In silico insights into prediction and analysis of potential novel pyrrolopyridine analogs against human MAPKAPK-2: a new SAR-based hierarchical clustering approach

Kranthi Kumar Konidala¹ · Uma Devi Bommu² · Suneetha Yeguvapalli² · Neeraja Pabbaraju¹

Received: 28 June 2017 / Accepted: 13 August 2018 / Published online: 23 August 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

In the present study, we have focused on to elucidate potential bioactive pyrrolopyridine (PYP23) analogs against human mitogen-activated protein kinase-activated protein kinase-2 (MK-2). Here, in silico methods and computational systems biology tools were used as rational strategies to predict novel PYP23 analogs against the MK-2. Initially, crystal structure (PDB-ID: 2P3G) consists stereochemical conflicts were rectified by structure-optimization approaches using the Modeller program, and a new optimized-high resolution model was generated. The stereochemical qualities of the predicted MK-2 model were judged; these showed that the model was reliable for docking assessments. SAR-based bioactivity analysis showed that among the 197 datasets only 15 candidates contained bioactivity data and were accepted as probable MK-2 inhibitors. Virtual screening and docking strategies of dataset compounds against the ligand-binding domain of MK-2 recognized 13 composites containing high binding affinity than known compounds. Furthermore, the comparative structure clustering, in silico toxicogenomics and QSAR-based anticancer properties prediction approaches were successful in the recognition of five best potential compounds such as 60118340, 60118338, 60117736, 60118473 and 60118322, which have great anticancer and drug-likeness with non-toxicity class indices. Leu⁷⁰, Glu¹³⁹, Leu¹⁴¹, Glu¹⁴⁵, Glu¹⁹⁰, Thr²⁰⁶ and Asp²⁰⁷ were found to be novel hotspot residues prominently involved in H-bonds framing with ligands. Interestingly, they have shown better molecular similarity with known bioactive PYP inhibitors. Thus, predicted five compounds can useful as possible chemotherapeutic agents for MK-2 and show similar molecular actions like known PYP inhibitors. Overall, these streamlined new methods may have great potential to reveal possible ligands toward other molecular targets and biomarkers.

Keywords MAP kinase-activated protein kinase-2 · Pyrrolopyrimidine inhibitors · SAR-based bioassay clustering · Virtual screening · Ensemble docking · ADMET analysis

Kranthi Kumar Konidala and Uma Devi Bommu: equal contribution.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13205-018-1405-x>) contains supplementary material, which is available to authorized users.

✉ Neeraja Pabbaraju
pneeraja.bio2015@gmail.com

¹ Division of Molecular Physiology, Department of Zoology, Sri Venkateswara University, Tirupati 517502, India

² Division of Cancer Informatics, Department of Zoology, Sri Venkateswara University, Tirupati 517502, India

Introduction

Mitogen-activated protein kinase-activated protein kinase-2 (MAPKAP-K2 or MK-2) is a crucial enzyme in the p38 α /MK-2 signaling pathway. It promotes several advancements resembling physiological and pathophysiological responses that include inflammation, cell cycle regulation, and cell migration (Gurgis et al. 2013). Generally, MK-2 significantly supports post-transcriptional mechanism and standardizes cytokine assembly by directly performing as a p38 α substrate (Cargnello and Roux 2011). Recent examinations have revealed that MK-2 has progressed as a foremost mediator in various cellular responses such as cancer, inflammation associated with proinflammatory cytokines, and rheumatoid arthritis (Fiore et al. 2015; Chikanza 2004). Potent p38 α is an essential gene that regulates a myriad of cellular

associated mechanisms via interference with MK-2 (Anton et al. 2014). Enormous investigations propose that the selective MK-2 inhibitors show potential efficacy as p38 blockers without disrupting other cellular pathways and possess low rates of adverse side effects. This sort of observation has been found in wild-type and knockout mice models through *in vivo* and *in vitro* experiments (Anderson et al. 2007; Xu et al. 2006).

