homo sapiens* and all other organisms were downloaded from Protein Data Bank (PDB), and cocrystals were extracted. The structures were clustered into different groups on the basis of sequence and source from which they were taken. Multiple sequence alignment (MSA) of sequences of PDB crystal structures was done using ClustalW to identify mutation prone regions and dissimilarities. Seven representative sequences of this set four from *H. sapiens* and three from *Mus musculus* were subjected to MSA to identify differences in key domains and active sites. This was followed by a detailed study of cocrystals and binding modes leading to the generation of fragment library for the design of type II p38 MAP kinase lead. In the next phase of lead design, a fragment library was generated from the p38 MAP kinase cocrystals and the FDA approved drugs so as to effectively exploit the existing structural preferences of the already established chemotypes that target subpockets within the overall active site. Based on our past experiences and applying the present strategy, new type II inhibitors have been designed by the rational application of molecular modeling methods.<sup>22–26</sup>

A number of lead design strategies spanning frontier areas of drug design like QSAR,<sup>27,28</sup> high-throughput screening,<sup>29</sup> pharmacophore mapping,<sup>30</sup> combinatorial libraries,<sup>31</sup> traditional drug design techniques,<sup>32,33</sup> etc., have been used either individually or as a combination to understand the nuances underlying specificity and selectivity in kinases, which are one of the most sought after class of druggable targets. Different aspects of kinases like size of gatekeeper residue,<sup>34</sup> metabolic properties of the target involved in signal transduction,<sup>35,36</sup> protonation switch in drug binding,<sup>37</sup> network analysis and fingerprint of structure and sequence of kinome,<sup>38</sup> recognition properties affecting selectivity,<sup>39</sup> chemogenomics,<sup>40</sup> and fragment-based drug design<sup>41–44</sup> have been probed. In spite of the ongoing efforts, kinase inhibitors lack specificity and selectivity. An attempt has been made to fill in these lacunae by exploiting the sequence, structure, active site, binding modes, and DFG conformation of the existing inhibitors and drugs for the design of kinase leads. We hereby propose for the first time a new multifacet lead design strategy for the design of type II p38 MAP kinase.

## 2. MATERIALS AND METHODS

**A. Homology Modeling.** Prime and Modeler 9v5 were used to build the breaks in the structures of p38 MAP kinase downloaded from PDB.<sup>45,46</sup> The template sequence was selected depending on the basis of the source of target sequence from among the seven representative sequences and the DFG conformation (Table 1). The quality of the constructed model was cross validated with Ramchandran plot (RAMPAGE),<sup>47</sup> energy profiles with ProSA,<sup>48</sup> and the secondary structure was determined with STRIDE.<sup>49</sup>

**B. Protein and Ligand Preparation.** The entire set of 138 structures downloaded from PDB and 31 FDA approved drugs of p38 MAP kinase were subjected to protein and

ligand preparation. Two different program packages, Sybyl 6.9<sup>50</sup> and Schrodinger,<sup>51</sup> have been used for the purpose of molecule preparation, active site search, and docking. For GOLD<sup>52</sup> and CDOCKER<sup>53</sup> docking, preparation of proteins and ligands was done using Sybyl 6.9. All the ligands were built and minimized using Tripos force field; MMFF94 atomic partial charges implementing Powell's conjugate gradient method with a distance-dependent dielectric constant, until a convergence gradient of 0.02 kcal/mol was accomplished. For protein preparation, all hydrogens were added using the Biopolymer module of Sybyl package, and inhibitors within the active site, heteroatoms, and all waters were removed. Proteins were minimized applying Kollman's all partial atomic charges, Powell's conjugate gradient method with distance-dependent dielectric constant value of 1.0, and a gradient convergence value of 0.05 kcal/mol. While for Glide, protein preparation wizard was used to prepare the proteins after adding hydrogens with default settings. The side-chain residues were refined using Prime. Ligands were submitted to the LigPrep module to generate a range of ionization states populated at a given pH range of  $7.4 \pm 2$  followed by Confgen.

**C. Docking.** Three docking protocols have been employed viz., Glide 4.5, GOLD 3.2, and CDOCKER 2.1. For Glide docking, the two options of docking namely simple precision (SP) and extra precision (XP) of 'Glide' module were used. The grid was generated by specifying M109 residue as grid center. The default SP docking settings were used, and the conformations obtained from SP were used as input for XP. Hydrophobic and hydrophilic maps were generated to get an idea of the solvent accessible regions. Docking with GOLD and CDOCKER was done with the default parameters using the backbone nitrogen of M109 residue as center. Ten poses were generated for each of the 138 cocrystal complexes with all the three docking protocols.

## 3. STRATEGY

This section comprises of the step by step approach used to build up a strategy to decipher the differences and the uniqueness of existing cocrystals and drugs underlying the sequence and structures of p38 MAP kinase for correlation and application in lead design.

**A. Sequence–Structure Analysis.** Structure of the receptor being an important factor in structure-based drug design, a meticulous study of the crystal structures, cocrystals, and their sequences of p38 MAP kinase will be a starting point for the design of lead. The p38 MAP kinase crystal structures from all organisms were thus downloaded from PDB.<sup>54</sup> The total downloaded data set (98) comprised of structures from *H. sapiens* (86), *M. musculus* (11), and one structure which we term as hybrid contains chains from both *H. sapiens* and *M. musculus* (Table S1, Supporting Information). A detailed study of all the crystal structures individually embarking upon issues like MODRES, breaks, sequence length, and side chains requiring refinement was carried out. The quality of crystal structures is widespread ranging from 1.45 to 2.8 Å with an exception of 2ONL, which has a resolution of 4 Å. However, there are few amino acid residues which need refinement. Seven modified cysteine residues are observed in six crystal structures (PDB: 2BAL, 2BAQ, 3D7Z, 3D83,

**Table 1.** Structure and Sequence Analysis of p38 MAP Kinase Structures From PDB<sup>a</sup>

S. no.	feature	<i>H. sapiens</i>	<i>M. musculus</i>
1	UNIPROT ID	<b>Q16539</b>	<b>P49137</b>
2	EC no.	2.7.11.24	2.7.11.24
3	gene	MAPK14	MAPK14
4	chain length	UNIPROT PDB 360	365 370 360 1–364 2–352 <sup>i</sup> 47–400 <sup>f</sup> 370–393 <sup>g</sup> 1–400 <sup>h</sup>
5	domain	kinase ATP TXY	24–308 (285AA) 30–38 (9AA) 180–182 (3AA)
6.	mutations	lowered kinase activity: A34Y, Y69H, A320T; loss of kinase activity: K53R, D168A, T175A, D177A, T180E, Y182F, W337R; hypothetical: R67G; emulation of active state, increase in activity: D176A, F327L	64–325(262AA) 70–78(9AA) – phosphorylation blocked: T180A, Y182A T180A, Y182A Y182F 25–308(284AA) 31–39(9AA) 180–182(3AA) lowered kinase activity: T180A, Y182A inactivation: T180A, Y182F 24–308(285AA) 30–38AA(9AA) 180–182AA(3AA) phosphorylation blocked: T180A, Y182F 24–308(285AA) 30–38AA(9AA) 180–182AA(3AA) inactivation: T180A, Y182F
7	no. of structures	protein co-crystal 82 124	3 – 1 – 3 6 12 8

<sup>a</sup> Deciphering the differences in gene; chain length; the three domains: kinase, ATP, activation loop (TXY); co-crystals and amino acid residues with propensity to mutate in the seven sequences of *H. sapiens* and *M. musculus* comprising the p38 MAP kinase crystal structures and FDA drugs. Details on chain type (a-o) specified in Table S3, Supporting Information.

3E92, 3E93). Cysteine residues are modified into *s*-hydroxy cysteine, *s*-mercapto cysteine, and hydroxy ethyl cysteine. Breaks are observed in different parts of the crystal structures of protein. A close observation reveals a concurrent appearance of breaks in the main DFG region (D168, F169, and G170) (Table S2, Supporting Information). The terminal end patches of sequences are found to be missing in most of the crystal structures. As these parts do not lie near the binding site region even in the coiled structure of the protein, their absence is not strongly felt.

Clustering of these structures of PDB according to sequence type emancipates that the structures comprise of seven different sequences from two organisms (*H. sapiens* and *M. musculus*) differing in sequence length constituting both homo- and heterochains (Table S3, Supporting Information). *H. sapiens* crystal structures comprise of four different sequences dominated by Q16539 constituting the MAPK14 gene, whereas the *M. musculus* crystal structures constitute a small number of crystal structures comprising mostly the P74811 sequence expressed by the gene Mapk14, which is analogous to the MAPK14 gene in *H. sapiens*. Thus there is a need to identify the commonalities and the differences within sequences of the same species and also across species.

The bound ligands were extracted from the downloaded PDB structures. There are in total 138 (H:130, M:8) cocrystals in the 84 (H:81, M:3) downloaded proteins; 10 proteins do not have cocrystals. These cocrystals were also clustered on the basis of sequence type to observe bias if any for a particular motif (Tables S4 and S5, Supporting Information). Majority of the structures (58) occupy the DFG-in conformation consisting chain A of Q16539 of *H. sapiens*.

