J Comput Aided Mol Des (2012) 26:409–423
DOI 10.1007/s10822-012-9569-7

Virtual screening using a conformationally flexible target protein: models for ligand binding to p38α MAPK

Natalie B. Vinh · Jamie S. Simpson · Peter J. Scammells · David K. Chalmers

Received: 20 December 2011 / Accepted: 26 March 2012 / Published online: 20 April 2012

Abstract We have used virtual screening to develop models for the binding of aryl substituted heterocycles to p38α MAPK. Virtual screening was conducted on a number of p38α MAPK crystal structures using a library of 46 known p38α MAPK inhibitors containing a heterocyclic core substituted by pyridine and fluorophenyl rings (structurally related to SB203580) and a set of decoy compounds. Multiple protonation states and tautomers of active and decoy compounds were considered. Each docking model was evaluated using receiver operating characteristic (ROC) curves and enrichment factors. The two best performing single crystal structures were found to be 1BL7 and 2EWA, with enrichment factors of 14.1 and 13.0 at 2 % of the virtual screen respectively. Ensembles of up to four receptors of similar conformations were generated, generally giving good or very good performances with high ROC AUCs and good enrichment. The 1BL7-2EWA ensemble was able to outperform each of its constituent receptors and gave high enrichment factors of 17.3, 12.0, 8.0 at 2, 5 and 10 % respectively, of the virtual screen. A ROC AUC of 0.94 was obtained for this ensemble. This method may be applied to other proteins where there are a large number of inhibitor classes with different binding site conformations.

Keywords p38α Mitogen-activated protein kinase · Virtual screening · Ensemble docking · Enrichment · Receiver operating characteristic (ROC) curves · Protein flexibility

Introduction

p38α mitogen-activated protein kinase (MAPK) is a serine/threonine kinase that plays an essential role in the regulation of pro-inflammatory signaling networks and the biosynthesis of cytokines, including tumor necrosis factor-α and interleukin-1β [1, 2]. These cytokines have been implicated in the development of chronic inflammatory diseases such as rheumatoid arthritis, inflammatory bowel disease, psoriasis and asthma [3, 4]. Small molecules that inhibit p38α MAPK may therefore provide a therapeutic strategy for the treatment of chronic inflammatory diseases and have been the subject of extensive research [5–19].

The p38α MAPK structure consists of *N*- and *C*-terminal domains which form the walls of a deep cleft that binds the coenzyme ATP [20]. Like other protein kinases, the ATP binding site is characterized by five regions including an adenine binding region, a sugar pocket, two hydrophobic regions (I and II) and a phosphate binding region (Fig. 1a) [21]. Activation of p38α MAPK occurs by the dual phosphorylation of residues Thr180 and Tyr182 in the activation loop [1]. A hinge region, which connects the two domains, makes key hydrogen bonding interactions with the adenine ring of ATP [22]. ATP also makes polar interactions with the glycine rich loop and activation loop to stabilize the phosphates and transition states generated during the activation process [22].

A large number of p38α MAPK inhibitors are ATP-competitive. These inhibitors target the active form of the

Electronic supplementary material The online version of this article (doi:10.1007/s10822-012-9569-7) contains supplementary material, which is available to authorized users.

N. B. Vinh · J. S. Simpson · P. J. Scammells · D. K. Chalmers (✉)
Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University (Parkville Campus), 381 Royal Parade, Parkville, VIC 3052, Australia
e-mail: David.Chalmers@monash.edu

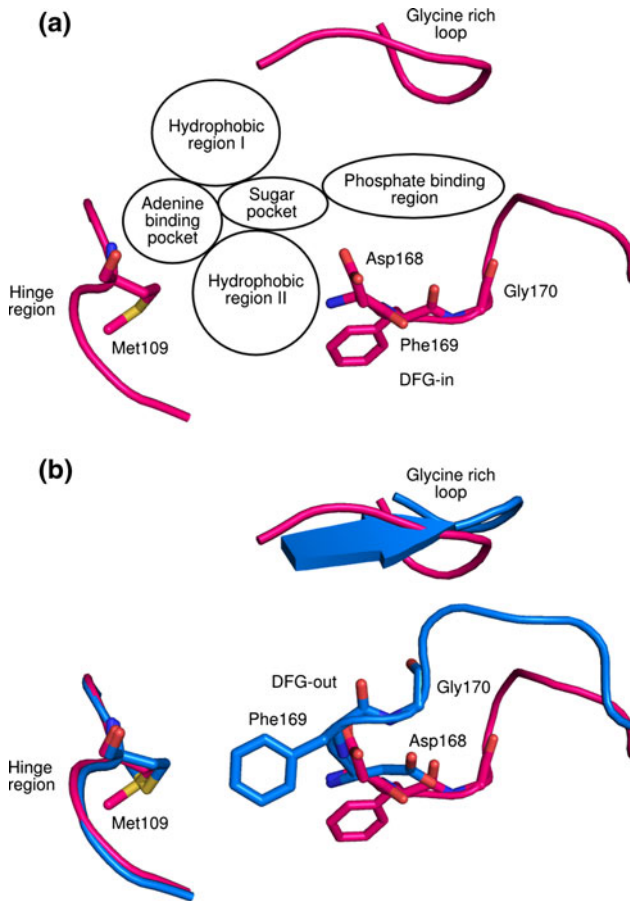

Fig. 1 Schematic of the ATP binding site **a** p38α MAPK adopting the DFG-in conformation **b** p38α MAPK adopting the DFG-out conformation

protein [23]. Within the activation loop of the active site is the conserved Asp168-Phe169-Gly170 (DFG) motif [24]. In its active form, the protein adopts a "DFG-in" conformation (Fig. 1a) [23]. However, a number of inhibitors are also known to utilize an allosteric binding site, causing a conformational reorganization of the protein to the "DFG-out" conformation (Fig. 1b) [25, 26]. In this state the Phe169 side chain of the DFG motif moves ~10 Å leaving a deep hydrophobic pocket within the protein [25–27].

Our research focuses on the design and development of small molecule inhibitors of p38α MAPK using a structure-based approach. We are particularly interested in one class of p38α MAPK inhibitors, exemplified by SB203580 (Fig. 2) [28], which possess a heterocyclic core containing pyridine and fluorophenyl substituents.

X-ray crystallography is the prime source of structural information for most structure-based molecular design projects. Crystal structures of protein–ligand complexes reveal the key interactions made between the protein and ligand which can then be exploited to design new compounds that improve existing interactions or gain new

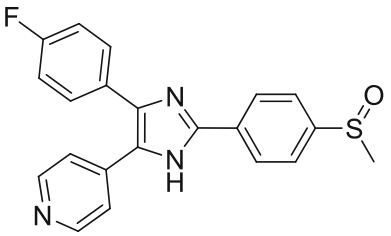

Fig. 2 Structure of SB203580

interactions and have better affinity and/or specificity [29]. However, limitations associated with protein X-ray crystal structures provide reason for caution to be taken when utilizing them in drug design. Uncertainty associated with crystallography arises from a number of sources, including the quality of the diffraction data, the quality of the model generated to fit the data, and importantly the unclear relationship between the crystallized protein conformation and the biologically relevant conformation [30, 31]. Since most protein crystals are not determined to full atomic resolution (1.2 Å) there is a degree of uncertainty in the atomic positions in all protein structural models. Significant uncertainties often arise when regions of the proteins are highly flexible or exist in multiple conformations causing difficulties in interpretation of the electron density map [31]. Additionally, crystal structures do not represent a flexible protein in the solution state, but in a static conformation driven by the crystallization conditions, in which the protein may be perturbed by the crystallization environment and crystal packing. Consequently, some areas of the protein may adopt different conformations in the crystal than what would be observed from the protein in its native environment.

Protein conformational flexibility and uncertainty about the biological relevance of particular protein–ligand complexes is a significant problem in the structure-based design of MAPK inhibitors. In December 2011, there were 175 crystal structures of p38α MAPK in the protein data bank (PDB); all of which provide evidence of ligand-induced conformational changes to the protein. This makes p38α MAPK a challenging target to model. Nevertheless, to assist our research, we wished to identify a structure, or structures, which could reliably model the binding of small molecule ligands to the protein. Our problem then became, which crystal structures provide the best models? Simple inspection is not sufficient to select models for docking and it is not clear that structures containing the most similar ligands are the best choice. There is need for a more rigorous assessment.

In this work we have used virtual screening to assess which crystal structure should be used as a model to accurately dock ligands into the protein. Virtual screening involves searching a database of compounds to identify

those that fit into the active site of the target protein [32]. Using databases containing known p38α MAPK inhibitors and drug-like decoys, a virtual screen can be used to assess a crystal structure on its ability to identify bioactives within the database.

Virtual screening has led to the identification of many actives in drug discovery projects. In the case of p38α MAPK, Cheeseright et al. [33] used virtual screening to identify bioactives within a database of commercially available compounds. Armen et al. [34] used molecular dynamics simulations to sample protein conformations and combined their model with ensemble docking for enrichment studies. In another study Rao et al. [35] also utilized ensemble docking to account for structural flexibility. Their research found that a small fraction of ensembles of up to three receptor structures of p38α MAPK gave improved enrichment of actives in a database screen [35]. However, they also noted that finding the right ensemble of receptor conformations is difficult and that ensembles consisting of structurally diverse conformations may give good or poor enrichment compared to ensembles consisting of similar conformations [35]. Unlike most virtual screening methods that aim to identify a novel scaffold for drug discovery, we have used virtual screening to identify a model most suitable for docking a certain class of inhibitor, in our case, compounds containing a heterocyclic core with pyridine and fluorophenyl substituents.

Methods

Molecular modeling was primarily carried out using the Schrödinger software suite (Schrödinger LLC, New York, NY, USA). Protein preparation and refinement was conducted using Maestro v9.1 and v9.2 (Schrödinger) and PyMOL (www.pymol.org). Ligands were prepared using LigPrep v2.4 (Schrödinger). Docking was performed using Glide v5.6 (Schrödinger). Physicochemical properties were predicted using QikProp v3.4 (Schrödinger).