Several immune checkpoint routines have demonstrated that MK-2 is a principal protein, and its amelioration improved the chemotherapeutic drug's efficiency in tumors in p53 lacking conditions (Dobbelstein and Moll 2014; Chen et al. 2007; Manke et al. 2005). The anticancer agent doxorubicin acts as an apoptosis-antagonizing transcription factor, which leads to control p53-driven apoptosis in cancer cells by interacting with MK-2. Thus, the targeted MK-2 of the p38 α /MK-2 pathway has been documented as a key therapeutic hallmark for malignant disorder remedy (Höpker et al. 2012). Because of this, in this signaling cascade, p38 consists of more than 60 active substrates. Among these, MK-2 is a down-regulated substrate in fewer dynamic pathways. Therefore, it is demonstrated as a prospective beneficial target toward moderating several cancers with low side effects (Tremple et al. 2013; Mourey et al. 2010). Nevertheless, chemotherapy medication such as doxorubicin and other chemotherapeutic agents show undesirable health effects in humans (Sancar et al. 2014). Consequently, one can define that the potential novel therapeutic candidates recognition is urgently required as for new inhibitors invention against MK-2. According to the previous studies, the pyrrolopyridine inhibitors (PYPs) have been recognized as familiar MK-2 inhibitors and show the good antagonization reaction by binding with MK-2 lead to control rheumatoid arthritis, malignancy, etc. (Anderson et al. 2007). However, toxicity potentials, pharmacological profiles and other molecular mechanisms of PYPs related to MK-2 are still unclear and disputable.

Novel drug discovery is an impending task for new drug invention; however, it involves high animal usage, laboratory wastes, time and cost. Combinatorial chemistry and computational systems biology approaches have been developed for therapeutic prospective novel molecules recognition with low time and cost (Sliwoski et al. 2013; Phatak et al. 2009). In the early stage of drug discovery, the structure–activity relationship (SAR)-based clustering and classification analysis is used as a promising approach. That is imposed in obtaining new druggables by structural similarity assessment during facilitating hit examination and computations from the millions of compounds (Kim et al. 2015a, b). Bioactivity data analysis can fundamentally provide checkpoints to propose novel molecular scaffolds with similar pharmacological behaviors of known compounds. Therefore, the defined molecular candidates may be acting as potential

next-generation inhibitors toward possible therapeutic targets (Putz et al. 2016).

In this study, we focused on bioactive PYP structural analogs evaluation and inhibition profile elucidation for MK-2 by using *in silico* approaches. SAR-based bioassays scrutiny and clustering approaches have been used for estimating 197 PYP inhibitors from the PubChem database. Computational chemoinformatics strategies that include molecular modeling, structure-based virtual screening and flexible docking were employed to forecast the binding affinity of PYPs with MK-2. In this study, to extend its scope of docking studies, comparative clustering analysis was established, which reveals the structural similarities of the best ligands with known inhibitors. Moreover, ligands contain anticancer properties, adverse health effects and ADMET features were analyzed to potentiate them as possible drug-like molecules. Efforts were made for PYP inhibitor screening to identify the new lead candidates with multiple screening approaches; thus, each and every filtration was cautiously judged before going to the next phase. Finally, distinct new molecular candidates may be used as templates for acquiring new scaffold entity, which can ameliorate the diverse p58 and p38 associated unregulated signaling pathways by mitigating MK-2. Altogether, the described methods and provided pinpoints could be useful as a virtual road map for prediction and analysis of potential novel templates towards therapeutic targets.

Materials and methods

Ligands prediction

The potential 197 pyrrolopyridine (PYPs) structural derivatives were virtually screened from PubChem database (Kim et al. 2015). Three-dimensional structure optimization approaches were built up to optimize the selected molecules by using ChemAxon standardizer module. Including the adopted bioavailability features, the protonation condition of compounds was adjusted to the neutral point (pH 7.0) for sustaining the flexibility and complexity of H-bonds, and adjustable analogs were prepared as an intent inhibitor dataset.

SAR-based bioactivity analysis

Chemical structure clustering is a fundamental approach to distinguish molecule similarity to elucidate the pharmacological actions of chemical compounds. In this study, the prepared dataset molecules were subjected to structure–activity relationship (SAR)-based clustering approach with regard to structural (fingerprints) similarities and 3D-structure optimization algorithms by using the PubChem

structure-clustering module (Kim et al. 2015a, b). Here, for preferred scaffolds, 318 bioassays (478 bioactivity end points) were investigated against 2265 to describe protein targets from the bioactivity database. Further, fingerprint-based hierarchical clustering was employed to categorize the compounds with retrieved experimental evidences. Moreover, 2D-fingerprint and 3D-structure optimization algorithms were established and the consequences expressed as dendrograms. These indicate how well our training set compounds were projected with bioassays data and to the other obscure bioactive compounds.