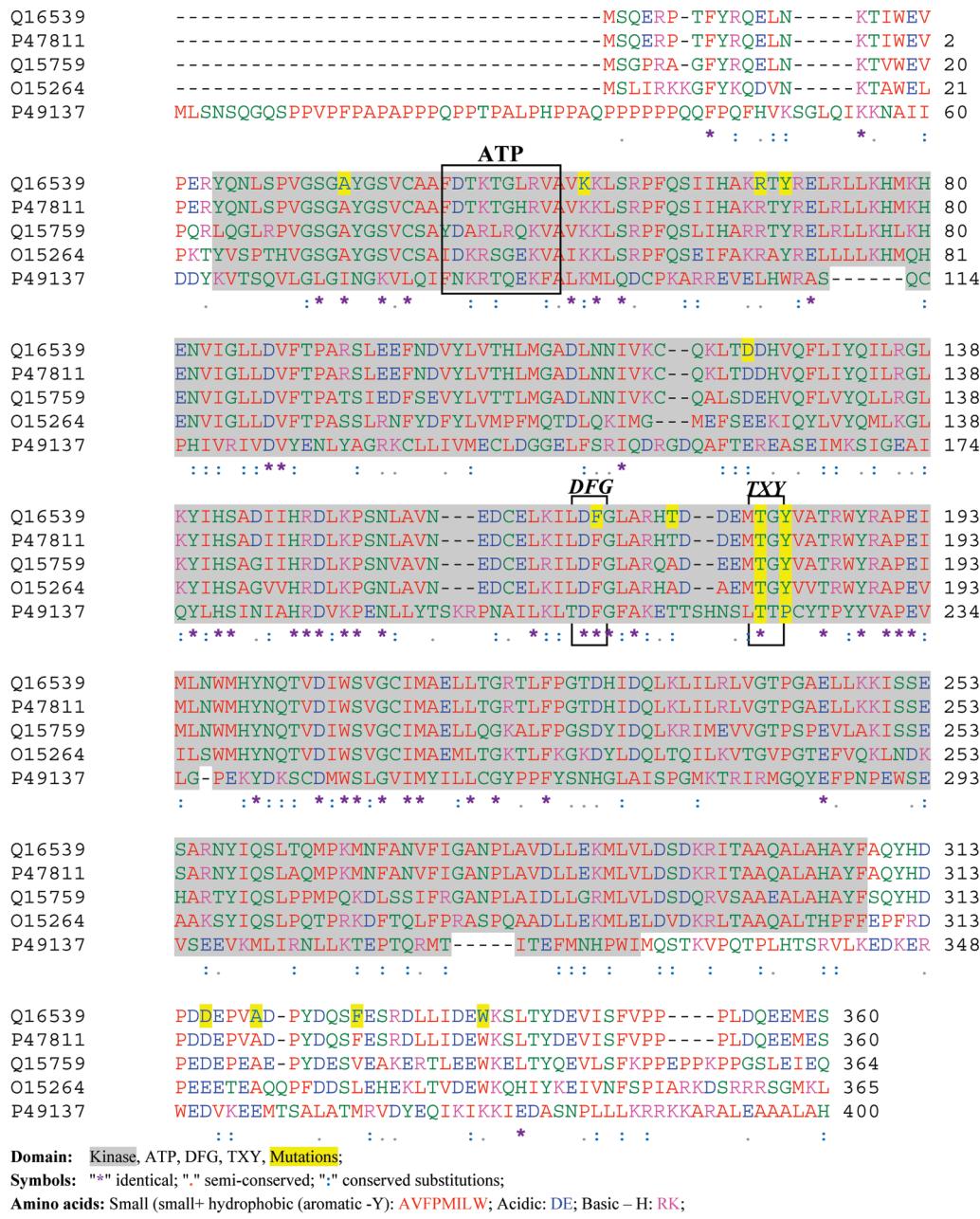
A total of 31 FDA approved MAP kinase drugs were collected and studied. These drugs were clustered into seven categories on the basis of isoform, gene, sequence type and chain length p38 $\alpha$ , MAPK14, Q16539, 360AA; p38 $\gamma$ , MAPK12, Q61C53, 357AA; ERK2, MAPK1, P28482, 360AA; ERK1, MAPK3, P27361, 379AA; MAPK10, Q449Y8, 319AA; JNK dual specificity, MAP2K7, O14733, 419A; and dual specificity, MAP2K1, A4QPA9, 393AA.

**B. Multiple Sequence Alignment (MSA).** An analysis of the representative set of p38 MAP kinase crystal structures sequences from UNIPROT to identify the differences among crystal structures and also among their chains was carried out. Sequences of these structures were downloaded in FASTA format from UNIPROT. A MSA of these sequences was done with ClustalW.<sup>55</sup> To understand position specific changes in amino acid residues and to identify patterns in prime domains like kinase, nucleotide, and activation loop, MSA was done with the representative set of five sequences viz., Q16539, P49137, O15264, Q15759, and P47811. The two sequences Q02078 and 1778153 being short peptide sequences were omitted. The MSA reflects >95% sequence similarity in most of the sequences except a few (Figure 1). During MSA it was observed that out of the three sets of sequences comprising the *M. musculus* crystal structures, two sequences (Q02078 and 1778153), are short amino acid fragments which are negligible. The major domains (kinase, ATP, and TXY) in these sequences are composed of conserved sets of amino acid residues reflecting high similarity in the binding site region. Mutation probability is found to be very high in the activation loop (TXY) and in

phenylalanine (F169) of the DFG-loop, extending a warning against the design of lead binding only to these loops, as such inhibitors are likely to be ineffective in case of mutation. Dissimilarities are observed in many regions in the MSA except the domains which may be due to the different sequences they comprise of. This led to probing into finer details such as genes, chain length, domain position, and mutation, which explains the reason for similarities and differences observed in MSA likewise also pointing to the regions which should be targeted for inhibitor binding (Table 1). The structures belong to different origin and subfamily and chain viz. *H. sapiens* (p38 $\alpha$ , p38 $\beta$ , p38 $\delta$ , MAPK2) and *M. musculus* (Mapk14, MEF2A, MKK3b). A cursory look at the genes explains that the similarity between sequences exists due to their being isoforms or belonging to the same subfamily, the reason for dissimilarity being the difference in sequence length. The high similarity observed in domains during MSA is confirmed, but the fact that the domains are present in the same analogous position in all sequences is also observed. The reason for mutation when traced shows lowering of kinase activity in case of mutation in one of the domains.

**C. Active Site Analysis.** Active site analysis and binding modes of inhibitors correlating the two DFG conformations were studied using docking for the entire data set. The active site of all structures were analyzed in terms of the position of the three amino acid residues (D168, F169, and G170) which constitute the DFG-loop with respect to the ATP and allosteric binding site, likewise classifying them on the basis of the DFG-loop conformation. A conformational transition leading to a shift of F169 side chain from its usual buried location (DFG-in) to a location that sterically interferes with the ATP binding site is noticed. This led to further probing of F169 which occupies the lipophilic space. F169 side chain usually moves ~10 Å around the ATP site sugar pocket leading to the reorganization of the DFG-loop in conjunction to the activation loop. The outward movement of F169 in most of the complexes is in conjunction with the inward movement of M78. A flip in the G110 residue of the DFG-loop is observed that causes the side chain of the G110 residue to fall inside the active site pocket creating room to accommodate a carbonyl anchor in the arylpyridazin type of templates having high dipole movement leading to enhanced activity. Small conformational changes have also been noticed due to the induced fit effect of inhibitors.

Consequently, the binding modes of the 138 cocrystals and 30 FDA approved drugs were studied by docking these cocrystals on to their receptors in different conformations in order to visualize and correlate the change in occupancy of the active site by the different conformers of the same inhibitor (Table S9, Supporting Information). The contribution of individual scoring components of DFG-in and -out structures reveals that H-bonding interactions are favorable for binding in the DFG-out active site due to the exposure of interactive residues like (T106, M109, D168, F169, E170). The DFG-in active site however favors van der Waal's (vdW) interaction due to the closely enclosed amino acid residues around the usual H-bond formers S104, T106, H107, and K110 (Table 2). Interactions with residues K53, E71, S104, T106, H107, M109, K110, D168, and A235, having high propensity to mutate, have been observed in many cocrystals repetitively cautioning the careful use of such fragments in lead design.



**Figure 1.** Comparative analysis of the MSA of the representative sequences constituting the p38 MAP kinase crystals from PDB with ClustalW2.1, depicting the highly conserved DFG-loop and domains (kinase, ATP, and TXY) with possible sites of mutation in the activation loop.

Type I ATP site inhibitors were found to interact mostly with the residues (S104, H107, M109, and K110) of the ATP binding pocket and the gatekeeper residue T106, whereas type II (allosteric site inhibitors) interacts with the two hydrophobic pockets along with the DFG-loop; the gatekeeper residue T106 is found to be inaccessible in most of the cases. A shift in the activation loop conformation toward the ATP binding site has been observed in case of inhibitors which interact with the DFG motif successfully, thus, mimicking substrate binding and preventing kinase activity.<sup>56</sup> Designing type II leads mimicking the interactions of its predecessor is likely to prove more effective, as subsequent analysis reveals that all type II p38 MAP kinase FDA drugs and cocrystal structures share a similar pharmacophore and exploit a conserved set of H-bonds. The most stable binding mode which was found to be similar to the cocrystal pose in a majority of the cases formed the basis for the formulation of fragment library (Figure 2). The cocrystals

and drugs were noticed to form H-bond interactions with key residues which have been proved to be responsible for inhibition through mutational studies experimentally. We validated this by mutating these amino acid residues in silico using Schrodinger. Receptor–ligand interactions ceased to form on mutation of the key interacting residues (K53, E71, G77, L83, K97, S104, T106, H107, M109, K110, K120, D168, M149, E180, E186, and A235).

#### 4. DESIGN

In this section, we delineate our approach for the design of specific kinase inhibitors. It starts with the description of generating the fragment library, followed by the approaches used to optimize the chemotype requirements for the DFG-in and -out conformations. Finally, a section on the in silico validation based on docking has been given.

**Table 2.** Evaluation of the Performance of (a) Glide 4.5 and (b) GOLD 3.2 Docking for ATP (DFG-in) and Allosteric (DFG-out) Site Binders of p38 MAP Kinase Used for Lead Design**a DFG-in (ATP site binders)**

S. No.	PDB id	Co-crystal	vdW	VdW Rank	Coul	Coul Rank	Lipo	Lipo Rank	HBond	Metal	Rewards	RotB	Site	Gscore	GRank	RMSD	RMSD Rank
1	1A9U	SB2	-38	6	-6.2	6	-2.2	2	-0.5	0	-1.5	0.4	-0.1	-6.63	3	1.2004	3
2	1BMK	SB5	-15.8	2	-3.6	3	-1.8	1	0	0	-2.4	0.8	-0.1	-4.84	7	1.3254	4
3	1DI9	MSQ	-36.6	5	-1	1	-2.2	3	0	0	-1.2	0.4	-0.1	-5.08	5	2.6724	8
4	3C5U	P41_362	-14.5	1	-6.3	7	-3.3	5	-0.3	0	-2	3.4	0	-3.88	8	1.5650	6
5	1WBW	L14	-23.2	3	-6.7	8	-3.3	6	-0.6	0	-2.3	3.4	0	-4.95	6	1.6920	7
6	1W84	L12	-36.5	4	-5.5	5	-4	8	-0.5	0	-2.7	1	0	-8.84	1	0.0688	1
7	1ZYJ	B15	-43	8	-1.1	2	-3.7	7	0	0	-1.8	1	0	-6.78	2	1.4239	5
8	2BAL	PQA	-42.9	7	-4.4	4	-2.3	4	0	0	-0.6	0.4	0	-5.33	4	1.0304	2