Crystal structure preparation

Forty-four p38α MAPK crystal structures (1A9U, 1BL6, 1BL7, 1BMK, 1DI9, 1KV1, 1M7Q, 1OUK, 1OUY, 1OVE, 1OZ1, 1W7H, 1W82, 1W83, 1W84, 1WBN, 1WBS, 1WBT, 1WBV, 1YQJ, 1ZZ2, 1ZZL, 2BAK, 2BAL, 2EWA, 2ZB0, 3CTQ, 3D7Z, 3E92, 3GFE, 3HUC, 3IPH, 3FL4, 3KF7, 3KQ7, 3FMK, 3FMM, 3LFF, 3MPT, 3NNU, 3NNV, 3NNW, 3NNX, 3NWW) [9, 13, 16, 25, 26, 36–55] were obtained from the PDB [56]. These included proteins with the DFG-in, DFG-out and hinge region peptide flip conformations. The crystal structures were visually assessed to check that binding site residues were resolved and that there were no mutations in or near the active site. Structures were aligned using PyMOL to 1A9U which is crystallized with SB203580. All small molecules and water molecules were removed. Maestro Protein Preparation Wizard default parameters were used for the addition of hydrogens, assignment of bond orders and partial charges. Prime v2.2 (Schrödinger) was used to add side chains to residues with missing atoms. Each protein was refined using exhaustive sampling, in which a sequence of restrained partial minimizations were used to optimize the orientation of the added hydrogens and relieve any strain due to steric clashes in the crystal structure.

Ligand preparation

Crystal structure ligands were prepared for docking by manually assigning bond orders and using the Maestro Protein Preparation Wizard for the addition of hydrogens.

A set of 46 p38α MAPK inhibitors (Fig. 3) was selected from a review of small molecular anti-cytokine agents by Wagner and Laufer [57]. Actives were selected to be structurally similar to SB203580 and to possess characteristics of the prototypical p38α MAPK inhibitor, generally bearing pyridine and fluorophenyl substituents. These p38α MAPK inhibitors have been tested in a variety of assays (Table 1) and therefore we were unable to make direct comparisons of docking scores with activity data of inhibitors. The active compounds were built in ChemBio-Draw Ultra 12.0 (CambridgeSoft) and prepared using LigPrep applying default settings for conversion to the three-dimensional (3D) structure. All ligands in the active set contain heterocyclic structures that can exist in multiple tautomeric and protonation states. To account for this, we used LigPrep to generate multiple tautomers and/or protonation states for each compound where low energy ionization and tautomeric states compatible with the pH range 7.0 ± 2.0 were generated and structures were minimized using the OPLS_2005 force field. Generally between 2 to 16 states were generated for each ligand. In total 197 tautomer/protonation states for the 46 actives were used in the virtual screen.

A set of 1,000 drug-like decoy compounds with an average molecular weight of 400 g/mol was obtained from Schrödinger. This set of drug-like compounds was selected from a larger database of approximately one million purchasable compounds and can be freely downloaded from http://www.schrodinger.com [58, 59]. This set was chosen because they are likely to reflect the size of our synthetic targets and has also been previously applied to kinases [59] and G-protein coupled receptors [60]. Although the affinity of the decoy compounds for p38α MAPK is unknown, the decoys are structurally dissimilar to the 46 active compounds. As such, the decoys do not possess the common

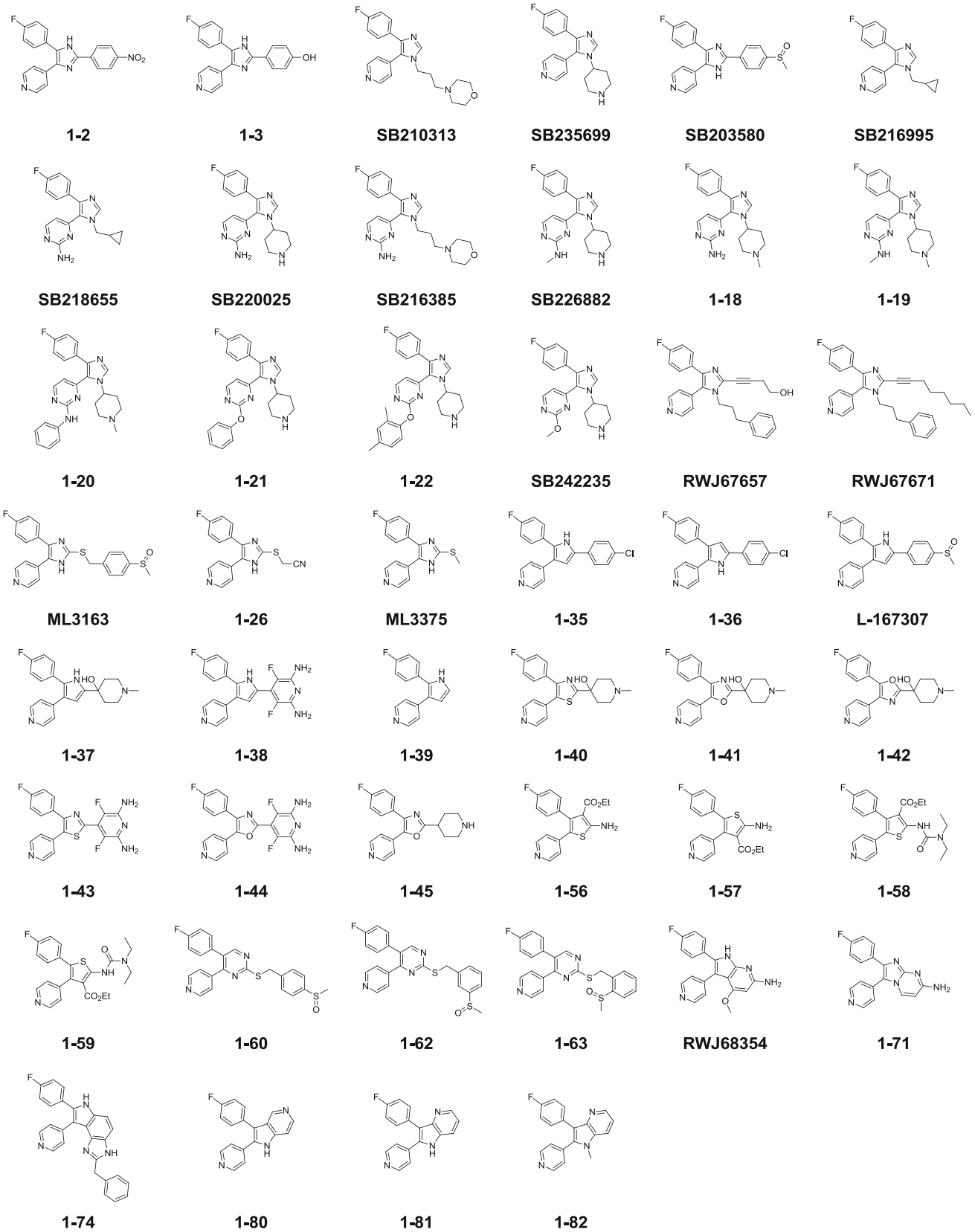

Fig. 3 Structures of known p38α MAPK inhibitors selected as the active compound set. *Numbering* is according to that used in the review by Wagner and Laufer [57]

Table 1 Inhibitory activity values for the p38α MAPK actives used in the database screen

Active	IC_{50} p38α inhibition	Active	IC_{50} p38α inhibition	IC_{50} TNF-α inhibition
SB203580[a]	74 nM[h]	1–36[d]	1.40 μM[k]	–
1–2[a]	89 nM[h]	L-167307[d]	5.1 nM[k]	–
1–3[a]	67 nM[h]	1–37[e]	0.13 μM[l]	–
SB210313[a]	1.3 μM[h]	1–38[e]	4 nM[l]	–
SB235699[a]	60 nM[h]	1–39[e]	0.80 μM[l]	–
SB216995[a]	0.16 μM[h]	1–40[e]	0.45 μM[l]	–
SB218655[a]	25 nM[h]	1–41[e]	0.35 μM[l]	–
SB220025[a]	19 nM[h]	1–42[e]	2.0 μM[l]	–
SB216385[a]	0.48 μM[h]	1–43[e]	0.14 μM[l]	–
SB226882[a]	27 nM[h]	1–44[e]	0.40 μM[l]	–
1–18[a]	86 nM[h]	1–45[e]	15 nM[l]	–
1–19[a]	0.10 μM[h]	1–56[f]	–	56 μM[m]
1–20[a]	3 nM[h]	1–57[f]	–	55 μM[m]
1–21[a]	19 nM[h]	1–58[f]	–	2.6 μM[m]
1–22[a]	6 nM[h]	1–59[f]	–	1.7 μM[m]
SB242235[a]	19 nM[h]	1–60[c]	36 μM[n]	–
SB203580[a]	79 nM[i]	1–62[c]	19 μM[n]	–
RWJ67657[b]	30 nM[i]	1–63[c]	5.1 μM[n]	–
RWJ67671[b]	1.7 μM[i]	RWJ68354[b]	1.5 μM[o]	7 nM[p]
SB203580[a]	0.29 μM[j]	1–71[b]	570 nM[o]	40 nM[p]
ML3163[c]	4.0 μM[j]	1–74[b]	22 nM[o]	60 nM[p]
1–26[c]	0.40 μM[j]	1–80[g]	3.1 μM[q]	–
ML3375[c]	0.63 μM[j]	1–81[g]	6.5 nM[q]	–
SB203580[a]	39 nM[k]	1–82[g]	98 nM[q]	–
1–35[d]	0.20 μM[k]			

Data is taken from Wagner and Laufer [57]

Direct comparisons can only be made between activities measured in the same assay. Some assays quote the activity of SB203580 as a control allowing general comparisons to be made between different assays

Compounds developed by: [a] GSK, [b] RW Johnson, [c] University of Tübingen, [d] Merck, [e] Novartis, [f] Nikken, [g] Roche. [h] Inhibition of activated mammalian p38α MAPK [36, 65–69]. [i] Inhibition of activated p38α MAPK [70]. [j] Inhibition of activated p38α MAPK [71]. [k] Inhibition of activated p38α MAPK [72]. [l] Inhibition of activated murine His-p38α MAPK [73–75]. [m] Inhibition of LPS-stimulated TNF-α production in rat whole blood [76]. [n] Inhibition of activated p38α MAPK [77]. [o] Inhibition of activated mouse p38α MAPK [78]. [p] Inhibition of LPS-stimulated TNF-α production from human peripheral blood mononuclear cells (PBMCs) [78]. [q] Inhibition of human recombinant activated p38α MAPK [9]

structural features required for binding to p38α MAPK and therefore are unlikely to be actives themselves. If a small number of decoy compounds do, in fact inhibit p38α MAPK then this should be reflected in all the crystal structures tested and should not unduly affect the relative ranking of crystal structures. The drug-like decoy set was prepared in the same way as the active compounds, including generation of multiple tautomeric forms and ionization states. This resulted in 4,041 tautomer/protonation states for the 1,000 decoys.