Protein structure preparation

From the bioassay data, a pyrrolopyridine inhibitor bound to MK-2 receptor crystal structure (PDB code: 2P3G) was preferred and predicted from the Protein Data Bank (PDB) (Anderson et al. 2007). The stereochemical quality appraisals discovered higher rates of structural conflicts in the crystal structure. However, the optimized high-resolution structure prophecy is more helpful for further structure-based virtual screening approaches. Subsequently, the optimized molecular-modeling method was employed to prepare a new reliable structure from native target protein. For structural optimization, molecular dynamics (MDs) and conjugate gradients (CGs) minimization with simulated annealing (SA) algorithms were applied and 100 new models of MK-2 were predicted using MODELLER v12.1 modeling package (Šali and Blundell 1993). Among these, one model was derived from discrete optimized protein energy (DOPE) score, and again 100 new structures of the selected one were obtained with the optimally satisfied special restraint method (variable target function method (VTFM)) (Šali and Blundell 1993). Further, other structural conflicts of the model were reduced with protein geometry optimization by employing GROMACS96 force field using Swiss PDB-viewer v4.1 program (Kaplan 2001) to obtain 0.05 kcal/mol energy congeners. The stereochemical quality of the streamlined model was judged by PROCHECK (Ramachandran plot analysis) (Laskowski et al. 1993), Verify_3D (Lüthy et al. 1992) and ERRAT programs (Colovos and Yeates 1993), respectively. Finally, the selected model was used as a virtual road map to obtain efficient PYP inhibitors through screening and docking examinations during in silico structure-based drug repositioning strategies.

Virtual screening and molecular docking

High potential druggable pockets of MK-2 were identified by AutoLigand module, which can provide flexible pocket residues, fill volume and shape related information. By utilizing empirical AutoDock energies, active site dimensions were placed at 83.21 (25), 261.63 (25) and 105.87 (25) Å of

the *x*, *y*, *z* center and axis sizes, correspondingly separated by 0.375 Å grid spacing. As per these specifications, the potential active site was generated by using AutoGrid v4.0 module. Further, compounds consisting bond errors were relaxed by applying the universal force fields (UFF), followed by energy minimization using OpenBabel module of the PyRx virtual screening program (Dallakyan and Olson 2014), and optimized leads were prepared as a seed-structures library. Next, the dataset compounds were virtually screened against the binding pocket of MK-2 protein by using AutoDock Vina software through the PyRx program. For the docking simulations, the Lamarckian genetic algorithm (LGA) and Solis-wets algorithm were used with empirical free energy function and maximum 100 exhaustiveness; and the docking simulation protocols were obtained from our previously published data (Kumar et al. 2017; Bommu et al. 2017). Subsequently, by using the comparative docking analysis with a known inhibitor, highly possible molecules were scrutinized and taken into consideration. Afterward, they were subjected to flexible docking (adaptable residues were used) approach with threefold docking replications. Finally, predictable candidates can be preferred as the best competitors for future immune checkpoints.

Probable toxicity asset estimation and evaluation

Toxicity assets of the ligands such as mutagenicity, carcinogenicity, developmental toxicity and skin sensitization features were assessed by CAESAR models using VEGA-NIC v1.4.1 program (Kumar et al. 2017; Bommu et al. 2017; Tusar et al. 2012). Here, we screened the approved PYPs against specified end points by assessing with experimental training and test set compounds data, to define the 'active' or 'inactive' state of compounds with a computational value range of 0-1. In addition, a representation space of the applicability domain (AD) is compound related training set information, which reveals how many times a compound is repeated and projected with PYPs dataset molecules. As per the end point analysis, 2 × 2 (positive and negative labels) confusion matrix was analyzed from the binary vector data of QSAR models that is defined by predicting the sensitivity (SE), specificity (SP), accuracy (ACC) and Matthews correlation coefficient (MCC) estimates.

In silico anticancer activity prediction

Anticancer activities of selected lead molecules were predicted from the robust CancerIN program (Singh et al. 2016). It forecasts the molecular activities alongside experimental data of NCI-60 cancer cell lines and provides reliable scores from different databases using similarity-based QSAR approach. Therefore, we used the Scan Library module to

represent the anticancer properties with potent drug scores of the selected PYP inhibitors.