**DFG-out (Allosteric site binders)**

S. No.	PDB id	Co-crystal	vdW	VdW Rank	Coul	Coul Rank	Lipo	Lipo Rank	HBond	Metal	Rewards	RotB	Site	Gscore	GRank	RMSD	RMSD Rank
1	1WBN	L09	-51.1	2	-17.4	8	-5.5	5	-1.6	0	-1.8	0.9	0	-13.24	1	0.2999	2
2	2BAJ	1PP	-55.3	5	-11	6	-5.4	4	-0.8	0	-0.9	0.5	-0.1	-11.08	4	0.4965	3
3	1W83	L11	-51.8	4	-9.6	5	-5.8	7	-0.7	0	-2.1	0.7	-0.1	-12	3	0.7727	5
4	1KV2	B96	-65.9	8	-6.3	1	-6.3	8	-0.5	0	-0.5	0.6	0	-11	5	1.1655	6
5	1W82	L10	-55.3	6	-6.8	3	-5.1	3	-0.6	0	-0.9	0.6	0	-9.81	7	0.5042	4
6	1WBV	L13	-51.4	3	-6.4	2	-4.9	2	-1.9	1	-2.3	0.7	0	-10.59	6	0.2879	1
7	2BAQ	PQB	-37.6	1	-11.3	7	-1.9	1	-5.5	6	-1.1	1	0	-5.91	8	1.9535	8
8	2BAK	AQZ	-61.2	7	-9.1	4	-5.5	6	-0.7	0	-2.1	0.4	0	-12.26	2	1.2353	7

Legend RMSD:

0 - 0.5 Å	0.5 - 1.0 Å	1.0 - 1.5 Å	1.5 - 2.0 Å	2.0 - 2.5 Å	>2.5 Å
-----------	-------------	-------------	-------------	-------------	--------

**b DFG-in (ATP site binders)**

S. No.	PDB id	Co-crystal	S(hb_int)	S(hb_ext)	S(hb_ext) Rank	S(int)*	S(int)* Rank	S(vdw_ext)	S(vdw_ext) Rank	Fitness	GOLD Rank	RMSD	RMSD Rank
1	1A9U	SB2	0	1.21	6	-13.74	7	38.35	3	40.21	5	0.678	3
2	1BMK	SB5	0	5.09	2	-12.04	6	28.81	7	32.67	8	0.847	4
3	1DI9	MSQ	0	1.72	4	-10.68	4	38.92	2	44.55	2	1.453	7
4	3C5U	P41_362	0	1.74	3	-0.95	1	24.29	8	34.19	7	1.051	5
5	1WBW	L14	0	0	7	-9.22	3	35.3	5	39.31	6	0.639	1
6	1W84	L12	0	0	8	-4.47	2	33.94	6	42.2	4	1.145	6
7	1ZYJ	B15	0	8.88	1	-10.7	5	36.3	4	48.1	1	2.031	8
8	2BAL	PQA	0	1.23	5	-17.88	8	42.87	1	42.29	3	0.650	2

**DFG-out (Allosteric site binders)**

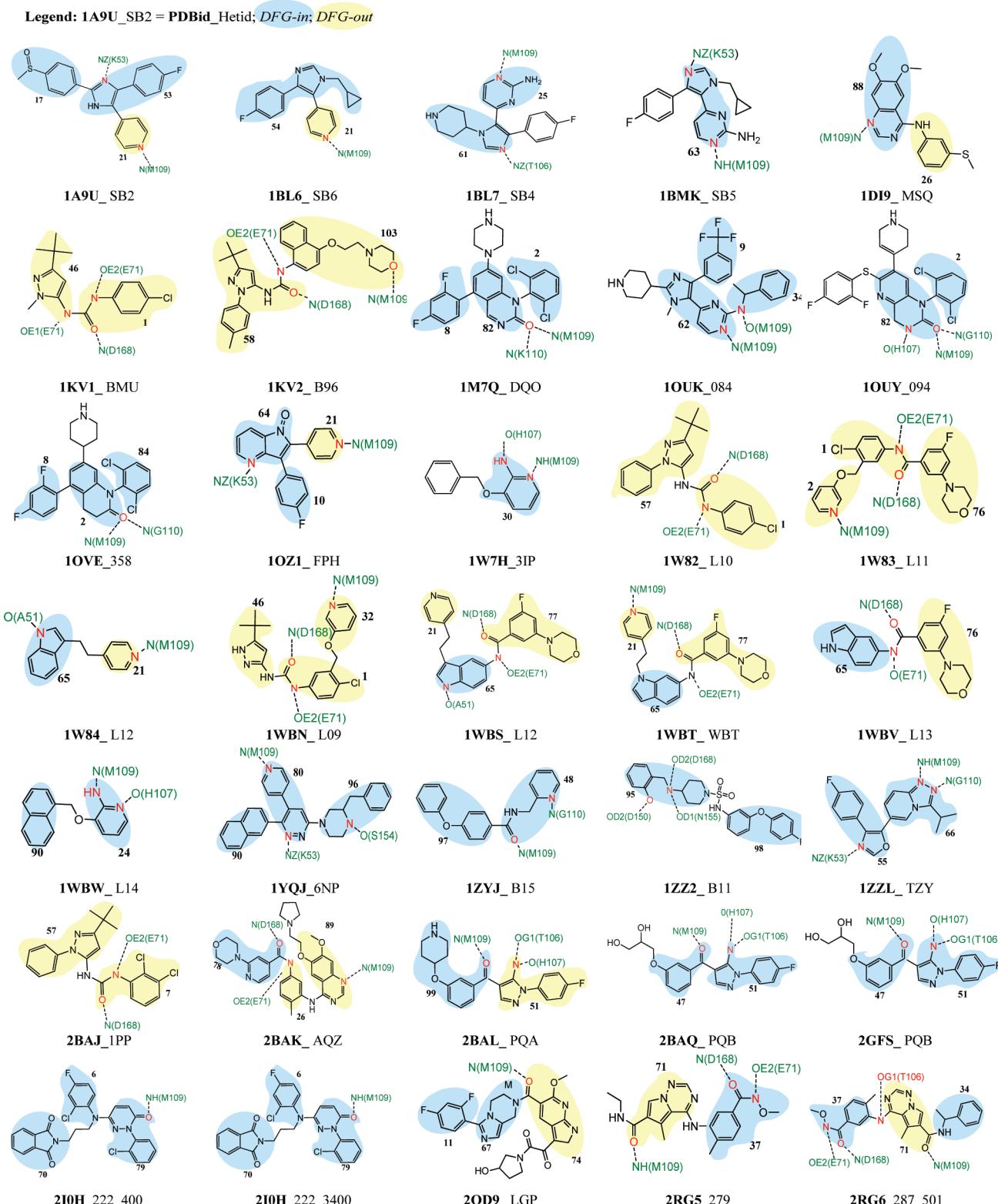
S. No.	PDB id	Co-crystal	S(hb_int)	S(hb_ext)	S(hb_ext) Rank	S(int)*	S(int)* Rank	S(vdw_ext)	S(vdw_ext) Rank	Fitness	GOLD Rank	RMSD	RMSD Rank
1	1WBN	L09	0	6.13	3	-15.11	2	42.77	4	49.82	2	0.909	2
2	2BAJ	1PP	0	2.9	6	-19.17	5	38.06	7	36.07	7	2.006	8
3	1W83	L11	0	0.09	8	-21.18	7	43.86	3	39.22	5	1.648	6
4	1KV2	B96	0	7.69	1	-40.82	8	46	2	30.12	8	1.369	4
5	1W82	L10	0	6	5	-13.89	1	39.51	5	46.43	3	1.014	3
6	1WBV	L13	0	6.01	4	-17.13	3	37.26	8	40.11	4	1.414	5
7	2BAQ	PQB	0	2.69	7	-19.37	6	38.86	6	36.76	6	1.943	7
8	2BAK	AQZ	0	7.5	2	-18.97	4	50.7	1	58.24	1	0.878	1

Legend RMSD:

0 - 0.5 Å	0.5 - 1.0 Å	1.0 - 1.5 Å	1.5 - 2.0 Å	2.0 - 2.5 Å	>2.5 Å
-----------	-------------	-------------	-------------	-------------	--------

**A. Generation of Fragment Library.** In the past decade, fragment-based drug discovery has led to the discovery of new scaffolds that were later combined or grown into high-affinity inhibitors.<sup>57-59</sup> Small molecule p38 MAP kinase inhibitors are widely employed as biological reagents and as leads in the design of drugs for rheumatoid arthritis. A particular challenge in fragment-based drug discovery is the

identification of initial binding fragments. The binding modes of cocrystals and FDA approved MAP kinase drugs have been used as a starting point to formulate a dictionary of keys (Tables S6 and S7, Supporting Information). Fragments were generated by pruning cocrystals and drugs on the basis of binding interactions and were further segregated according to the site or part of the pocket occupied. The library

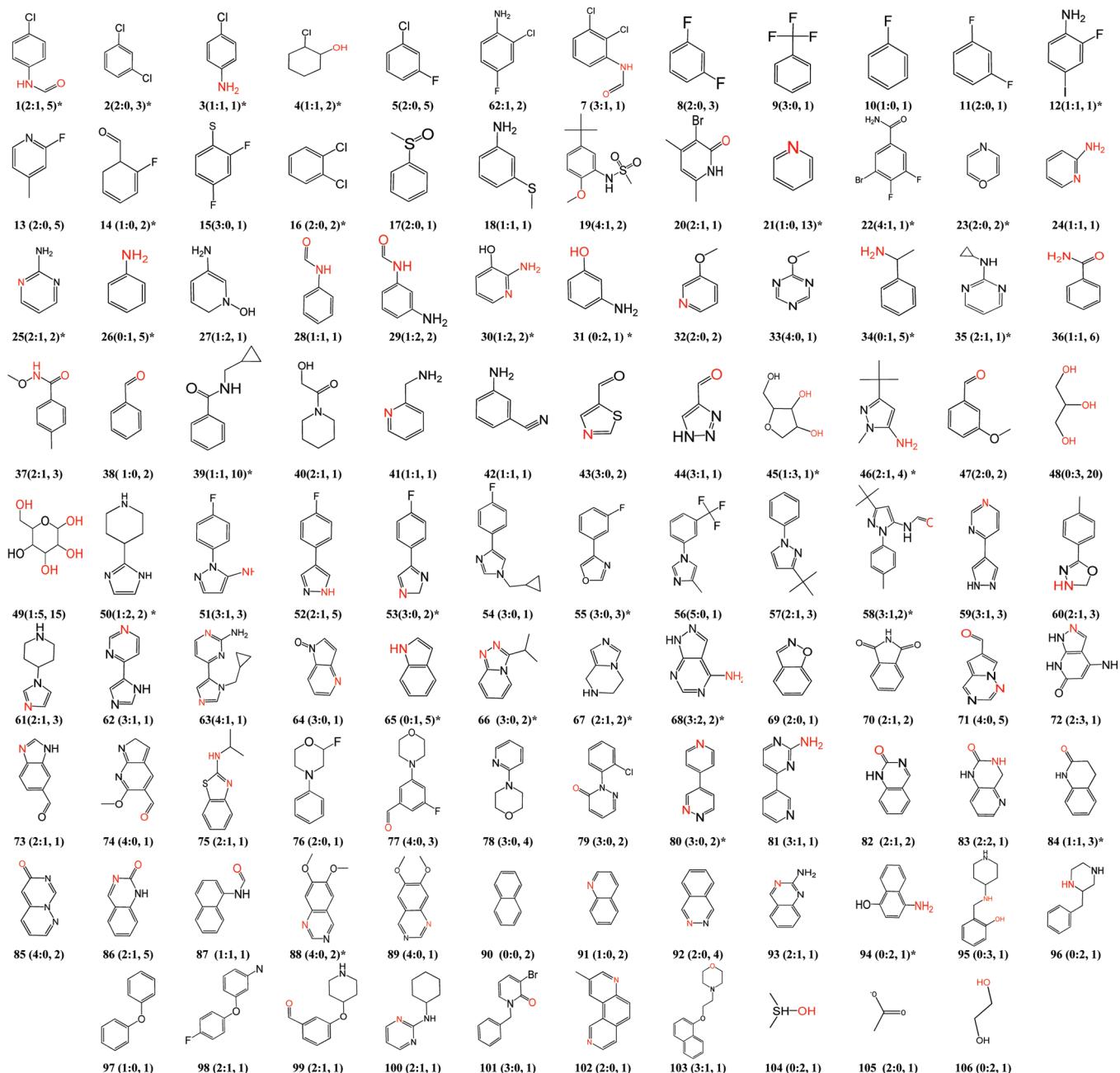


**Figure 2.** Receptor–ligand interactions of p38 MAP kinase cocrystal complexes from PDB obtained on docking with Glide 4.5 and their segregation into fragments on the basis of DFG conformation.