Cognate ligand docking

Crystal structure ligands were redocked into the p38α MAPK crystal structure using Glide extra precision (XP) method. RMSD calculations were made using the superposition tool in Maestro, in which the docked pose was superimposed onto the crystal structure ligand conformation using substructure recognition SMARTS [61].

Virtual screening

The final virtual screening database consisted of 4,238 compounds (197 actives and 4,041 decoys) including multiple tautomeric and ionization states for each compound. Docking calculations were performed using the Glide XP method which uses a rigid receptor and flexible ligand [62]. The docking site was defined as a 30 × 30 × 30 Å box centered on the average of coordinates of the native ligand present in the DFG-in structures. These coordinates were calculated as x (1.5418), y (12.3467), z (29.6070) in 1A9U. Each compound in the database was docked into the forty-

four crystal structures with one pose per ligand collected. These compounds were ranked and scored using GlideScore. Only the best scoring (lowest GlideScore) tautomer/ionization state of each compound was retained for enrichment studies (46 actives and 1,000 decoys).

Analysis of model performance

Receiver operating characteristic (ROC) curves and enrichment factors were calculated to compare the performance of each crystal structure. In a successful virtual screening model, the docking program must rank active ligands ahead of the decoy compounds and should correctly predict the bound ligand structures. ROC curves are a useful tool to characterize the ability of a model to select active compounds and discard inactive compounds [63]. A ROC curve plots the percentage rank of true positives (actives found) on the Y axis against the percentage rank of false positives (decoys found) on the X axis. The area under the ROC curve was used to measure the overall performance of each model. If the plot is a diagonal line, the AUC has a value of 0.5. This curve results from a model that cannot discriminate active and decoy compounds. An ideal performance, where all active compounds are identified before any decoys, gives a curve that rises vertically from (0, 0) to (0, 100) and then runs to (100, 100). This curve has an AUC of 1.0. In addition to having a large AUC, better models will rank active compounds more highly. Therefore enrichment factors were used as another measure of assessing a model's performance. Enrichment factors were calculated at 2, 5 and 10 % of the virtual screen using the following equation [64]:

$$EF^{x\%} = \left(Hits_{sampled}/N_{sampled}\right)/\left(Hits_{total}/N_{total}\right)$$

where Hits refers to the number of actives and N refers to the number of compounds. An enrichment factor of 1.0 indicates that the crystal structure is randomly identifying actives from the database. An enrichment factor less than 1.0 means the crystal structure is performing worse than random. The maximum possible enrichment factor (EFmax) at 2, 5 and 10 % of the screen were calculated as 22.7, 19.7 and 10.0 respectively using the equation:

$$EFmax^{x\%} = \left(\min\left(N_{sampled},\ Hits_{total}\right)/N_{sampled}\right)/\left(Hits_{total}/N_{total}\right)$$

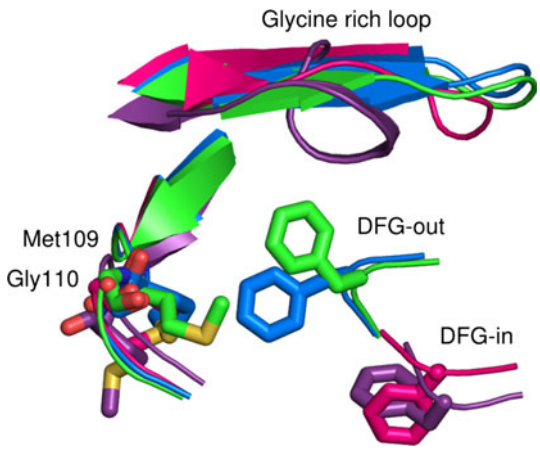

Fig. 5 The different conformations of p38α MAPK considered; DFG-in (*pink*, 1BL7), DFG-in with Met109-Gly110 peptide flip in the hinge region (*purple*, 1OVE), DFG-out (*blue*, 1WBN), DFG-out with Met109-Gly110 peptide flip in the hinge region (*green*, 1WBS)

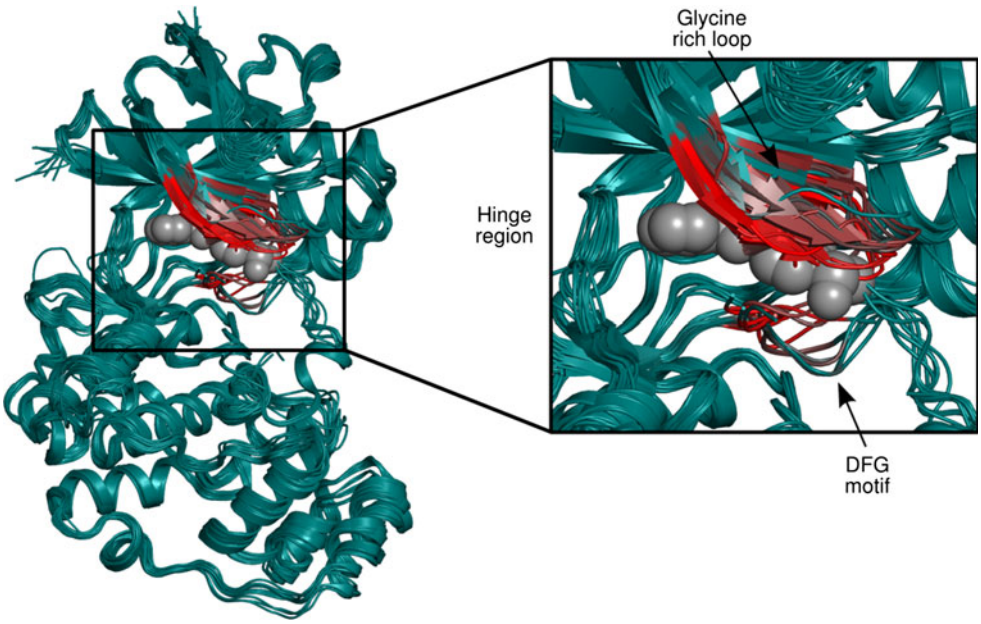

Fig. 4 An overlay of twenty p38α MAPK crystal structures. The glycine rich loop and DFG motif are *highlighted* with a gradient from teal to red with red depicting the greatest deviation from the 1A9U structure (teal)

Table 2 Cognate docking and virtual screening results for forty-four p38α MAPK crystal structures

PDB	Res. (Å)[a]	Set[b]	Ligand in inhibitor class?	Protein RMSD (Å) to 1A9U	Cognate Ligand RMSD (Å)	ROC AUC	EF$^{2\%}$ [c]	EF$^{5\%}$ [d]	EF$^{10\%}$ [e]	Result[f]
1A9U	2.50	1	Yes	0.00	1.50	0.77	7.6	3.4	3.0	Good
1BL6	2.50	1	Yes	0.30	1.59	0.71	3.2	5.1	3.5	Good
1BL7	2.50	1	Yes	0.24	0.64	0.83	**14.1**	9.9	6.1	**V. good**
1BMK	2.40	1	Yes	0.21	0.41	0.64	1.1	0.4	1.7	Poor
1DI9	2.60	1	No	0.41	0.55	0.83	3.2	2.6	2.8	Good
1OUK	2.50	1	No	0.45	0.64	0.86	5.4	4.7	4.8	Good
1OZ1	2.10	1	Yes	0.38	**0.21**	0.93	10.8	9.9	7.6	**V. good**
1W7H	2.21	1	No	0.40	0.43	0.47	0.0	0.0	0.2	Poor
1W84	2.20	1	No	0.33	0.33	0.80	6.5	4.7	3.7	Good
1YQJ	2.00	1	No	0.36	0.69	0.78	5.4	5.6	3.9	Good
2BAL	2.10	1	No	0.73	0.87*	0.84	6.5	6.0	5.2	Good
2EWA	2.10	1	Yes	0.69	0.85	**0.97**	13.0	**12.4**	**8.9**	**V. good**
3D7Z	2.10	1	No	0.34	0.45	0.27	0.0	0.0	0.0	Poor
3FL4	1.80	1	No	0.32	1.41	0.68	2.2	1.7	1.9	Poor
3FMK	1.70	1	No	0.33	1.44	0.89	11.9	9.0	6.1	**V. good**
3FMM	2.00	1	No	0.35	0.44	0.60	1.1	1.7	1.9	Poor
3GFE	2.10	1	No	0.44	**0.21**	0.88	8.7	7.3	6.1	Good
3IPH	2.10	1	No	0.40	0.51	0.77	1.1	2.1	2.6	Poor
3MPT	1.89	1	No	0.29	1.48	0.65	0.0	0.9	1.1	Poor
3NWW	2.09	1	No	0.73	1.69	0.91	6.5	8.9	6.7	Good
1M7Q	2.40	2	No	0.36	1.52	0.81	1.1	3.4	2.8	Poor
1OUY	2.50	2	No	0.41	0.39	0.64	0.0	0.0	1.1	Poor
1OVE	2.10	2	No	0.44	0.46	0.74	0.0	1.3	1.7	Poor
1ZZ2	2.00	2	No	0.86	1.16	0.79	2.2	2.6	2.4	Poor
1ZZL	2.00	2	No	0.40	0.31	0.59	1.1	3.4	3.0	Poor
2ZB0	2.10	2	No	0.29	0.51	0.50	1.1	0.9	1.5	Poor
3E92	2.00	2	No	0.32	0.32	0.44	1.1	0.9	0.6	Poor
3KF7	2.00	2	No	0.89	1.06	0.93	5.4	6.9	6.9	Good
1W83	2.50	3	No	0.50	0.64	0.51	1.1	1.7	1.3	Poor
1WBN	2.40	3	No	0.42	0.73	0.64	2.2	2.6	2.4	Poor
2BAK	2.20	3	No	0.64	1.86*	0.80	4.3	3.4	3.7	Good
3CTQ	1.95	3	No	0.77	0.43	0.46	1.1	0.4	0.9	Poor
3HUC	1.80	3	No	0.62	1.11	0.83	1.1	2.6	3.5	Poor
3KQ7	1.80	3	No	0.22	1.10	0.57	0.0	0.4	0.6	Poor
3NNU	2.40	3	No	0.42	0.90	0.56	1.1	0.9	1.1	Poor
3NNW	1.89	3	No	0.48	0.26*	0.54	2.2	1.3	1.1	Poor
1KV1	2.80	4	No	0.77	0.59	0.44	0.0	0.9	0.6	Poor
1W82	2.20	4	No	0.50	0.40	0.50	2.2	1.7	1.9	Poor
1WBS	1.80	4	No	0.55	0.38	0.53	0.0	0.0	0.4	Poor
1WBT	2.00	4	No	0.47	0.33	0.49	0.0	0.4	0.9	Poor
1WBV	2.00	4	No	0.38	0.49	0.58	1.1	0.4	0.6	Poor
3LFF	1.50	4	No	0.71	0.84	0.48	1.1	1.7	1.7	Poor
3NNV	2.10	4	No	0.36	0.48	0.45	0.0	0.0	0.6	Poor
3NNX	2.28	4	No	0.98	0.83	0.43	1.1	1.7	1.1	Poor