ADMET and drug-likeness screening

The physicochemical properties (Lipinski's rule of five principles), pharmacokinetics (PKs), drug-likeness, medicinal chemistry, synthesis accessibility and possible molecular descriptors of the defined PYPs were predicted from the robust SwissADME program (Daina et al. 2017). The SAR-based absorption, distribution, metabolism, excretion and toxicity properties (ADMET properties) of PYPs were also evaluated by the admetSAR program (Cheng et al. 2012). This provides the adaptable region scores of compounds, which can potentiate them as novel drug candidates for future wet-biology experiments.

Results and discussion

Dataset optimization, curation and preparation

Initial data curation is an essential task for novel chemometrics approaches, which optimizes the drug-related analogs. Hence, a set of 197 PYP inhibitors was subjected to the structure optimization method, which exposed effective ligands from the dataset. Among the 197 scaffolds (Table S1), 194 compounds were shown to have the most reliable 3D structures, and the other 3 compounds were discarded by the standard principles of the ChemAxon standardizer software. The adjusted protonation state of ligands (pH = 7.0) can promote the H-bonds framing with the receptor domain. Therefore, the optimized candidates were prepared as a seed structure library for further screening approaches; thus, indices scaffolds proceed as reference leads.

Fingerprint-based bioactivity examination and clustering

On the basis of structure–activity, the bioassay properties of PYPs were analyzed by employing SAR-based bioassay approach for scrutinizing the selectivity and specificity of the target compounds. The chemical structure, biological outcome and target information of active substances generated were also explored as possible clusters using the PubChem SAR module. The possible outcomes of cluster counts showed that the 478 compounds out of 490 were structurally aligned with 15 biologically active PYPs, classified as potential bioactive substances against protein targets such as 11715767 (PHA-767491), 11992146 (PYP23), 16719862 (PYP9), 16720200 (PYP25), 25190891 (PYP84), 25190890 (PYP81), 16720208 (PYP30), 16720202

(PYP26), 16720095 (PYP24), 16719970 (PYP11), 16719968 (PYP10), 11630374 (PYP scaffold 3), 49830880 (Kinome_3177), 10313607 (ChEMBL510545) and 25190897 (PYP scaffold 2) (Fig. 1a). Overall, these predictions were carefully assessed with 318 bioassays, 478 bioactivities and 2265 protein targets. The 15 PYP candidates prominently showed effective inhibitory concentration (IC_{50})

Fig. 1 SAR-based hierarchical clustering analysis of potential PYPs. **a** 15 predicted bioactive known candidates, **b** 2D Tanimoto similarity (substructure fingerprint scores)-based clustering of PYPs with known compounds, **c** 3D Tanimoto similarity (shape + feature; shape optimized; 10 generated conformers per compound)-based clustering of PYPs with known compounds

against mitogen-activated protein kinase family proteins. After investigations of all lead factors, PYP23 (11992146) from explored dataset was ranked with the best cluster class for MK-2; thus, these have been considered as probable reference drugs and target protein for structure-based drug repositioning approaches.

In addition, cluster similarity data were utilized for similarity ensemble generations between all compounds, which can provide other PYP inhibitors probably projected with known bioactive compounds with reliable Tanimoto similarity scores. If a compound has high similarity with a known bioactive scaffold, it may have similar biological action like an active dataset compound. Consequently, we implemented the 2D-fingerprint and 3D-structure optimization approaches against active PYPs through structure clustering methods.

A statistically significant score of 0.68 with 95% confidence interval was used to generate 2D-structure clusters by applying Tanimoto fingerprint similarities, and the resulting data showed 0.95 to 1.0 distance matrix. For 3D-structure clustering, 3D-shape with feature (0.88 and 1.03), ST-optimized (0.74 and 0.85) and CT-optimized (0.30 and 0.39) scores with 95% confidence interval were, respectively, used for one and ten conformers and the resulting data showed 0.80–1.4 distance matrix. According to principle loadings, 2D and 3D data were subjected to *k*-means clustering through an agglomerative hierarchical clustering algorithm, which can provide potential PYPs dataset from known compound similarity ensembles as dendrograms. As shown in Figs. 1b and 2d dendrogram, 100, 24 and 70 compounds structurally clustered with known bioactive candidates as three classes

Fig. 2 Stereochemical quality evaluations of optimized model protein. **a** Ramachandran plot statistics, **b** 3D profile scores of target and optimized model protein, **c** overall quality factor of the adjusted

model structure, **d** secondary structure superimposition of the target model and the template (red: template; green: modeled protein of MK-2)

with 0.894043–0.094043 average cutoff, which consisted of 4/100, 6/24 and 5/70 known active compounds in each class. As shown in Figs. 1c and 3d dendrogram, 61, 68 and 65 compounds structurally clustered with known bioactive candidates as three classes with 0.880429–0.080429 average cutoff, which consisted of 6/61, 7/68 and 2/65 known active compounds in each group. These consequences revealed that the equivocal weighted groups contained active data equivalent of 3D structures more reliable than that of 2D structures. Henceforth, the potential 194 3D structures were considered as a lead library for screening and molecular docking strategies.