constitutes 210 fragments from *H. sapiens*, 3 from *M. musculus*, and 31 from FDA approved drugs. A total of 106 nonredundant sets of keys have been formulated with an overlap of 13 drug fragments (Figure 3). Nature and frequency of occurrence of individual fragments of the library revealed the participation of acceptors and donors of the same fragment in different scaffolds in similar set of interactions (Table S8, Supporting Information). This observation led to

the understanding that a given fragment has a tendency to occupy the same part of the binding site irrespective of its position in the parent scaffold and thus displays interactions with same residues. The fragment library was therefore eventually segregated into ATP and allosteric site binders in order to tailor these categorized fragments into new leads.

**B. Design of New Leads.** Majority of type I p38 MAP kinase cocrystals and drugs are heterocyclic and flat moieties



**Figure 3.** Fragment library of 106 fragments generated from 138 cocrystals from PDB and 31 FDA drugs of p38 MAP kinase on the basis of binding modes and interactions.

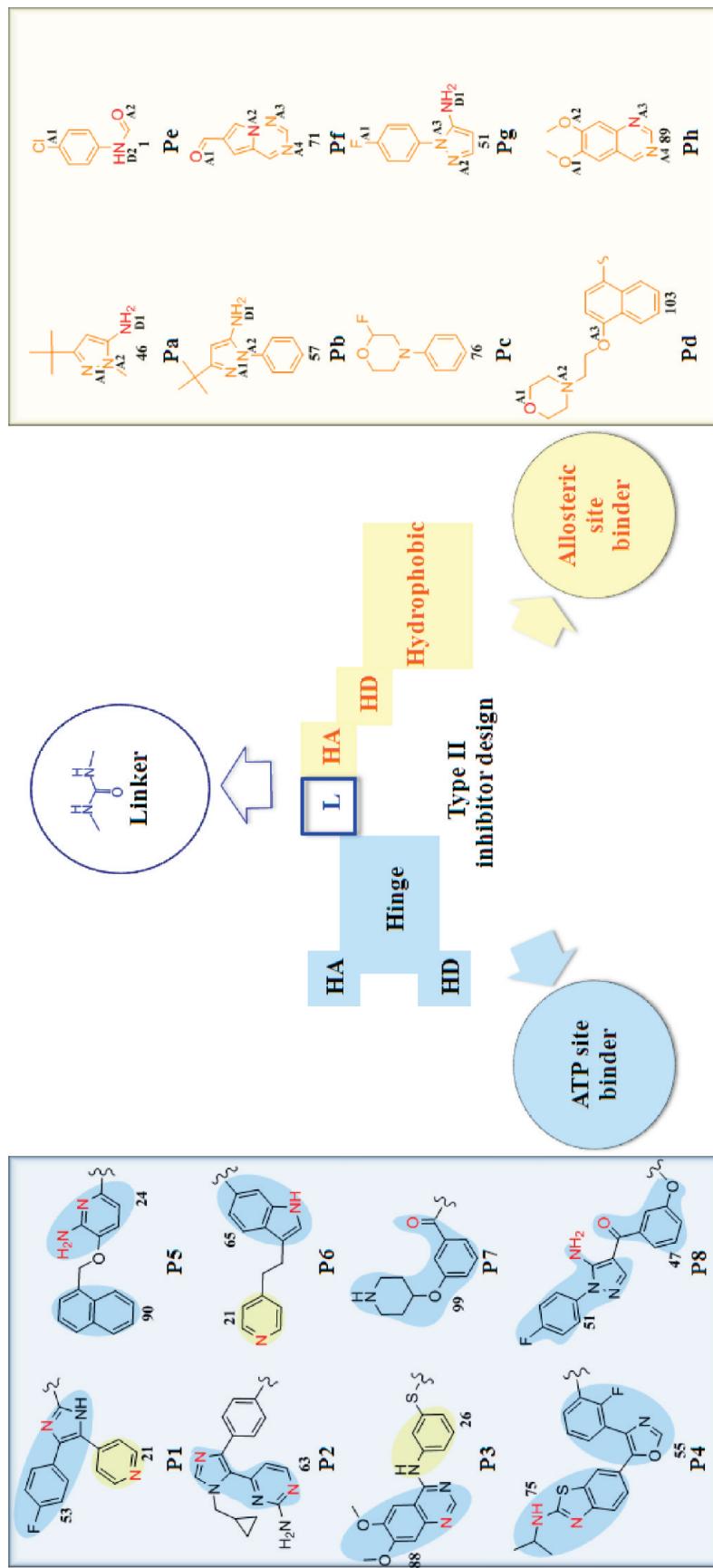
with hydrophobic nature, imitating ATP binding. Allosteric p38 MAP kinase inhibitors encompass lengthy scaffolds constituting dense ring systems with H-bond acceptors and donors along with a hydrophobic moiety, which occupies the space created by the shift of F169 on transition of DFG-loop from in to out.<sup>60–62</sup> New type II p38 MAP kinase leads were designed from type I considering the following points observed on docking diverse type I and II inhibitors onto different conformations of p38 MAP kinase:

- ATP site binder: A well established p38 MAP kinase type I inhibitor capable of inhibiting kinase independently along with an inclination to form H-bonds with amino acid residues of the hinge region and hydrophobic interactions in the adenine region of the ATP binding site.
- Linker: A short chemical moiety which can be used as a hook to join the loose ends of ATP and allosteric site

binders and likewise also interact with the DFG-loop.

- Allosteric site binder: A unique scaffold to bind exclusively to the allosteric site of p38 MAP kinase consisting of a H-bond donor–acceptor pair and a hydrophobic motif (Figure 4).

The design of leads followed a step by step protocol (Chart 1). The main idea is to put together observations of different aspects like: similarity, mutational nature of sequence, and domains to avoid inclusion of fragments which have a tendency to form bonds with residues like the F169 of the DFG-loop and the T180 and Y182 of the activation loop; structural details like quality, MODRES, breaks in the choice of receptor for docking; probing the active site to delineate the features of the ATP and allosteric site in relation to the changes induced on the binding of inhibitors for effective



**Figure 4.** Formulation of a blue print for the design of type II p38 MAP kinase inhibitors. The design encompasses three parts starting with ATP site binders garnered from p38 MAP kinase cocrystals and FDA drugs (sky blue) attached to a common linker (navy blue) and extended with an allosteric site binders (orange) picked up from the fragment library generated. A total of 64 type II inhibitors were grown from type I using all possible combinations of the eight ATP and allosteric site binders with the common linker.

**Chart 1.** Fragment-Based Drug Design Strategies Employed for the Design of Type II p38 MAP Kinase Inhibitors from Type I Co-Crystals and FDA Drugs