The best docking results are highlighted in bold

[a] Resolution [b] Set 1: DFG-in conformation; Set 2: DFG-in conformation with Met109-Gly110 flip in the hinge region; Set 3: DFG-out conformation; Set 4: DFG-out conformation with Met109-Gly110 flip in hinge region. [c] Maximum enrichment factor at 2 % = 22.7. [d] Maximum enrichment factor at 5 % = 19.7. [e] Maximum enrichment factor at 10 % = 10.0. [f] Poor performance defined as $EF^{2\%}/EFmax^{2\%} \leq 0.10$; good performance as $0.10 < EF^{2\%}/EFmax^{2\%} < 0.40$; very good performance as $EF^{2\%}/EFmax^{2\%} \geq 0.40$. * Second/third best pose used for RMSD calculation

Ensemble docking

An ensemble analysis of the top performing crystal structures was carried out. Ensemble scores were calculated by taking the average of the GlideScore of each ligand in multiple receptors. The similarity of ligand poses in different receptors was measured by calculating RMS deviations between ligand poses in superimposed receptors.

Physicochemical properties

The physicochemical properties of the actives and decoys were predicted using the fast mode in Qikprop (see supplementary information). The main differences between the two sets were that, on average, the decoys contained a larger number of rotatable bonds (decoy mean 2.3, actives mean 5.9) and a higher polar surface area (decoy mean 95 $Å^2$, actives mean 55 $Å^2$). These differences in the average properties of active and decoy sets are identical for all systems studied and are therefore not expected to affect comparisons of virtual screening results made between p38α MAPK crystal structures.

Results and discussion

Our current research focuses on the structure-based design of p38α MAPK inhibitors containing a core heterocycle with pyridine and fluorophenyl substituents. To assist inhibitor design, we required a docking model that was able to predict the bound conformations of compounds belonging to the chemical class of interest. However, the large number of crystal structures published in the PDB and the conformational variability of p38α MAPK made choosing a crystal structure for docking difficult. An overlay of p38α MAPK crystal structures, shown in Fig. 4, illustrates the significant ligand induced conformational variation observed in the p38α MAPK active site. Aside from the movement of the DFG motif in the activation loop, major conformational changes occur in the glycine rich loop and the Met109 and Gly110 residues in the hinge region, where a flip in the peptide bond can occur (Fig. 5) [38, 39].

To assess which crystal structure to use as a model, we have performed a virtual screening study using a library of actives and decoys. The study was performed on selected

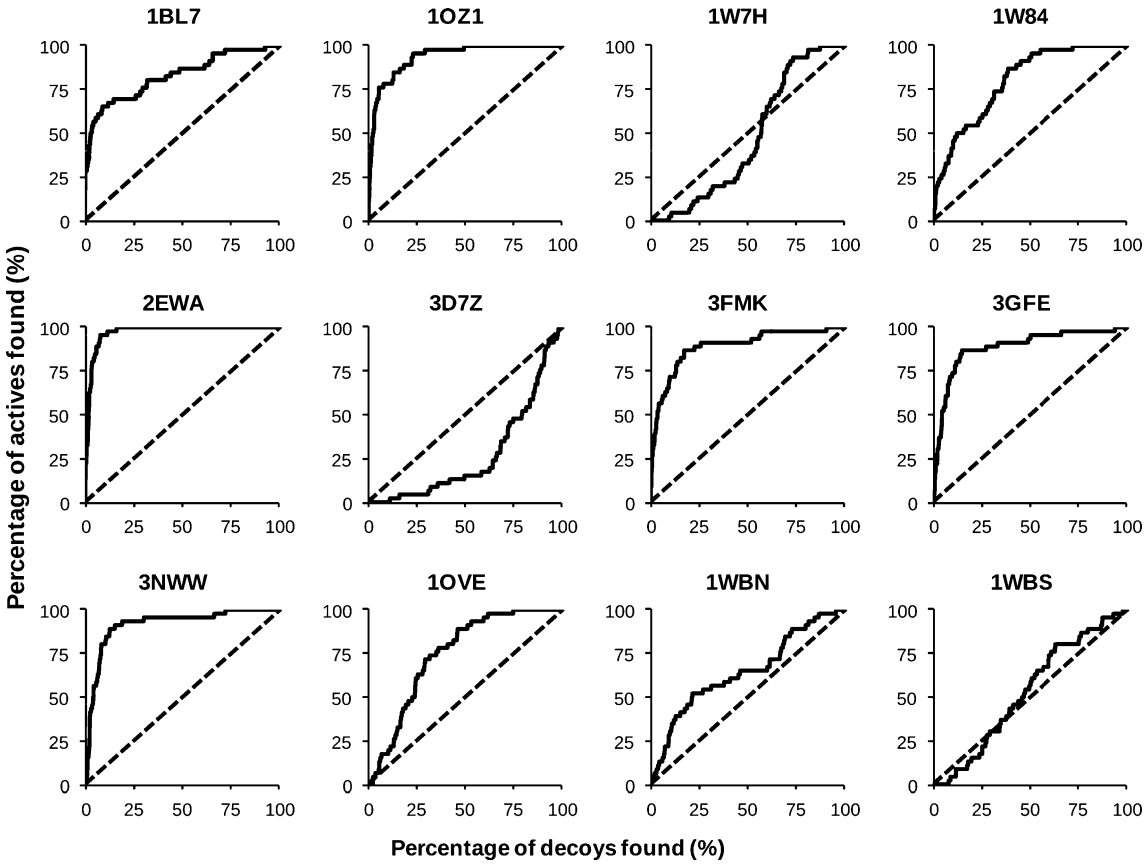

Fig. 6 ROC curves obtained from the virtual screen; DFG-in (1BL7, 1OZ1, 1W7H, 1W84, 2EWA, 3FMK, 3D7Z, 3GFE, 3NWW), DFG-in with Met109-Gly110 peptide flip (1OVE), DFG-out (1WBN) and DFG-out with Met109-Gly110 peptide flip (1WBS)

crystal structures with different binding site conformations in order to identify the conformation that can best dock ligands of this class and filter down the number of crystal structures to analyse. Subsequently, ensembles of the best performing receptor conformation were used to improve enrichment and find the most suitable model for p38α MAPK inhibitors containing heterocycles substituted with a pyridine and fluorophenyl ring.

Virtual screening

The virtual screens were conducted using a library of 1,000 drug-like decoy compounds enriched with 46 known p38α MAPK inhibitors having structures that contain or resemble heterocycles with pyridine and fluorophenyl substituents (Fig. 3). Table 2 compiles the results of our virtual screening study (see above), totaling forty-four p38α MAPK structures. Table 2 includes an evaluation of each docking model based on the enrichment factor obtained at 2 % of the virtual screen. Models are denoted as 'poor' when the enrichment was equal to or less than 10.0 % of $EFmax^{2\%}$, 'good' when the enrichment was between 10.0 and 40.0 % of $EFmax^{2\%}$ and 'very good' when the enrichment was equal to or more than 40.0 % of $EFmax^{2\%}$. Despite conformational variation in the active site, the p38α MAPK structures are all very similar, having protein heavy atom RMS deviations of less than 1.00 Å. All cognate ligands docked in the crystal structure orientation and had an acceptable RMSD of less than 2.0 Å.

Our virtual screen was conducted on p38α MAPK crystal structures representing four major conformations; the DFG-in conformation (set 1), DFG-in with a Met109-Gly110 flip in the hinge region (set 2), DFG-out conformation (set 3) and DFG-out with a Met109-Gly110 flip in the hinge region (set 4) (Fig. 5). Within the binding site, residues in the glycine rich loop are seen to adopt different conformations; this is particularly evident for Tyr35, which is often not resolved. The different configurations of Tyr35 were not considered for the selection of crystal structures but are discussed further below. Figure 6 shows ROC curves for the virtual screens performed on four p38α MAPK binding site conformations. The structure 1BL7 identified the most active compounds with a ROC AUC of 0.83 and an enrichment factor of 14.1 at 2 % of the database screen (Table 2). This structure belonged to set 1 having the DFG-in conformation. 1WBN, having the DFG-out arrangement (set 3), only identified a small number of actives with an enrichment factor of 2.2 at 2 % database screen. Crystal structures 1OVE and 1WBS, which have a flip in the Met109-Gly110 peptide bond of the hinge region, gave little enrichment. It is evident that p38α MAPK inhibitors containing pyridine and fluorophenyl substituents bind to the active form of p38α MAPK in the DFG-in conformation. This is in agreement with previous reports that the backbone amide NH of Met109 makes a key hydrogen bond with the pyridinyl ring nitrogen of some inhibitors including SB203580 [36]. Therefore a peptide flip in Met109-Gly110 residues of the hinge region are unfavorable for heterocycles bearing pyridine and fluorophenyl substituents. Additionally, movement to the DFG-out conformation results in a smaller ATP binding pocket as a result of the shift in the Phe169 residue, so it was not surprising that the actives did not bind to either of the DFG-out conformations tested.