Protein structure optimization and validation

Since the SAR-based bioactivity exploration, PYP inhibitors have effectively interacted with MAP kinase family proteins, which show potent anticancer activities to ameliorate

peculiar malignant disorders. Recent studies have reported that PYP23 mitigates MK-2 action, which reduces tumor necrosis factor α (TNF α) secretion and controls malignancy and rheumatoid arthritis. These investigations were carried out in vivo and in vitro experiments. Additionally, PYPs potentially showed anticancer effects by interfering with MAP kinase pathway proteins and acted as a barrier to mitigate unaccounted cell signals in malignant cells (Fiore et al. 2015; Anderson et al. 2007; Leelananda and Lindert 2016).

According to examinations and crucial viewpoints, human MK-2 protein (PDB ID: 2P3G) consisting of pyrrolopyridine inhibitor (F10) crystal structure was chosen as the therapeutic target for PYPs dataset screening. Protein quality was first verified using PROCHECK, Verify_3D, and ERRAT programs that shows the target protein to have structural errors (Table 1). However, 2.0 Å resolution and R-factor ≥ 20.0 may be useful for ligands binding. Thus, the defined structure was subjected to structure optimization and

Fig. 3 **a** AutoLigand module revealed pocket cavity (red box) of MK-2 protein with known PYP inhibitor (F10), **b** Flexible amino acid residues were allowed into the active zone, which potentiate to PYP inhibitors binding, to prompt the complexity and stability of protein

Table 1 Stereochemical qualities of native and optimized model of MK-2 protein structures

Stereochemical quality descriptors	Prediction scores	
	Original structure (PDB code: 2P3G)	Modeled structure (Model no: MK-2-6)
Ramachandran plot statistics	69.4	93.7 ^a
% of Most favoured regions		
% of Additional allowed regions	25.4	6.3
% of Generously allowed regions	4.8%	0.0 ^a
% of Disallowed regions	0.4%	0.0 ^a
Verify_3D—% of averaged 3D-1D scores ≥ 0.2	82.98	91.88% ^a
ERRAT—overall quality factor	61.303	91.667 ^a

^aIncreased factors quality in optimized protein model

modeling approaches to reduce stereochemical clashes and to provide a valuable model for screening aspects. For protein optimization, a suitable VTF method was implemented to generate 100 new models of target MK-2 structure using Python scripts through Modeller program. Based on DOPE potentials, one model was picked, and again 100 new models were generated against the selected one using the spatial restraint method. This method is used for rectifiable C α —C α , N—C α and C—C α distances of amino acid residues, expressed as conditional probability density functions (pdfs). Finally, an optimally satisfied model (MK-2.B99990006 = − 36839.996094 DOPE) was chosen as a reliable model. The other structural conflicts of model were reduced with geometry minimization by Swiss-Pdb Viewer package.

Ramachandran plot statistics of the model was 93.7%; 6.3% of the residues were found in the most favoured and additional allowed regions, respectively, and residues were not found in generously allowed and disallowed regions (Fig. 2a). Further, 91.88% of the residues had an average of 3D-1D score, revealed from Verify_3D program. The 3D environment plot showed that the average residue scores adopted above zero (Fig. 2b) suggested that the model retains good environmental profile and fine folded side chain and main regions. The increased protein overall quality factor score of 92.126 was observed after energy minimization of the model computed by non-bonded atom–atom interactions from ERRAT program (Fig. 2c). Generally, for low-resolution (2.5 to 3 Å) structures, 91% of overall quality factors are suitable for drug discovery (Dallakyan and Olson 2014). Accordingly, our new model is good and reasonable for potent druggables prediction. The energy value of protein was 0.05 kcal/mol with 1.5 Å structural resolution and had an acceptable secondary structure-like template (Fig. 2d). Overall, protein quality assessments assured that the predicted model had better reliability than the native structure (Table 1). Thus, the high quality MK-2 model is may be utilized as a virtual path for computational screening and docking estimations.