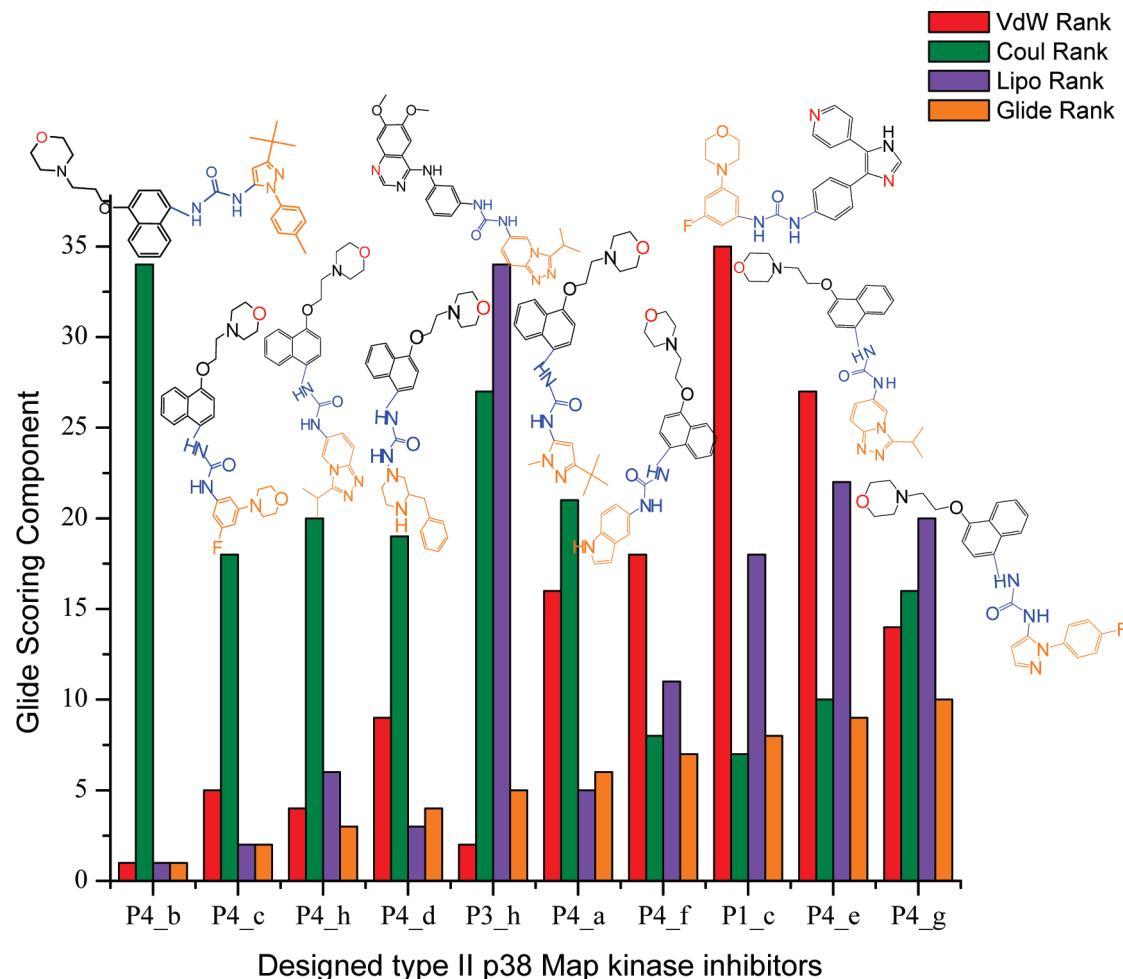
	<b>Objective</b>	<b>Result</b>	<b>Inference</b>
<b>STRATEGY</b>	<b>Sequence - Structure analysis</b> of co-crystals and binding sites of crystal structures and FDA drugs of p38 Map kinase	Crystal structures : 98 Co-crystals : 132 FDA drugs : 31 Source : 2 Sequence type : 7 Gene type : 7 Quality : 2.0-2.8 Å	<ul style="list-style-type: none"> <li>For the choice of receptor for SBDD in p38 Map kinase, the <i>DFG</i> conformation and co-crystal is the basis as most of the structure share high sequence similarity</li> <li>Type and chemical nature of co-crystals encourage the use of diverse scaffolds</li> </ul>
	<b>Sequence analysis</b> Multiple sequence alignment - ClustalW	<ul style="list-style-type: none"> <li>&gt;95% similarity</li> <li>Highly conserved functional domains (ATP, kinase, <i>TXY</i>)</li> <li>High propensity for mutation in both ATP and allosteric binding site</li> </ul>	<ul style="list-style-type: none"> <li>Conserved nature of functional domains and their propensity for mutation indicate the need for binding interactions with other regions like <i>DFG</i> (168-170), Gatekeeper residue (T106) for long term use of designed inhibitor</li> </ul>
	<b>Active site analysis</b> <i>DFG</i> conformation and binding modes of inhibitors	<ul style="list-style-type: none"> <li>Identification of region specific fragments</li> </ul>	<ul style="list-style-type: none"> <li>Binding modes of those parts of inhibitor bind to ATP site is same in both type I and II</li> <li>Type I inhibitor scaffold can be grown into type II</li> </ul>
<b>DESIGN</b>	<b>Dictionary of fragments</b> Generation of fragment library from co-crystals and FDA approved drugs	Co-crystals : 213 (H: 210, M: 3, Hybrid: 0) FDA drugs : 31 Total : 244 Frequency : 1-20 HA : 1-5 HD : 1-3	<ul style="list-style-type: none"> <li>Existing p38 Map kinase inhibitors although diverse still share similarity in fragments</li> <li>A given fragment has a tendency to occupy the same part of the binding site irrespective of its position in the parent scaffold and thus displays interactions with same residues</li> <li>The design of inhibitor using region specific fragments is likely to be more successful</li> </ul>
	<b>Formulation of blue-print for p38 MAP kinase inhibitor</b>	3 main parts <ul style="list-style-type: none"> <li>Head (ATP region)</li> <li>Linker (<i>DFG</i> loop)</li> <li>Tail (Allosteric region)</li> </ul>	<ul style="list-style-type: none"> <li><b>Head:</b> A well established p38 MAP kinase type I inhibitor capable of inhibiting kinase independently along with an inclination to form H-bonds with amino acid residues of the hinge region and hydrophobic interactions in the adenine region of the ATP binding site</li> <li><b>Linker:</b> A short chemical moiety which can be used as a hook to join the loose ends of ATP and allosteric site binders and likewise also interact with the <i>DFG</i> loop</li> <li><b>Tail:</b> A unique scaffold to bind exclusively to the allosteric site of p38 MAP kinase consisting of a hydrogen bond donor-acceptor pair and a hydrophobic motif</li> </ul>
	<b>Design of new leads</b>	<ul style="list-style-type: none"> <li>41 ATP site binders with 106 fragments, a total of 4346 leads</li> <li>8 ATP and 8 allosteric site fragments, a total of 64 leads</li> </ul>	<ul style="list-style-type: none"> <li><b>Validation :</b> Ensemble docking</li> <li><b>Synthetic feasibility:</b> Synthesis of individual fragments independently, Covalent linkage of the ATP site binder, urea linker, allosteric site Nucleophilic substitution reactions</li> <li><b>Drug likeliness</b></li> </ul>

utilization of the binding pocket; analysis of structure, binding mode and active site cocrystals and FDA drugs to generate a fragment library; and formulation of a blue-print to design an inhibitor which can bind effectively to ATP site, allosteric site with special emphasis to gatekeeper residue T106 and D168, and G170 of the *DFG*-loop.

The above methodology was put together to graft structural features that are present in different structures into a common template. The segregated fragments were docked onto the *DFG*-out conformations to delineate their affinity for particular part of the active site. However, considering the limitation of docking protocols in prioritizing fragments, the final subset was garnered on combined retrospective analysis. The final design was initiated by the selection of ATP site binders which are well established and exhibit high kinase selectivity, with a probability to be extended using an allosteric site binder and possibility of covalent linkage with its allosteric counterpart through a linker. A combination of

the 41 ATP site binders with 106 fragments gives a total of 4346 leads, which itself is a very huge starting point. Based on our interests, we sorted out eight ATP and eight allosteric site fragments to proceed for inhibitor design. Each of the eight ATP and allosteric site binders were thus picked up on the basis of the formulated blue -print such that the ATP site binder adds to inhibition, whereas the allosteric site binder brings in specificity (Figure 4). These fragments were assembled into 64 new leads using a common urea linker (Figure SF1, Supporting Information).

**C. Validation.** Evaluation of the designed inhibitors is necessary to visualize their binding, occupancy of the active site, and contribution of atoms of individual fragments in H-bond formation. The p38 MAP kinase structures crystallized in *DFG*-out conformation were considered since their cocrystals have been designed to target the *DFG*-out conformation. There are variations in active site conformation of the 138 p38 MAP kinase crystal structures due to the



**Figure 5.** Contribution of scoring components of the 10 best designed type II p38 kinase leads on docking with Glide 4.5.

induced fit nature of the diverse cocrystals. Receptor rearrangement on the binding of ligand (induced-fit) is a common and well-known feature of kinase inhibitors causing local movement of side chains to large domain shifts. Choice of receptor to be used is therefore a serious issue, wherein increasing the flexibility of receptor is necessary to predict accurate ligand binding affinities.<sup>63,64</sup> Although molecular dynamics brings in protein flexibility, it requires high-computational time. Therefore an ensemble of protein structures exclusively in DFG-out conformation has been taken in order to span maximum conformational space.<sup>65,66</sup> The ensemble constitutes one representative of source protein of each of the eight allosteric site fragments viz. a:46:1KV1, b:57:1KV2, c:77:2BAK, d:96:1YQJ, e:2:3FC1, f:65:1WBV, g:51:2BAQ, and h:66:3HP5 (Figure 5). The 64 designed inhibitors were thus evaluated in silico by docking them onto the active site of these eight proteins. Regular rigid docking using Glide was followed by the induced fit docking (IFD) of Schrodinger in an attempt to incorporate protein flexibility. The designed inhibitors were evaluated on the basis of their binding. The influence of induced conformational changes on docking results become apparent when a ligand is docked to a receptor cocrystallized with another inhibitor. Small variations in the structure of the binding pocket have a large impact on docking geometries, so ensemble docking experiments are useful to assess the magnitude of this influence on the designed inhibitors (Table 3). Visualization of the docked poses shows that designed inhibitors effectively

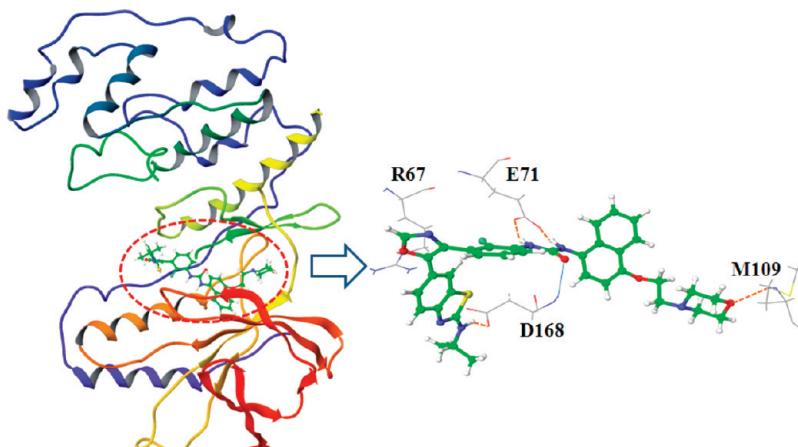
occupy the ATP, DFG, and allosteric pockets of the active site (Figure 6). The ATP site binder occupies the ATP binding site and forms H-bonds with R67, E71, and the front end of D168 of the DFG-loop, which forms a gateway for the inhibitor to enter the hydrophobic allosteric pocket adjacent to it in the DFG-out conformation. The linker occupies the DFG-loop forming H-bond with D168, whereas the allosteric site binder fits in the allosteric pocket forming H-bond with M109. Close contacts have also been observed in many cases with the gatekeeper residue T106. Comparison of the binding modes of type I and type II designed inhibitors validates the effective interactions of the type II designed inhibitor's ATP fragment with the hinge region and that of the allosteric fragment with the DFG-loop and the hydrophobic pockets opened up on transition of F169 forming DFG-out conformation (Figure 7).