Virtual screening using eight representative crystal structures from each of sets 2, 3 and 4 collectively gave poor enrichment of actives (Table 2) and supported the finding that set 1 consisted of structures with the relevant binding site conformation. Structures belonging to set 1 therefore became the focus of the study. The structures in set 1 were chosen to include those containing ligands bearing pyridine

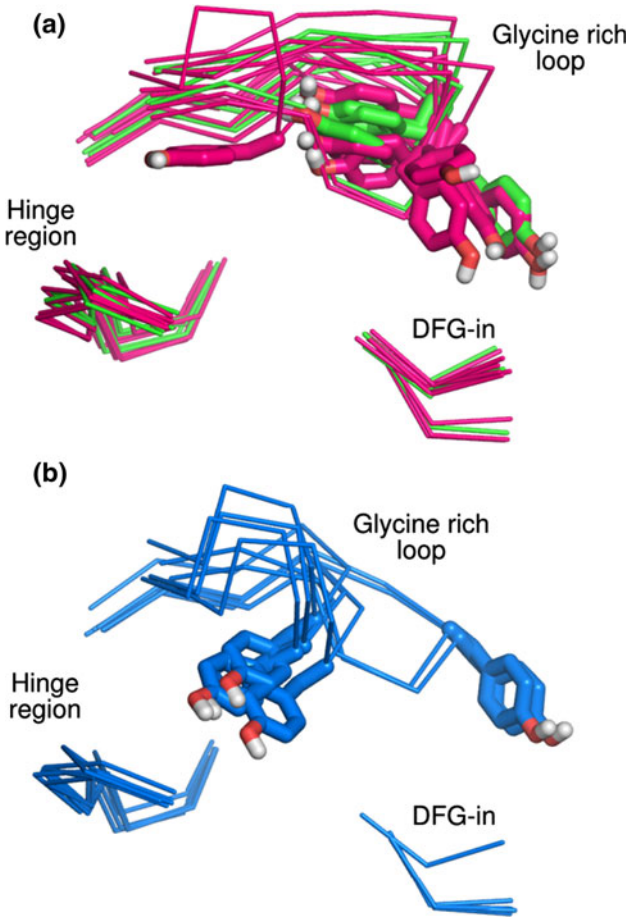

Fig. 7 The active site of p38α MAPK crystal structures from set 1. Tyr35 of the glycine rich loop is *highlighted* in stick representation. **a** Structures which performed well (*pink*: 1A9U, 1BL6, 1DI9, 1OUK, 1W84, 1YQJ, 2BAL, 3GFE and 3NWW) or very well (*green*: 1BL7, 1OZ1, 2EWA, 3FMK) in virtual screening; **b** Structures which performed poorly in virtual screening (*blue*: 1BMK, 1W7H, 3D7Z, 3FL4, 3FMM, 3IPH and 3MPT)

and fluorophenyl substituents but also included other high resolution complexes with structurally different ligands covering most of the inhibitor classes. Figure 6 shows structure 2EWA clearly gave the best discrimination having a ROC AUC of 0.97 and an enrichment factor of 13.0 at 2 % of the virtual screen. Other structures that performed well include 1OZ1 and 3FMK with enrichment factors of 10.8 and 11.9 respectively at 2 % of the virtual screen. 1W84, 3GFE and 3NWW all gave good performances with ROC AUCs of 0.80, 0.88 and 0.91. Some of the crystal structures produced no enrichment of actives at 2 % of the screen even though they adopted the DFG-in conformation. Crystal structures 1W7H and 3D7Z performed worse than random with ROC AUCs of 0.47 and 0.27 respectively. Interestingly, overlays of these crystal structures onto 1A9U all gave protein heavy atom RMS deviations between 0.34 and 0.40 Å, yet their performance in the virtual screen was greatly variable.

The large variation in the performances of the DFG-in crystal structures from set 1 is attributed to differences in the conformation of the glycine rich loop. Figure 7 shows an overlay of these structures depicting the additional

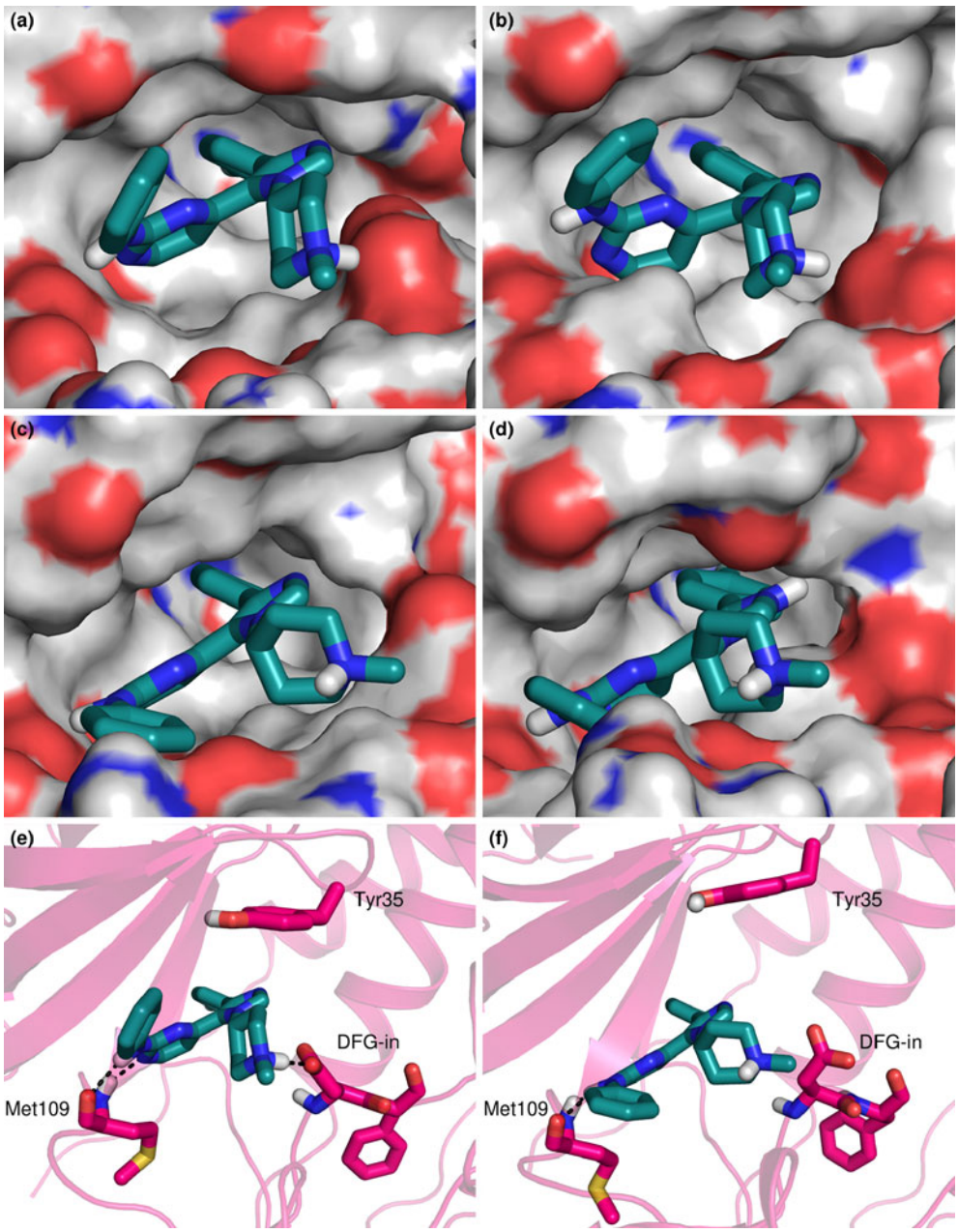

Fig. 8 Docked pose of active 1–20 into; **a** 1BL7 **b** 1OZ1 **c** 2EWA and **d** 3FMK. Hydrogen bonding interactions of active 1–20 with active site residues of the top performing structures; **e** 1BL7 and **f** 2EWA

conformational variation in the glycine rich loop with Tyr35 residue side chain highlighted in stick representation. This Tyr35 side chain is oriented toward the ligand in the binding pocket, suggesting a role as a possible hydrogen bond donor or providing stacking interactions with the actives. Figure 7a shows crystal structures 1A9U, 1BL6, 1BL7, 1DI9, 1OUK, 1OZ1, 1W84, 1YQJ, 2BAL, 2EWA, 3FMK, 3GFE and 3NWW which gave good to very good enrichments, and Fig. 7b overlays structures 1BMK, 1W7H, 3D7Z, 3FL4, 3FMM, 3IPH and 3MPT which gave poor or no enrichment. Of the four top performing structures, three have similar binding site conformations (1BL7, 1OZ1 and 2EWA). These structures all contain an inhibitor of the compound class of interest. Comparison of 1BL7, 1OZ1 and 2EWA shows a 1–2 Å translation of residues in the glycine rich loop with largest differences observed between residues 33–36. 3FMK on the other hand, contains a pyridopyrimidinone ligand and differs greatly from the other three structures in the glycine rich loop. Comparison of the four structures also reveals a small translation between residues of the hinge region and a significant difference in the conformation of the Met109 side chain. In the activation loop, the Asp168 side chain of the DFG motif is also in a variety of different orientations. It is also important to note that residues Gly170-Ala184 were not resolved for 2EWA; however, these residues are outside the binding site.