Structure-based virtual screening and re-docking

Structure-based virtual screening and molecular docking approaches are more efficient and less expensive programs in drug discovery manual. These events are useful for understanding the small molecules interaction abilities with core cavity and functional residues of the specific protein target (Leelananda and Lindert 2016). Therefore, virtual screening and flexible re-docking processes were established for PYPs dataset screening against the active site of the MK-2 model via AutoDock Vina in PyRx program. Afterward, the best binding leads were selected while comparing with reference drug binding energy, and they followed conducting flexible re-docking with three replication frameworks.

Binding mode analysis

According to docking simulations, 13 lead scaffolds were identified and had high binding affinities than known ligands (Table S2), which include 60118340, 60118338, 60117736, 60118473, 60117986, 60118322, 60118420, 60118576, 60118671, 60118849, 87959182, 60118372 and 60118793, consisting − 10.6, − 10.5, − 10.3, − 10.2, − 10.1, − 10.1, − 10.1, − 10.1, − 10.1, − 10.1, − 10.1, − 10.0 and − 10.0 kcal/mol^{−1} $\Delta G_{\text{binding}}$ energies, respectively (Table 2). Frequency strategies obviously defined that Leu⁷⁰, Glu¹³⁹, Leu¹⁴¹, Glu¹⁴⁵, Glu¹⁹⁰, Thr²⁰⁶ and Asp²⁰⁷ residues were repeatedly found in H-bonds formation with lead scaffolds and selectively acted as H-bond acceptors or donors to from flexible interactions with reliable bond distances (Å) and bond angles (°) with ligands. Thus, the considered compounds were bound like reference inhibitors (Table 2, Table S2). From the previous enormous outcomes, terminal nitrogen and lactam nitrogen atoms of Lys⁹³ and Asp²⁰⁷ were, respectively, from hydrogen bonds with ligand oxygen atoms (Anderson et al. 2007). Other flexible van der Waals (vdW) interactions possibly sustain the complexity and stability of the protein–ligand complexes (Fig. 3). Therefore, our docking consequences of potential PYP inhibitors strongly agreed with previous binding mechanism and found novel hotspot residues of the receptor that protracts vdW and H-bond interactions.

Comparative compounds clustering analysis

Furthermore, cluster classification methods were employed to find the structural similarity of the best leads with the bio-active neighbours with binding affinities. As shown in Fig. 4, the low binding compounds such as R5, R12, and R13 were clustered as one class (class-I) and not defined any structural similarities with known inhibitors (R—rank of compounds, see Table 2). The other high binding affinity leads resembling R1, R4, R2, R10, R9, R11 (PYP26, ChEMBL510545 and Kinome_3177) and R3 R8, R6, R7, (PYP84, PYP30, PYP11, PYP10, PYP23, PYP24, PYP9, PYP25), respectively, formed as two clusters (class-II and class-III) with known inhibitors. From this analysis, only class-II and class-III compounds are proposed as novel ligands and they might act as reference skeletons. This can promote the inhibition of malignant disorders through reducing the secretion of several tumor-inducing factors (e.g., TNF α) by interfering with MK-2. High probable H bonds and vdW interactions retain the complex stability that could control single alteration to generate drug resistance. As per these consequences, R1, R2, R3, R4 and R6 compounds were projected and proposed as the best inhibitors. Molecular interactions of the considered ligands are shown in Fig. 5. The identified leads were scrutinized with further chemometric tools.

Table 2 AutoDock binding energies and flexible H-Bond interactions of best potential lead molecules