**D. Drug-Likeness of the Designed Ligands.** The designed inhibitors were tested further in terms of their drugability and ADMET properties. ADMET properties: hepato toxicity (HT), cytochrome P450 inhibition (CYP P450), plasma protein binding (PPB), human intestinal absorption (HIA), aqueous solubility (Sol), and blood brain barrier (BBB) of the best 64 designed p38 MAP kinase inhibitors calculated with TOPKAT module of Discovery Studio 2.1 (Table 4). These properties play an important role in metabolic regulation in proteins including kinases.<sup>67</sup> Assessments of the various toxic effects of chemicals based on their molecule structure and calculated properties

**Table 3.** Scoring Components of Glide 4.5 on Docking with the DFG-Out Conformation of p38 MAP Kinase for the Top Ranking Designed Type II Inhibitors<sup>a</sup>

molecule	vdW	VdW rank	coul	coul rank	lipo	lipo rank	H-bond	metal	rewards	RotB	site	Gscore	glide rank
1 P4_d	-67.4	<b>1</b>	-6.9	<b>34</b>	-6.6	<b>1</b>	-0.6	0	-1	0.6	-0.1	<b>-11.99</b>	<b>1</b>
2 P4_c	-62.8	<b>5</b>	-8.8	<b>18</b>	-6.1	<b>2</b>	0.7	0	-1.2	0.7	-0.1	<b>-11.93</b>	<b>2</b>
3 P4_h	-63.5	<b>4</b>	-8.8	<b>20</b>	-5.6	<b>6</b>	-0.7	0	-1.7	0.6	-0.1	<b>-11.93</b>	<b>3</b>
4 P4_b	-60.7	<b>9</b>	-8.8	<b>19</b>	-6.1	<b>3</b>	-0.8	0	-0.9	0.8	-0.1	<b>-11.44</b>	<b>4</b>
5 P3_h	-64	<b>2</b>	-7.9	<b>27</b>	-4.7	<b>34</b>	-0.4	0	-2.1	0.5	-0.1	<b>-11.27</b>	>5
6 P4_a	-57.4	<b>16</b>	-8.7	<b>21</b>	-5.6	<b>5</b>	-0.8	0	-1.4	0.7	-0.1	<b>-11.27</b>	<b>6</b>
7 P4_f	-56.8	<b>18</b>	-10	<b>8</b>	-5.3	<b>11</b>	-0.8	0	-1.2	0.6	-0.1	<b>-11.08</b>	<b>7</b>
8 P1_c	-51.6	<b>35</b>	-10.3	<b>7</b>	-5.1	<b>18</b>	-0.3	0	-1.6	0.7	-0.3	<b>-10.75</b>	<b>8</b>
9 P4_e	-53.8	<b>27</b>	-9.8	<b>10</b>	-5	<b>22</b>	-0.8	0	-1.3	0.7	-0.1	<b>-10.68</b>	<b>9</b>
10 P4_g	-59.2	<b>14</b>	-8.9	<b>16</b>	-5.1	<b>20</b>	-0.7	0	-1	0.7	-0.1	<b>-10.55</b>	<b>10</b>
11 P3_a	-60.7	<b>8</b>	-8.4	<b>23</b>	-5.1	<b>19</b>	-0.4	0	-1.2	0.5	-0.1	<b>-10.49</b>	<b>11</b>
12 P8_b	-59.9	<b>12</b>	-5.7	<b>48</b>	-5	<b>24</b>	-0.3	0	-1.8	0.5	0	<b>-10.49</b>	<b>12</b>
13 P8_c	-58	<b>15</b>	-7.6	<b>31</b>	-4.8	<b>31</b>	-0.3	0	-1.8	0.7	-0.3	<b>-10.46</b>	<b>13</b>
14 P3_c	-54.6	<b>22</b>	-9.4	<b>12</b>	-4.3	<b>48</b>	-0.5	0	-1.9	0.5	-0.1	<b>-10.40</b>	<b>14</b>
15 P3_f	-59.6	<b>13</b>	-7.8	<b>28</b>	-4.9	<b>26</b>	-0.4	0	-1.3	0.5	-0.1	<b>-10.35</b>	<b>15</b>
16 P2_c	-55.2	<b>20</b>	-9.2	<b>14</b>	-4.4	<b>45</b>	-0.2	0	-1.9	0.7	-0.2	<b>-10.19</b>	<b>16</b>
17 P1_g	-48.9	<b>43</b>	-9.9	<b>9</b>	-4.9	<b>25</b>	-0.6	0	-1.2	0.7	-0.2	<b>-10.12</b>	<b>17</b>
18 P5_c	-53.1	<b>29</b>	-6.6	<b>36</b>	-5.3	<b>12</b>	-0.3	0	-1.6	0.8	-0.1	<b>-10.07</b>	<b>18</b>
19 P3_b	-52.5	<b>32</b>	-8.2	<b>25</b>	-5.3	<b>10</b>	-0.5	0	-0.7	0.4	-0.1	<b>-10.03</b>	<b>19</b>
20 P7_c	-56.8	<b>19</b>	-7.7	<b>29</b>	-4.6	<b>40</b>	-0.5	0	-1.3	0.9	-0.1	<b>-9.67</b>	<b>20</b>
21 P3_d	-38.5	<b>59</b>	-8.2	<b>26</b>	-5.5	<b>8</b>	-0.3	0	-1.3	0.6	-0.1	<b>-9.66</b>	<b>21</b>
22 P7_b	-62.7	<b>6</b>	-6.4	<b>38</b>	-5.3	<b>16</b>	-0.2	0	-0.8	0.8	0	<b>-9.63</b>	<b>22</b>
23 P1_h	-46.8	<b>52</b>	-8.2	<b>24</b>	-4.1	<b>51</b>	-0.5	0	-1.9	0.6	-0.1	<b>-9.58</b>	<b>23</b>
24 P2_d	-47	<b>50</b>	-13.8	<b>2</b>	-4.6	<b>37</b>	-0.5	0	-0.7	0.7	-0.1	<b>-9.58</b>	<b>24</b>
25 P5_g	-50.2	<b>41</b>	-6.7	<b>35</b>	-5	<b>23</b>	-0.3	0	-1.4	0.9	-0.1	<b>-9.48</b>	<b>25</b>
26 P1_e	-47.1	<b>49</b>	-8.9	<b>15</b>	-4.5	<b>42</b>	-0.5	0	-1.3	0.6	-0.2	<b>-9.47</b>	<b>26</b>
27 P7_h	-56.9	<b>17</b>	-7.7	<b>30</b>	-4.1	<b>55</b>	-0.4	0	-1.6	0.9	-0.1	<b>-9.39</b>	<b>27</b>
28 P1_f	-49.9	<b>42</b>	-6.3	<b>39</b>	-4.8	<b>29</b>	-0.5	0	-1.1	0.6	-0.1	<b>-9.38</b>	<b>28</b>
29 P6_c	-60.9	<b>7</b>	-3.9	<b>57</b>	-4.3	<b>49</b>	-0.2	0	-1.7	0.7	-0.2	<b>-9.37</b>	<b>29</b>
30 P8_d	-60.4	<b>10</b>	-4.2	<b>54</b>	-4.8	<b>32</b>	-0.1	0	-1.4	0.7	-0.2	<b>-9.36</b>	<b>30</b>
31 P7_d	-37.9	<b>60</b>	-15.2	<b>1</b>	-4.9	<b>28</b>	-0.3	0	-0.8	1	-0.1	<b>-9.35</b>	<b>31</b>
32 P6_d	-60.3	<b>11</b>	-4.1	<b>56</b>	-4.8	<b>30</b>	0	0	-1.4	0.7	-0.2	<b>-9.33</b>	<b>32</b>

<sup>a</sup> Ranks have been assigned to each scoring components of the 64 designed inhibitors to delineate their contributions individually.



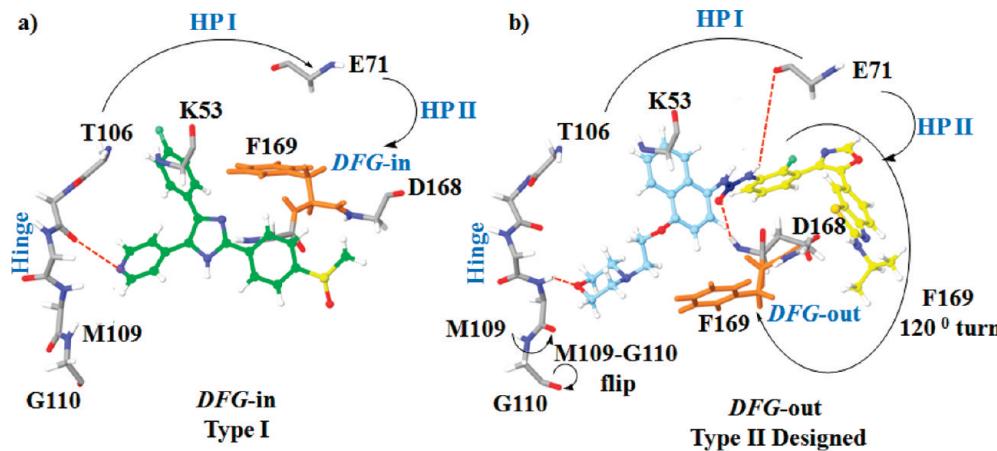
**Figure 6.** Interactions of the designed type II p38 kinase inhibitor P4\_d on docking with Glide 4.5. The ATP site binder occupies the ATP binding site and forms H-bonds with R67, E71, and the front end of DFG-loop interacting with D168. The linker occupies the DFG-loop forming H-bond with D168, whereas the allosteric site binder fits in the allosteric pocket forming H-bond with M109.

were used to develop quantity structure toxicity relationships (QSTR) and were tested on 14 models viz., developmental toxicity potential (DTP); mutagenicity (Ames test); rodent carcinogenicity; rat chronic oral LOAEL; skin sensitization (GPMT); skin irritancy; rat oral LD50; maximum tolerated dosage; fathead minnow LC50; *Daphnia magna* EC50; VlogP; ocular irritation; inhalational LC50; and aerobic biodegradability. Analysis of the calculated properties and developed QSTR models shows

that the designed inhibitors satisfy the drug-likeness and ADMET requirements.

## CONCLUSIONS

A new fragment-based lead design method constraining the binding properties of inhibitors and chemotype specificity of binding pockets in the two DFG-loop conformations has been presented. Correlation of the study of sequence, structure, and active site of both type I and II inhibitors from



Legend: b) Type II designed inhibitor: Head; Linker; Tail

**Figure 7.** Binding modes of the two types of p38 MAP kinase inhibitors: (a) Type I, ATP site binders and (b) Type II, designed inhibitors. The figure illustrates: (a) ATP site inhibitor interacting with hinge region; the F169 in the DFG-in conformation blocking the hydrophobic pocket I and II. (b) The designed type II inhibitor comfortably occupies the larger area of the active site created on the 1200 turn of F169 leading to the transition into DFG-out conformation and opening the hydrophobic back pockets I and II. A small M109–G110 flip is also observed. The designed type II p38 MAP kinase inhibitor interacts with the hinge region (head), DFG-loop (linker), and hydrophobic pocket (tail).