Of the twenty DFG-in structures, six were crystallized with a ligand containing a heterocycle substituted with pyridine and fluorophenyl groups and the other fourteen contained diverse ligands. Since the active set in the database screen were specifically selected to represent inhibitors containing a heterocycle with pyridine and fluorophenyl substituents, it might have been expected that these six crystal structures already containing an inhibitor of this class would perform the best. This is true in most, but not all cases. Of the six crystal structures containing this substituted heterocycle, five gave good or very good results. Of the top performing crystal structures (1BL7, 1OZ1, 2EWA and 3FMK), three were bound to a compound of the inhibitor class of interest. However, structures 1OUK, 1W84, 1YQJ, 2BAL, 3GFE and 3NWW, containing structurally diverse ligands, were able to perform better than some of the structures complexed with a ligand of the inhibitor class. The observation that the six crystal structures do not outrank crystal structures with structurally diverse ligands, highlights the difficulty in finding the right model for docking into p38α MAPK. Even structures with an RMSD of greater than 2.0 Å from 1BL7 in the loop region (residues 30–40) can perform well in the virtual screen. The conformational variability of the receptor is evident in Fig. 7 where the glycine rich loop of the top performing structures adopts a range of conformations.

Since it is well established through crystallographic and mutagenesis studies that the pyridinyl part of the ligand makes a hydrogen bonding interaction with Met109 and the fluorophenyl portion of the ligand interacts with the hydrophobic region I [7, 8, 22, 36], cases where the actives did not dock in this orientation were considered to be incorrect. From inspection of the ligands docked into the best performing structures, all ligands docked in the correct conformation using structure 1OZ1. Forty-four of the forty-six actives docked into 2EWA were in the expected binding orientation, whereas ten of the forty-six actives docked into 1BL7 bound in the wrong orientation. Of these ten, six contained the pyridine or pyridine-like substituent binding in hydrophobic region I. For 3FMK five of the forty-six actives did not dock in the expected orientation.

Figure 8a–d shows active compound 1–20 docked into structures 1BL7, 1OZ1, 2EWA and 3FMK respectively. The docked poses of this compound are similar to each other and are consistent with literature in that the pyrimidine substituent and fluorophenyl ring bind in the correct orientation. The 2-anilinopyrimidine group of active 1–20 extends into hydrophobic region II but also gains polar contacts through a hydrogen bonding interaction with the backbone carbonyl of Met109. The conformation of the aniline ring differs between docked structures. Additional interactions may be made with the *N*-methylpiperidine group and Asp168 of the DFG motif, observed when docked to 1BL7 and 1OZ1. Figure 8e, f shows the hydrogen bonding interactions made within the binding site of the two best performing models. Compared to 2EWA, the 1BL7 structure makes an additional hydrogen bonding

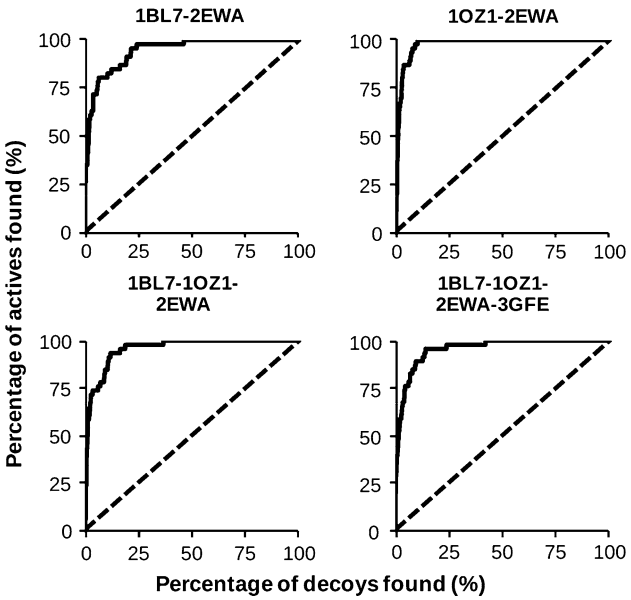

Fig. 9 ROC plots for high-scoring ensembles of two to four receptors

Table 3 ROC AUCs and enrichment factors obtained for ensembles of p38α MAPK structures and pairwise RMSD calculations for the active ligands across the receptors that make up the ensemble

Ensemble	ROC AUC	$EF^{2\%}$ [a]	$EF^{5\%}$ [b]	$EF^{10\%}$ [c]	Mean RMSD (Å)	Pairs with RMSD < 3.0Å (%)[d]
1A9U-1BL7	0.83	13.0	7.3	5.2	3.24	60.9
1A9U-1OZ1	0.91	10.8	7.7	6.1	2.71	73.9
1A9U-2EWA	0.94	9.7	9.0	7.4	3.46	65.2
1A9U-3FMK	0.87	9.7	6.9	5.4	4.32	43.5
1A9U-3GFE	0.87	9.7	7.3	5.4	4.34	45.7
1BL7-1OZ1	0.91	15.2	10.7	7.1	2.74	71.7
1BL7-2EWA	0.94	**17.3**	12.0	8.0	3.27	65.2
1BL7-3FMK	0.89	15.2	10.3	6.9	3.65	65.2
1BL7-3GFE	0.89	15.2	7.7	6.3	3.79	54.4
1OZ1-2EWA	**0.98**	14.1	**12.9**	**8.7**	2.78	71.7
1OZ1-3FMK	0.94	15.2	11.6	7.4	3.37	60.9
1OZ1-3GFE	0.94	11.9	10.3	7.4	3.52	54.4
2EWA-3FMK	0.96	14.1	12.0	8.4	3.01	67.4
2EWA-3GFE	0.96	14.1	11.6	8.2	2.75	71.7
3FMK-3GFE	0.91	10.8	10.3	7.6	**2.51**	**80.4**
1A9U-1BL7-1OZ1	0.89	11.9	9.4	6.5	2.90	54.4
1A9U-1BL7-2EWA	0.92	11.9	9.9	7.6	3.33	47.8
1A9U-1BL7-3FMK	0.88	13.0	9.0	6.1	3.74	39.1
1A9U-1BL7-3GFE	0.87	14.1	7.7	6.1	3.79	37.0
1A9U-1OZ1-2EWA	0.96	10.8	11.6	7.8	2.98	56.5
1A9U-1OZ1-3FMK	0.93	10.8	9.4	6.9	3.47	39.1
1A9U-1OZ1-3GFE	0.92	11.9	8.6	7.1	3.53	39.1
1A9U-2EWA-3FMK	0.94	11.9	10.7	7.6	3.60	32.6
1A9U-2EWA-3GFE	0.94	13.0	10.3	7.1	3.52	37.0
1A9U-3FMK-3GFE	0.89	9.7	8.2	6.9	3.73	30.4
1BL7-1OZ1-2EWA	0.96	**16.2**	**12.9**	7.6	2.93	54.4
1BL7-1OZ1-3FMK	0.93	**16.2**	12.0	7.4	3.26	47.8
1BL7-1OZ1-3GFE	0.93	**16.2**	9.9	7.1	3.35	39.1
1BL7-2EWA-3FMK	0.94	**16.2**	12.0	7.8	3.31	47.8
1BL7-2EWA-3GFE	0.94	**16.2**	11.6	7.6	3.27	43.5
1BL7-3FMK-3GFE	0.91	15.2	10.3	6.9	3.32	47.8
1OZ1-2EWA-3FMK	**0.97**	14.1	**12.9**	**8.7**	3.05	41.3
1OZ1-2EWA-3GFE	**0.97**	13.0	12.4	8.4	3.02	45.7
1OZ1-3FMK-3GFE	0.94	10.8	11.2	7.4	3.14	43.5
2EWA-3FMK-3GFE	0.95	11.9	11.6	8.2	**2.76**	**60.9**
1A9U-1BL7-1OZ1-2EWA	0.94	13.0	11.2	7.4	3.03	**43.5**
1A9U-1BL7-1OZ1-3FMK	0.91	14.1	10.3	7.4	3.34	34.8
1A9U-1BL7-1OZ1-3GFE	0.91	14.1	9.4	7.1	3.39	32.6
1A9U-1BL7-2EWA-3FMK	0.93	14.1	10.3	7.6	3.49	26.1
1A9U-1BL7-2EWA-3GFE	0.93	14.1	9.4	7.6	3.48	28.3
1A9U-1BL7-3FMK-3GFE	0.89	14.1	8.6	6.9	3.64	28.3
1A9U-1OZ1-2EWA-3FMK	**0.96**	13.0	12.0	7.8	3.28	28.3
1A9U-1OZ1-2EWA-3GFE	**0.96**	11.9	11.6	7.6	3.26	30.4
1A9U-1OZ1-3FMK-3GFE	0.93	13.0	9.9	6.9	3.46	26.1
1A9U-2EWA-3FMK-3GFE	0.94	10.8	11.2	7.6	3.40	21.7
1BL7-1OZ1-2EWA-3FMK	**0.96**	15.2	**12.9**	7.6	3.14	34.8
1BL7-1OZ1-2EWA-3GFE	**0.96**	**16.2**	11.5	7.8	3.14	32.6

Table 3 continued

Ensemble	ROC AUC	$EF^{2\%}$ [a]	$EF^{5\%}$ [b]	$EF^{10\%}$ [c]	Mean RMSD (Å)	Pairs with RMSD < 3.0Å (%)[d]
1BL7-1OZ1-3FMK-3GFE	0.93	15.2	11.2	7.6	3.27	32.6
1BL7-2EWA-3FMK-3GFE	0.94	**16.2**	11.6	8.0	3.16	39.1
1OZ1-2EWA-3FMK-3GFE	**0.96**	13.0	11.6	**8.7**	**2.99**	32.6

The best results for each set of ensembles are highlighted in bold

[a] Maximum enrichment factor at 2 % = 22.7. [b] Maximum enrichment factor at 5 % = 19.7. [c] Maximum enrichment factor at 10 % = 10.0. [d] Numbers of comparisons made for ensembles of two, three and four receptors were 46, 138 and 276 respectively

interaction, with the protonated piperidine group hydrogen bonded to Asp168.

Ensemble docking

Having identified four crystal structures that give 'very good' enrichment of active compounds, we investigated whether the discrimination of active compounds could be further improved using ensemble docking. Ensemble docking refers to docking compounds into multiple receptor conformations [79, 80]. Such methods allow incorporation of receptor flexibility and may result in better pose prediction of bound ligands and improve virtual screening enrichment of actives [80]. Several groups have performed ensemble docking studies, combining receptors with similar conformations as well as receptors with diverse conformations. Usually the active set in the virtual screening library comprises of diverse ligands [35, 80]. In our case, we are interested in a specific inhibitor class and therefore a set of related compounds was chosen as our active set.

Ensembles of two to four receptors were chosen by combining the individual models which gave ROC AUCs greater than 0.75 and $EF^{2\%}$ greater than 7.0. Figure 9 and Table 3 show the results of the ensemble docking study. Generally all the tested ensembles gave good or very good enrichments with ROC AUCs all being greater than 0.80. The two-member ensemble 1BL7-2EWA produced the best enrichment of actives with an enrichment factor of 17.3 at 2 % of the virtual screen, a significant improvement on their constituent receptors, and also maintained a high ROC AUC value of 0.94. Several three- and four-member ensembles gave an enrichment factor of 16.2 at 2 % of the database screen, which is better than results obtained for single receptor models. Although combinations of three and four receptors gave good results, they failed to improve on the best two-member ensemble. To assess whether the different models provided concordant predictions of the bound ligand poses, we measured pairwise RMSD between the ligands for the receptors used in the ensemble docking study (Table 3). For the pair 3FMK-3GFE it was found that 80.4 % of the active ligands (37 compounds) docked with an RMSD between poses of less than 3.0 Å, indicating that the large majority of the actives dock into these crystal structures in similar orientations. For most ensembles of two receptors more than 50 % of the poses were similar. In our best model, the 1BL7-2EWA ensemble, 65.2 % of the active ligands (30 compounds) docked in similar poses. The most concordant three-member ensemble, 2EWA-3FMK-3GFE, had a mean pairwise RMSD of 2.79 Å. The 1A9U-1BL7-1OZ1-2EWA ensemble docked 43.5 % of the active ligands (20 compounds) with an RMSD between poses of less than 3.0 Å.

Conclusion

Virtual screening studies are a useful way to develop predictive models for ligand binding to conformationally flexible proteins such as p38α MAPK. In this work we have used virtual screening to examine a large number of crystal structures having different binding site conformations and identified four crystal structures (1BL7, 1OZ1, 2EWA and 3FMK) that effectively distinguish active compounds belonging to a particular structural class, exemplified by SB203580, from drug-like decoy compounds. In these model structures, p38α MAPK is in its DFG-in conformation. However additional features are also important for recognition of this structural class. This includes the conformation of the glycine rich loop and activation loop. In particular the Tyr35 side chain, which is directed towards the ligand in the binding pocket, seems to be an important residue for interacting with inhibitors. This residue appears in a similar conformation in 1BL7, 1OZ1 and 2EWA but different in 3FMK. Importantly, we observe that not all structures which are in the DFG-in arrangement give predictive docking results. Studies on structures with the DFG-in conformation indicate that those containing ligands in the same compound class as SB203580 do not uniformly perform the best. Some crystal structures containing structurally different ligands outperformed those complexed to ligands of the SB203580 class. Interestingly, not all high resolution structures performed well. These results highlight the importance of careful selection of

docking models for structure based drug design projects. Validation of the docking method, where the chosen crystal structure contains a ligand in the same compound class to the molecules that are being docked, is required in order to gain accurate results.

The predictive ability of virtual screening can be improved by considering an ensemble of model structures. In this work we find that combining the results of the most predictive structures significantly improves the results. Combinations of two-, three- and four-receptor ensembles were able to improve enrichment relative to the single receptor. The best two-member ensemble 1BL7-2EWA gave an enrichment factor of 17.3 at 2 % of the database screen. Several three- and four-member receptor combinations also performed well but did not improve on the 1BL7-2EWA ensemble. The 1BL7-2EWA structures agree well on the bound ligand poses; an important aid to drug design.

We are now using the 1BL7-2EWA ensemble model in further docking studies to identify protein–ligand interactions that can be exploited in the design of new p38α MAPK inhibitors related to SB203580. Using this crystal structure ensemble we can better predict the bound conformation of the designed ligands allowing us to capture and improve existing protein–ligand interactions. Identification of ligands that dock well into the protein structure will be followed up with synthesis and subsequently evaluated using biological assays.

Acknowledgments N. B. Vinh is supported by an Australian Postgraduate Award (APA) scholarship. The authors acknowledge computational resources and technical support provided by Victorian Partnership for Advanced Computing (VPAC).

References

1. Schieven GL (2005) Curr Top Med Chem 5(10):921
2. Keshet Y, Seger R (2010) Methods Mol Biol 661:3
3. Yong HY, Koh MS, Moon A (2009) Expert Opin Investig Drugs 18(12):1893
4. Cuenda A, Rousseau S (2007) Biochim Biophys Acta 1773(8):1358
5. Pettus LH, Wurz RP (2008) Curr Top Med Chem 8(16):1452
6. Wu B, Wang HL, Pettus L, Wurz RP, Doherty EM, Henkle B, McBride HJ, Saris CJM, Wong LM, Plant MH, Sherman L, Lee MR, Hsieh F, Tasker AS (2010) J Med Chem 53(17):6398
7. Dominguez C, Tamayo N, Zhang D (2005) Expert Opin Ther Pat 15(7):801
8. Jackson PF, Bullington JL (2002) Curr Top Med Chem 2(9):1011
9. Trejo A, Arzeno H, Browner M, Chanda S, Cheng S, Comer DD, Dalrymple SA, Dunten P, Lafargue J, Lovejoy B, Freire-Moar J, Lim J, McIntosh J, Miller J, Papp E, Reuter D, Roberts R, Sanpablo F, Saunders J, Song K, Villasenor A, Warren SD, Welch M, Weller P, Whiteley PE, Zeng L, Goldstein DM (2003) J Med Chem 46(22):4702
10. Goldstein DM, Gabriel T (2005) Curr Top Med Chem 5(10):1017
11. Hynes J Jr, Leftheris K (2005) Curr Top Med Chem 5(10):967
12. Lee MR, Dominguez C (2005) Curr Med Chem 12(25):2979
13. Tamayo N, Liao L, Goldberg M, Powers D, Tudor YY, Yu V, Wong LM, Henkle B, Middleton S, Syed R, Harvey T, Jang G, Hungate R, Dominguez C (2005) Bioorg Med Chem Lett 15(9):2409
14. Peifer C, Wagner G, Laufer S (2006) Curr Top Med Chem 6(2):113
15. Laufer SA, Hauser DRJ, Domeyer DM, Kinkel K, Liedtke AJ (2008) J Med Chem 51(14):4122
16. Aston NM, Bamborough P, Buckton JB, Edwards CD, Holmes DS, Jones KL, Patel VK, Smee PA, Somers DO, Vitulli G, Walker AL (2009) J Med Chem 52(20):6257
17. Selness SR, Devraj RV, Monahan JB, Boehm TL, Walker JK, Devadas B, Durley RC, Kurumbail R, Shieh H, Xing L, Hepperle M, Rucker PV, Jerome KD, Benson AG, Marrufo LD, Madsen HM, Hitchcock J, Owen TJ, Christie L, Promo MA, Hickory BS, Alvira E, Naing W, Blevis-Bal R (2009) Bioorg Med Chem Lett 19(20):5851
18. Goldstein DM, Kuglstatter A, Lou Y, Soth MJ (2010) J Med Chem 53(6):2345
19. Mavunkel BJ, Perumattam JJ, Tan X, Luedtke GR, Lu Q, Lim D, Kizer D, Dugar S, Chakravarty S, Xu YJ, Jung J, Liclican A, Levy DE, Tabora J (2010) Bioorg Med Chem Lett 20(3):1059
20. Wang Z, Harkins PC, Ulevitch RJ, Han J, Cobb MH, Goldsmith EJ (1997) Proc Natl Acad Sci USA 94(6):2327
21. Traxler P (1998) Expert Opin Ther Pat 8(12):1599
22. Tong L, Pav S, White DM, Rogers S, Crane KM, Cywin CL, Brown ML, Pargellis CA (1997) Nat Struct Biol 4(4):311
23. Angiolini M (2011) Future Med Chem 3(3):309
24. Filomia F, De Rienzo F, Menziani MC (2010) Bioorg Med Chem 18(18):6805
25. Pargellis C, Tong L, Churchill L, Cirillo PF, Gilmore T, Graham AG, Grob PM, Hickey ER, Moss N, Pav S, Regan J (2002) Nat Struct Biol 9(4):268
26. Gill AL, Frederickson M, Cleasby A, Woodhead SJ, Carr MG, Woodhead AJ, Walker MT, Congreve MS, Devine LA, Tisi D, O'Reilly M, Seavers LCA, Davis DJ, Curry J, Anthony R, Padova A, Murray CW, Carr RAE, Jhoti H (2005) J Med Chem 48(2):414
27. Karcher SC, Laufer SA (2009) Curr Top Med Chem 9(7):655
28. Cuenda A, Rouse J, Doza YN, Meier R, Cohen P, Gallagher TF, Young PR, Lee JC (1995) FEBS Lett 364(2):229
29. Hubbard RE (ed) (2006) Structure based drug discovery: an overview. The Royal Society of Chemistry, Cambridge
30. Wlodawer A, Minor W, Dauter Z, Jaskolski M (2008) FEBS J 275(1):1
31. Acharya KR, Lloyd MD (2005) Trends Pharmacol Sci 26(1):10
32. Andricopulo AD, Salum LB, Abraham DJ (2009) Curr Top Med Chem 9(9):771
33. Cheeseright TJ, Holm M, Lehmann F, Luik S, Gottert M, Melville JL, Laufer S (2009) J Med Chem 52(14):4200
34. Armen RS, Chen J, Brooks CL (2009) J Chem Theory Comput 5(10):2909
35. Rao S, Sanschagrin PC, Greenwood JR, Repasky MP, Sherman W, Farid R (2008) J Comput-Aided Mol Des 22(9):621
36. Wang Z, Canagarajah BJ, Boehm JC, Kassisa S, Cobb MH, Young PR, Abdel-Meguid S, Adams JL, Goldsmith EJ (1998) Structure 6(9):1117
37. Shewchuk L, Hassell A, Wisely B, Rocque W, Holmes W, Veal J, Kuyper LF (2000) J Med Chem 43(1):133
38. Stelmach JE, Liu L, Patel SB, Pivnichny JV, Scapin G, Singh S, Hop CECA, Wang Z, Strauss JR, Cameron PM, Nichols EA, O'Keefe SJ, O'Neill EA, Schmatz DM, Schwartz CD, Thompson CM, Zaller DM, Doherty JB (2003) Bioorg Med Chem Lett 13(2):277