**Table 4.** ADMET Properties of the Best 64 Designed p38 MAP Kinase Inhibitors Calculated with TOPKAT Module of Discovery Studio 2.1<sup>a</sup>

S. no.	molecule	HT <sup>b</sup>	CYP P450 <sup>c</sup>	PPB <sup>d</sup>	HIA <sup>e</sup>	Sol. <sup>f</sup>	BBB <sup>g</sup>	S. no.	molecule	HT <sup>b</sup>	CYPP450 <sup>c</sup>	PPB <sup>d</sup>	HIA <sup>e</sup>	Sol. <sup>f</sup>	BBB <sup>g</sup>
1	P1_a	1	0	0	2	2	2	33	P5_a	0	0	0	2	2	0
2	P1_b	0	0	2	2	1	4	34	P5_b	0	0	2	2	2	2
3	P1_c	1	0	2	3	2	3	35	P5_c	0	1	2	3	2	2
4	P1_d	0	0	2	1	1	4	36	P5_d	0	1	2	1	2	2
5	P1_e	0	0	1	2	2	4	37	P5_e	0	0	1	2	2	1
6	P1_f	1	0	0	2	2	2	38	P5_f	0	0	0	2	2	0
7	P1_g	1	0	2	2	1	4	39	P5_g	0	0	2	2	1	2
8	P1_h	0	0	2	2	2	1	40	P5_h	0	0	0	2	2	0
9	P2_a	0	1	0	2	2	4	41	P6_a	1	0	1	2	1	1
10	P2_b	1	1	2	1	2	2	42	P6_b	1	0	2	1	1	2
11	P2_c	1	1	2	3	2	2	43	P6_c	1	0	1	3	1	1
12	P2_d	0	1	1	2	2	3	44	P6_d	0	0	1	2	1	1
13	P2_e	0	1	1	2	2	3	45	P6_e	1	1	1	2	2	1
14	P2_f	0	1	2	1	2	3	46	P6_f	1	0	1	1	2	1
15	P2_g	1	1	2	2	1	4	47	P6_g	1	0	2	2	1	2
16	P2_h	1	1	0	1	2	4	48	P6_h	1	0	1	1	1	1
17	P3_a	0	0	0	2	2	2	49	P7_a	0	0	0	2	2	0
18	P3_b	1	0	2	2	1	4	50	P7_b	1	0	2	2	2	2
19	P3_c	0	0	2	2	2	1	51	P7_c	1	0	2	2	2	2
20	P3_d	1	0	2	1	1	4	52	P7_d	1	0	1	1	2	1
21	P3_e	0	0	1	2	2	4	53	P7_e	0	1	1	2	2	1
22	P3_f	1	0	0	2	2	2	54	P7_f	0	0	0	2	2	0
23	P3_g	1	0	2	2	1	4	55	P7_g	1	0	2	2	1	2
24	P3_h	0	0	2	2	2	4	56	P7_h	0	0	1	2	2	1
25	P4_a	1	0	0	3	2	3	57	P8_a	1	1	0	3	2	0
26	P4_b	1	0	2	3	2	2	58	P8_b	1	1	2	3	2	2
27	P4_c	0	0	1	4	2	4	59	P8_c	1	1	0	4	2	0
28	P4_d	1	0	0	2	2	2	60	P8_d	1	1	2	2	2	2
29	P4_e	0	1	2	0	2	4	61	P8_e	0	1	2	0	2	2
30	P4_f	0	1	0	3	2	2	62	P8_f	1	1	0	3	2	0
31	P4_g	1	0	2	2	1	4	63	P8_g	1	0	2	2	1	2
32	P4_h	1	0	0	3	2	4	64	P8_h	1	0	0	3	2	0

<sup>a</sup> ADMET properties: hepato toxicity (HT), cytochrome P450 inhibition (CYP P450), plasma protein binding (PPB), human intestinal absorption (HIA), aqueous solubility (Sol.), and blood brain barrier (BBB). <sup>b</sup> HT: 0 (nontoxic); 1(toxic). <sup>c</sup> CYPP450: 0 (noninhibitor); 1 (inhibitor). <sup>d</sup> PPB: 0 (binding <90%); 1 (binding ≥90%); 2 (binding ≥95%). <sup>e</sup> HIA: 0 (good); 1 (moderate); 2 (low); 3 (very low). <sup>f</sup> Sol.: 0 (extremely low [ $\log S_w < -8.0$ ]); 1 (low [ $-6.0 < \log S_w < -4.0$ ]); 2 (good [ $-4.0 < \log S_w < -2.0$ ]); 3 (optimal [ $-2.0 < \log S_w < -8.0$ ]); 4 (too soluble [ $\log S_w < -8.0$ ]). <sup>g</sup> BBB: 0 (very high [ $\geq 0.7$ ]); 1 (high [ $>0.7$ ]); 2 (medium ( $<0$ )); 3 (low [ $\leq -0.52$ ]); 4 (undefined).

PDB and FDA drugs has been used to decipher the design of new type II leads from type I. Kinase specificity is as important as efficacy in developing drug candidates, and thus designing type II inhibitors with optimal binding to allosteric site is highly important. The availability of huge number of

crystal structures of p38 MAP kinase in different conformations provides a basis to understand the structural requirement. Sequence analysis reveals residues T180 and V182 in the activation loop have high propensity for mutation, and therefore interactions with them should not be taken as

chemotype specific. Structure analysis delineates the presence of breaks in regions flanking the DFG-loop. Study of active site and binding interactions reveals that a fragment has a tendency to occupy the same part of the active site irrespective of the type of scaffold it constitutes. The detailed study of these structures proved useful not only for the formulation of fragment library but also for the detailed evaluation of structure-based approaches in lead identification, optimization, and virtual screening. The 64 new leads have been generated targeting optimal binding to both ATP and allosteric sites along with a linker binding to the DFG-loop. The eight prioritized allosteric site binders are specific for p38 MAP kinase, however the eight ATP site binding fragments can be picked up as the head portion for any other kinase since the ATP site is conserved. The idea of linking ATP site binder (for inhibition) with allosteric site binder (for selectivity) can be effectively used for the design of kinase inhibitors with increased specificity. The current fragment library generation in combination with other structure-based approaches can be employed as a general strategy for new lead generation not only for kinases but also for other targets.

#### ACKNOWLEDGMENT

This paper is dedicated to Dr. J. S. Yadav on the occasion of his 60<sup>th</sup> birthday. DST is thanked for the Swarnajayanti Fellowship to G.N.S. and the Women Scientist Fellowship to P.B. DBT is thanked for financial assistance.

**Supporting Information Available:** Tables S1–S5 provide details of structure and sequence; analysis and evaluation of docking score components of GOLD, Glide, and CDOCKER in terms of rmsd, Score, and DFG-loop has been given in Table S9. The fragment library generated from cocrystals and FDA drugs, binding interactions along with details of the nature and the frequency of occurrence of individual fragments has been incorporated in Tables S6–S8. Table S10 contains the QSTR models generated. This information is available free of charge via the Internet at <http://pubs.acs.org>.

#### REFERENCES AND NOTES

- Mannig, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The protein kinase complement of the human genome. *Science* **2002**, *298*, 1912–1934.
- Kulkarni, R.; Achaiah, G.; Sastry, G. N. Novel targets for anti-inflammatory and anti-arthritis agents. *Curr. Pharm. Des.* **2006**, *12*, 2437–2454.
- Noble, M. E. M.; Endicott, J.; Johnson, L. K. Protein kinase inhibitors: insights into drug design from structure. *Science* **2004**, *303*, 1800–1805.
- Morphy, R. Selectively Nonselective Kinase Inhibition: Striking the Right Balance. *J. Med. Chem.* **2010**, *53*, 1413–1437.
- Simard, J. R.; Grutter, C.; Pawar, V.; Aust, B.; Wolf, A.; Rabiller, M.; Wulfert, S.; Robubi, A.; Kluter, S.; Ottmann, C.; Rauh, D. High-throughput screening to identify inhibitors which stabilize inactive kinase conformations in p38 $\alpha$ . *J. Am. Chem. Soc.* **2009**, *131*, 18478–18488.
- Vulpetti, A.; Bosotti, R. Sequence and structural analysis of kinase ATP pocket residues. *Farmaco* **2004**, *59*, 759–765.
- Huse, M.; Kuriyan, J. The conformational plasticity of protein kinases. *Cell* **2002**, *109*, 275–282.
- Smyth, L. A.; Collins, I. Measuring and interpreting the selectivity of protein kinase inhibitors. *J. Chem. Biol.* **2009**, *2*, 31–151.
- Jänne, P. A.; Gray, N.; Settleman, J. Factors underlying sensitivity of cancers to small-molecule kinase inhibitors. *Nat. Rev. Drug Discovery* **2009**, *8*, 709–723.
- Angell, R. M.; Angell, T. D.; Bamborough, P.; Brown, D.; Brown, M.; Buckton, J. B.; Cockerill, S. G.; Edwards, C. D.; Jones, K. L.; Longstaff, T.; Snee, P. A.; Smith, K. J.; Somers, D. O.; Walker, A. L.; Willson, M. Biphenyl amide p38 kinase inhibitors 2: optimization and SAR. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 324–328.
- Mol, C. D.; Fabbro, D.; Hosfield, D. J. Structural insights into the conformational selectivity of STI-571 and related kinase inhibitors. *Curr. Opin. Drug Discovery Dev.* **2004**, *7*, 639–648.
- Schindler, T.; Bornmann, W.; Pellicena, P.; Miller, W. T.; Clarkson, B.; Kuriyan, J. Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. *Science* **2000**, *289*, 1938–1942.
- Pargellis, C.; Tong, L.; Churchill, L.; Cirillo, P. F.; Gilmore, T.; Graham, A. G.; Grob, P. M.; Hickey, E. R.; Moss, N.; Pav, S.; Regan, J. Inhibition of p38 MAP kinase by utilizing a novel allosteric binding site. *Nat. Struct. Biol.* **2002**, *9*, 268–272.
- Wan, P. T.; Garnett, M. J.; Roe, S. M.; Lee, S.; Niculescu, D. D.; Good, V. M.; Jones, C. M.; Marshall, C. J.; Springer, C. J.; Barford, D.; Marais, R. Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell* **2004**, *116*, 855–867.
- Manley, P. W.; Cowan, J. S. W.; Mestan, J. Advances in the structural biology, design and clinical development of VEGF-R kinase inhibitors for the treatment of angiogenesis. *Biochim. Biophys. Acta* **2004**, *1697*, 17–27.
- Gill, A. L.; Frederickson, M.; Cleasby, A.; Woodhead, S. J.; Carr, M. G.; Woodhead, A. J.; Walker, M. T.; Congreve, M. S.; Devine, L.; Tisi, A. D.; O'Reilly, M.; Seavers, L. C. A.; Davis, D. J.; Curry, J.; Anthony, R.; Padova, A.; Murray, C. W.; Carr, R. A. E.; Jhoti, H. Identification of novel p38 $\alpha$  MAP kinase inhibitors using fragment based lead generation. *J. Med. Chem.* **2005**, *48*, 414–426.
- Cumming, J. G.; et al. Novel, potent and selective anilinoquinazoline and anilinopyrimidine inhibitors of p38 MAP kinase. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5389–5394.
- Heron, N. M.; Anderson, M.; Blowers, D. P.; Breed, J.; Eden, J. M.; Green, S.; Hill, G. B.; Johnson, T.; Jung, F. H.; McMiken, H. H.; Mortlock, A. A.; Pannifer, A. D.; Paupit, R. A.; Pink, J.; Roberts, N. J.; Rowsell, S. SAR and inhibitor complex structure determination of a novel class of potent and specific Aurora kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1320–1323.
- Kumar, S.; Boehm, J.; Lee, J. C. p38 MAP kinases: key signalling molecules as therapeutic targets for inflammatory diseases. *Nat. Rev. Drug Discovery* **2003**, *2*, 717–726.
- Schindler, J. F.; Monahan, J. B.; Smith, W. G. p38 pathway kinases as anti-inflammatory drug targets. *J. Dent. Res.* **2007**, *86*, 800–811.
- Imajo, M.; Tsuchiya, Y.; Nishida, E. Regulatory mechanisms and functions of map kinase signaling pathways. *IUBMB Life* **2006**, *58*, 312–317.
- Badrinarayan, P.; Srivani, P.; Sastry, G. N. Design of 1-arylsulfamido-2-alkylpiperazine derivatives as secreted PLA2 inhibitors. *J. Mol. Model.* **2010**, DOI: 10.1007/s00094-010-0752-2.
- Kulkarni, R. G.; Srivani, P.; Achaiah, G.; Sastry, G. N. Strategies to design of pyrazolyl urea derivatives for p38 kinase inhibition: A molecular modeling study. *J. Comput.-Aided Mol. Des.* **2007**, *25*, 155–166.
- Kulkarni, R. G.; Achaiah, G.; Sastry, G. N. Molecular modeling studies of phenoxypryimidinyl imidazoles as p38 kinase inhibitors using QSAR and docking. *Eur. J. Med. Chem.* **2008**, *43*, 830–838.
- Srivani, P.; Srinivas, E.; Raghu, R.; Sastry, G. N. Molecular modeling studies of pyridopurinone derivatives - Potential phosphodiesterase 5 Inhibitors. *J. Mol. Graphics Modell.* **2007**, *26*, 378–390.
- Srivani, P.; Usharani, D.; Jemmis, E. D.; Sastry, G. N. Subtype selectivity in phosphodiesterase 4 (PDE4): A bottleneck in rational drug design. *Curr. Pharm. Des.* **2008**, *14*, 3854–3872.
- Romeiro, N. C.; Albuquerque, M. G.; Alencastro, R. B.; Ravi, M.; Hopfinger, A. J. Free-energy force-field three-dimensional quantitative structure-activity relationship analysis of a set of p38-mitogen activated protein kinase inhibitors. *J. Mol. Model.* **2006**, *12*, 855–868.
- Romeiro, N. C.; Albuquerque, M. G.; Alencastro, R. B.; Ravi, M.; Hopfinger, A. J. Construction of 4D-QSAR models for use in the design of novel p38-MAPK inhibitors. *J. Comput.-Aided Mol. Des.* **2005**, *19*, 385–400.
- Reddy, A. S.; Pati, S. P.; Kumar, P. P.; Pradeep, H. N.; Sastry, G. N. Virtual screening in drug discovery - A computational perspective. *Curr. Protein Pept. Sci.* **2007**, *8*, 329–351.
- Xiao, Z.; Varma, S.; Xiao, Y. D.; Tropsha, A. Modeling of p38 mitogen-activated protein kinase inhibitors using the CatalystTM HypoGen and k-nearest neighbor QSAR methods. *J. Mol. Graphics Modell.* **2004**, *23*, 129–138.
- Chen, J.; Zhang, Z.; Stebbins, J. L.; Zhang, X.; Hoffman, R.; Moore, A.; Pellecchia, M. A Fragment-based approach for the discovery of isoform-specific p38 inhibitors. *ACS Chem. Biol.* **2007**, *2*, 329–336.