39. Fitzgerald CE, Patel SB, Becker JW, Cameron PM, Zaller D, Pikounis VB, O'Keefe SJ, Scapin G (2003) Nat Struct Biol 10(9):764
40. Hartshorn MJ, Murray CW, Cleasby A, Frederickson M, Tickle IJ, Jhoti H (2005) J Med Chem 48(2):403
41. Michelotti EL, Moffett KK, Nguyen D, Kelly MJ, Shetty R, Chai X, Northrop K, Namboodiri V, Campbell B, Flynn GA, Fujimoto T, Hollinger FP, Bukhtiyarova M, Springman EB, Karpusas M (2005) Bioorg Med Chem Lett 15(23):5274
42. McClure KF, Abramov YA, Laird ER, Barberia JT, Cai W, Carty TJ, Cortina SR, Danley DE, Dipesa AJ, Donahue KM, Dombroski MA, Elliott NC, Gabel CA, Han S, Hynes TR, LeMotte PK, Mansour MN, Marr ES, Letavic MA, Pandit J, Ripin DB, Sweeney FJ, Tan D, Tao Y (2005) J Med Chem 48(18):5728
43. Sullivan JE, Holdgate GA, Campbell D, Timms D, Gerhardt S, Breed J, Breeze AL, Bermingham A, Pauptit RA, Norman RA, Embrey KJ, Read J, VanScyoc WS, Ward WHJ (2005) Biochemistry 44(50):16475
44. Vogtherr M, Saxena K, Hoelder S, Grimme S, Betz M, Schieborr U, Pescatore B, Robin M, Delarbre L, Langer T, Wendt KU, Schwalbe H (2006) Angew Chem Int Ed Engl 45(6):993
45. Angell R, Aston NM, Bamborough P, Buckton JB, Cockerill S, deBoeck SJ, Edwards CD, Holmes DS, Jones KL, Laine DI, Patel S, Smee PA, Smith KJ, Somers DO, Walker AL (2008) Bioorg Med Chem Lett 18(15):4428
46. Cogan DA, Aungst R, Breinlinger EC, Fadra T, Goldberg DR, Hao MH, Kroe R, Moss N, Pargellis C, Qian KC, Swinamer AD (2008) Bioorg Med Chem Lett 18(11):3251
47. Baldwin I, Bamborough P, Haslam CG, Hunjan SS, Longstaff T, Mooney CJ, Patel S, Quinn J, Somers DO (2008) Bioorg Med Chem Lett 18(19):5285
48. Wurz RP, Pettus LH, Xu S, Henkle B, Sherman L, Plant M, Miner K, McBride H, Wong LM, Saris CJM, Lee MR, Chmait S, Mohr C, Hsieh F, Tasker AS (2009) Bioorg Med Chem Lett 19(16):4724
49. Simard JR, Getlik M, Grutter C, Schneider R, Wulfert S, Rauh D (2010) J Am Chem Soc 132(12):4152
50. Jerome KD, Rucker PV, Xing L, Shieh HS, Baldus JE, Selness SR, Letavic MA, Braganza JF, McClure KF (2010) Bioorg Med Chem Lett 20(2):469
51. Goldstein DM, Soth M, Gabriel T, Dewdney N, Kuglstatter A, Arzeno H, Chen J, Bingenheimer W, Dalrymple SA, Dunn J, Farrell R, Frauchiger S, La Fargue J, Ghate M, Graves B, Hill RJ, Li F, Litman R, Loe B, McIntosh J, McWeeney D, Papp E, Park J, Reese HF, Roberts RT, Rotstein D, San Pablo B, Sarma K, Stahl M, Sung ML, Suttman RT, Sjogren EB, Tan Y, Trejo A, Welch M, Weller P, Wong BR, Zecic H (2011) J Med Chem 54(7):2255
52. Soth M, Abbot S, Abubakari A, Arora N, Arzeno H, Billedeau R, Dewdney N, Durkin K, Frauchiger S, Ghate M, Goldstein DM, Hill RJ, Kuglstatter A, Li F, Loe B, McCaleb K, McIntosh J, Papp E, Park J, Stahl M, Sung ML, Suttman R, Swinney DC, Weller P, Wong B, Zecic H, Gabriel T (2011) Bioorg Med Chem Lett 21(11):3452
53. Down K, Bamborough P, Alder C, Campbell A, Christopher JA, Gerelle M, Ludbrook S, Mallett D, Mellor G, Miller DD, Pearson R, Ray K, Solanke Y, Somers D (2010) Bioorg Med Chem Lett 20(13):3936
54. Ahn YM, Clare M, Ensinger CL, Hood MM, Lord JW, Lu WP, Miller DF, Patt WC, Smith BD, Vogeti L, Kaufman MD, Petillo PA, Wise SC, Abendroth J, Chun L, Clark R, Feese M, Kim H, Stewart L, Flynn DL (2010) Bioorg Med Chem Lett 20(19):5793
55. Lin S, Wrobleski ST, Hynes J Jr, Pitt S, Zhang R, Fan Y, Doweyko AM, Kish KF, Sack JS, Malley MF, Kiefer SE, Newitt JA, McKinnon M, Trzaskos J, Barrish JC, Dodd JH, Schieven GL, Leftheris K (2010) Bioorg Med Chem Lett 20(19):5864
56. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) Nucleic Acids Res 28(1):235
57. Wagner G, Laufer S (2006) Med Res Rev 26(1):1
58. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) J Med Chem 47(7):1739
59. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2004) J Med Chem 47(7):1750
60. McRobb FM, Capuano B, Crosby IT, Chalmers DK, Yuriev E (2010) J Chem Inf Model 50(4):626
61. Daylight Chemical Information Systems, Inc. http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html Accessed 21 Nov 2011
62. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT (2006) J Med Chem 49(21):6177
63. Triballeau N, Acher F, Brabet I, Pin J-P, Bertrand H-O (2005) J Med Chem 48(7):2534
64. Pearlman DA, Charifson PS (2001) J Med Chem 44(4):502
65. Adams JL, Boehm JC, Gallagher TF, Kassis S, Webb EF, Hall R, Sorenson M, Garigipati R, Griswold DE, Lee JC (2001) Bioorg Med Chem Lett 11(21):2867
66. Boehm JC, Bower MJ, Gallagher TF, Kassis S, Johnson SR, Adams JL (2001) Bioorg Med Chem Lett 11(9):1123
67. Boehm JC, Smietana JM, Sorenson ME, Garigipati RS, Gallagher TF, Sheldrake PL, Bradbeer J, Badger AM, Laydon JT, Lee JC, Hillegass LM, Griswold DE, Breton JJ, Chabot-Fletcher MC, Adams JL (1996) J Med Chem 39(20):3929
68. Gallagher TF, Seibel GL, Kassis S, Laydon JT, Blumenthal MJ, Lee JC, Lee D, Boehm JC, Fier-Thompson SM, Abt JW, Soreson ME, Smietana JM, Hall RF, Garigipati RS, Bender PE, Erhard KF, Krog AJ, Hofmann GA, Sheldrake PL, McDonnell PC, Kumar S, Young PR, Adams JL (1997) Bioorg Med Chem 5(1):49
69. Gallagher TF, Fier-Thompson SM, Garigipati RS, Sorenson ME, Smietana JM, Lee D, Bender PE, Lee JC, Laydon JT, Griswold DE, Chabot-Fletcher MC, Breton JJ, Adams JL (1995) Bioorg Med Chem Lett 5(11):1171
70. Thurmond RL, Wadsworth SA, Schafer PH, Zivin RA, Siekierka JJ (2001) Eur J Biochem 268(22):5747
71. Laufer SA, Striegel HG, Wagner GK (2002) J Med Chem 45(21):4695
72. de Laszlo SE, Visco D, Agarwal L, Chang L, Chin J, Croft G, Forsyth A, Fletcher D, Frantz B, Hacker C, Hanlon W, Harper C, Kostura M, Li B, Luell S, MacCoss M, Mantlo N, O'Neill EA, Orevillo C, Pang M, Parsons J, Rolando A, Sahly Y, Sidler K, Widmer WR, O'Keefe SJ (1998) Bioorg Med Chem Lett 8(19):2689
73. Revesz L, Blum E, Di Padova FE, Buhl T, Feifel R, Gram H, Hiestand P, Manning U, Rucklin G (2004) Bioorg Med Chem Lett 14(13):3595
74. Revesz L, Di Padova FE, Buhl T, Feifel R, Gram H, Hiestand P, Manning U, Wolf R, Zimmerlin AG (2002) Bioorg Med Chem Lett 12(16):2109
75. Revesz L, Di Padova FE, Buhl T, Feifel R, Gram H, Hiestand P, Manning U, Zimmerlin AG (2000) Bioorg Med Chem Lett 10(11):1261
76. Fujita M, Hirayama T, Ikeda N (2002) Bioorg Med Chem 10(10):3113
77. Laufer SA, Wagner GK (2002) J Med Chem 45(13):2733
78. Rupert KC, Henry JR, Dodd JH, Wadsworth SA, Cavender DE, Olini GC, Fahmy B, Siekierka JJ (2003) Bioorg Med Chem Lett 13(3):347
79. Leach AR, Shoichet BK, Peishoff CE (2006) J Med Chem 49(20):5851
80. Craig IR, Essex JW, Spiegel K (2010) J Chem Inf Model 50(4):511