- (32) Natarajan, S. R.; Heller, S. T.; Nam, K.; Singh, S. B.; Scapin, G.; Patel, S.; Thompson, J. E.; Fitzgerald, C. E.; Keefeb, S. J. p38 MAP kinase inhibitors. Part 6: 2-Arylpyridazin-3-ones as templates for inhibitor design. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 5809–5813.
- (33) Regan, J.; Capolino, A.; Cirillo, P. F.; Gilmore, T.; Graham, A. G.; Hickey, E.; Kroenke, R. R.; Madwed, J.; Moriak, M.; Nelson, R.; Pargellis, C. A.; Swinamer, A.; Torcellini, C.; Tsang, M.; Moss, N. Structure-activity relationships of the p38r MAP kinase inhibitor 1-(5-tert-Butyl-2-p-tolyl-2H-pyrazol-3-yl)-3-[4-(2-morpholin-4-yl-ethoxy) naphthalen-1-yl]urea (BIRB 796). *J. Med. Chem.* **2003**, *46*, 4676–4686.
- (34) Getlik, M.; Grutter, C.; Simard, J. R.; Kluter, S.; Rabiller, M.; Rode, H. B.; Robubi, A.; Rauh, D. Hybrid compound design to overcome the gatekeeper T338M mutation in cSrc#. *J. Med. Chem.* **2009**, *52*, 3915–3926.
- (35) Collins, I.; Workman, P. Design and development of signal transduction inhibitors for cancer treatment: Experience and challenges with kinase targets. *Curr. Signal Transduction Ther.* **2006**, *1*, 13–23.
- (36) Davidson, W.; Frego, L.; Peet, G. W.; Kroenke, R. R.; Labadia, M. E.; Lukas, S. M.; Snow, R. J.; Jakes, S.; Grygon, C. A.; Pargellis, C.; Werneburg, B. G. Discovery and Characterization of a substrate selective p38R inhibitor. *Biochemistry* **2004**, *43*, 11658–11671.
- (37) Shana, Y.; Seeliger, M. A.; Eastwood, M. P.; Frankb, F.; Xua, H.; Jensen, M.; Drora, R. O.; Kuriyan, J.; Shawa, D. E. A conserved protonation-dependent switch controls drug binding in the Abl kinase. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 139–144.
- (38) Huang, D.; Zhou, T.; Lafleur, K.; Nevado, C.; Caflisch, A. Kinase selectivity potential for inhibitors targeting the ATP binding site: a network analysis. *Bioinformatics* **2010**, *26*, 198–204.
- (39) Soliva, R.; Gelpí, J. L.; Almansa, C.; Virgili, M.; Orozco, M. Dissection of the Recognition properties of p38 MAP kinase. Determination of the binding mode of a new pyridinyl-heterocycle inhibitor family. *J. Med. Chem.* **2007**, *50*, 283–293.
- (40) Vieth, M.; Higgs, R. E.; Robertson, D. H.; Shapiro, M.; Gragg, E. A.; Hemmerle, H. Kinomics—structural biology and chemogenomics of kinase inhibitors and targets. *Biochim. Biophys. Acta* **2004**, *1697*, 243–257.
- (41) Dey, F.; Caflisch, A. Fragment-based *de novo* ligand design by multiobjective evolutionary optimization. *J. Chem. Inf. Mod.* **2008**, *48*, 679–690.
- (42) Musafia, B.; Senderowitz, H. Bioactive conformational biasing: a new method for focusing conformational ensembles on bioactive-like conformers. *J. Chem. Inf. Mod.* **2009**, *49*, 2469–2480.
- (43) Sutherland, J. J.; Higgs, R. E.; Watson, I.; Vieth, M. Chemical fragments as foundations for understanding target space and activity prediction. *J. Med. Chem.* **2008**, *51*, 2689–2700.
- (44) Hartshorn, M. J.; Murray, C. W.; Cleasby, A.; Frederickson, M.; Tickle, I. J.; Jhoti, H. Fragment-based lead discovery using x-ray crystallography. *J. Med. Chem.* **2005**, *48*, 403–413.
- (45) Zhu, K.; Shirts, M. R.; Friesner, R. A. Improved methods for side chain and loop predictions via the protein local optimization program: variable dielectric model for implicitly improving the treatment of polarization effects. *J. Chem. Theory Comput.* **2007**, *3*, 2108–2119.
- (46) Fiser, A.; Do, R. K.; Sali, A. Modeling of loops in protein structures. *Protein Sci.* **2000**, *9*, 1753–1773.
- (47) Lovell, S. C.; Davis, I. W.; Arendall III, W. B.; de Bakker, P. I. W.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. *Proteins: Struct., Funct., Genet.* **2002**, *50*, 437–450.
- (48) Wiederstein, M.; Sippl, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **2007**, *35*, W407–W410.
- (49) Heinig, M.; Fishman, D. STRIDE a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res.* **2004**, *32*, W500–W502.
- (50) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489.
- (51) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
- (52) 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.
- (53) Wu, G.; Robertson, D. H.; Brooks III, C. L.; Vieth, M. Detailed analysis of grid-based molecular docking: a case study of CDOCKER - CHARMM- based MD docking algorithm. *J. Comput. Chem.* **2003**, *24*, 1549–1562.
- (54) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (55) 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.
- (56) Zuccotto, F.; Ardin, E.; Casale, E.; Angiolini, M. Through the “Gatekeeper door”: exploiting the active kinase conformation. *J. Med. Chem.* **2010**, *53*, 2681–2694.
- (57) Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R. Fragment-based lead discovery. *Nat. Rev. Drug Discovery* **2004**, *3*, 660–672.
- (58) Congreve, M.; Chessari, G.; Tisi, D.; Woodhead, A. J. Recent developments in fragment-based drug discovery. *J. Med. Chem.* **2008**, *51*, 3661–3680.
- (59) Chen, J.; Zhang, Z.; Stebbins, J. L.; Zhang, X.; Hoffman, R.; Moore, A.; Pellecchia, M. A fragment-based approach for the discovery of isoform-specific p38 inhibitors. *ACS Chem. Bio.* **2007**, *2*, 329–336.
- (60) Bamborough, P.; Drewry, D.; Harper, G.; Smith, G. K.; Schneider, K. Assessment of chemical coverage of kinase space and its implications for kinase drug discovery. *J. Med. Chem.* **2008**, *51*, 7898–7914.
- (61) Sutherland, J. J.; Higgs, R. E.; Watson, I.; Vieth, M. Chemical fragments as foundations for understanding target space and activity prediction. *J. Med. Chem.* **2008**, *51*, 2689–2700.
- (62) Davis, A. M.; Teague, S. J. Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 736–749.
- (63) Teague, S. J. Implications of protein flexibility for drug discovery. *Nat. Rev. Drug Discovery* **2003**, *2*, 527–541.
- (64) Bakan, A.; Bahar, I. The intrinsic dynamics of enzymes plays a dominant role in determining the structural changes induced upon inhibitor binding. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106* (34), 14349–14354.
- (65) Craig, I. R.; Essex, J. W.; Spiegel, K. Ensemble docking into multiple crystallographically derived protein structures: an evaluation based on the statistical analysis of enrichments. *J. Chem. Inf. Model.* **2010**, *50*, 511–524.
- (66) Huang, S. Y.; Zou, X. Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking. *Proteins: Struct., Funct., Bioinf.* **2007**, *66*, 399–421.
- (67) Srinivas, E.; Murthy, J. N.; Rao, A. R.; Sastry, G. N. Recent advances in molecular modeling and medicinal chemistry aspects of phosphoglycoprotein. *Curr. Drug Metab.* **2006**, *7*, 205–217.

CI100340W