Rank	PubChem ID	$\Delta G_{\text{binding energy}}$ (kcal/mol ⁻¹)	Hydrogen bonds	Distance (Å) ^e	Angle(°) ^f
1	60118340^B	− 10.6	Leu ¹⁴¹ CAN ^d :::OC ^a	3.13	177.22
			Glu ¹⁹⁰ OE1 ^d :::NC ^a	4.02	138.67
			Thr ²⁰⁶ OG1 ^d :::NC ^a	3.88	116.52
			Asp ²⁰⁷ OD2 ^a :::NC ^{a,d}	3.27	115.61
2	60118338^B	− 10.5	Leu ⁷² CAN ^d :::NC ^a	3.58	130.44
			Glu ¹⁹⁰ OE1 ^a :::NC ^{a,d}	3.55	142.07
3	60117736^C	− 10.3	Glu ¹³⁹ CO ^a :::NC ^{a,d}	3.96	122.52
			Leu ¹⁴¹ CAN ^d :::NC ^{a,d}	3.68	128.90
			Thr ²⁰⁶ OG1 ^d :::NC ^a	3.93	165.30
			Asp ²⁰⁷ OD2 ^a :::NC ^d	3.30	131.73
4	60118473^B	− 10.2	Ilc ⁷⁴ CAN ^d :::OC ^a	3.89	149.51
			Asp ¹⁴² OD1 ^{a,d} :::CN ^{a,d}	3.23	112.51
			Asn ¹⁹¹ OD1 ^a :::CN ^{a,d}	3.20	111.34
			Asp ²⁰⁷ OD2 ^d :::CN ^d	3.63	112.65
5	60117986 ^A	− 10.1	Leu ⁷⁰ CO ^a :::NC ^d	3.42	150.99
			Leu ¹⁴¹ CAN ^d :::NC ^a	3.47	156.97
			Glu ¹⁴⁵ CN ^d :::OC ^a	3.09	160.09
			Glu ¹⁴⁵ OE2 ^{a,d} :::NC ^{a,d}	3.09	132.25
			Thr ²⁰⁶ OG1 ^d :::NC ^a	3.03	155.40
6	60118322^C	− 10.1	Glu ¹³⁹ CO ^a :::NC ^{a,d}	3.933.93	119.18
			Leu ¹⁴¹ CAN ^d :::NC ^{a,d}	3.58	127.75
			Asp ²⁰⁷ OD2 ^a :::NC ^d	3.33	119.73
7	60118420 ^C	− 10.1	Glu ¹³⁹ CO ^a :::NC ^{a,d}	3.99	112.45
			Leu ¹⁴¹ CAN ^d :::NC ^{a,d}	3.53	126.68
			Thr ²⁰⁶ OG1 ^d :::NC ^a	4.00	128.68
			Asp ²⁰⁷ OD2 ^a :::NC ^d	3.48	106.72
8	60118576 ^C	− 10.1	Glu ¹³⁹ CO ^a :::NC ^{a,d}	3.88	116.97
			Leu ¹⁴¹ CAN ^d :::NC ^{a,d}	3.54	127.03
			Thr ²⁰⁶ OG1 ^d :::NC ^a	3.89	167.89
			Asp ²⁰⁷ OD2 ^a :::NC ^d	3.58	103.20
9	60118671 ^B	− 10.1	Glu ¹⁹⁰ OE1 ^d :::NC ^a	3.05	153.78
10	60118849 ^B	− 10.1	Leu ¹⁴¹ CAN ^d :::OC ^a	2.82	123.20
			Glu ¹⁹⁰ OE1 ^a :::NC ^{a,d}	3.10	150.24
11	87959182 ^B	− 10.1	Glu ¹⁹⁰ OE1 ^d :::NC ^a	3.04	153.33
12	60118372 ^A	− 10	Leu ⁷⁰ CO ^a :::NC ^d	3.40	136.83
			Leu ⁷² CAN ^d :::OC ^a	3.21	116.63
			Glu ¹³⁹ CO ^a :::NC ^{a,d}	3.77	114.28
			Leu ¹⁴¹ CAN ^d :::NC ^d	3.45	124.73
			Thr ²⁰⁶ OG1 ^d :::NC ^a	3.78	127.55
13	60118793 ^A	− 10	Leu ⁷² CAN ^d :::OC ^a	3.15	114.51
			Glu ¹³⁹ CO ^d :::NC ^{a,d}	3.58	117.18
			Leu ¹⁴¹ CAN ^a :::NC ^{a,d}	3.47	119.94
			Thr ²⁰⁶ OG1 ^d :::NC ^a	3.72	125.46

a H-bond acceptor atoms, *d* H-bond donor atoms, *e* distance between donor and acceptor atoms, *f* angle between donor, acceptor and hydrogen atom, *bold* selected best ranked compounds

^AComparative cluster analysis (class—I)

^BClass—II

^CClass—III

QSAR modeling for toxicity end point evaluation

Adverse toxicity properties such as mutagenicity, carcinogenicity, developmental toxicity and skin sensitization were assessed for selective five chemicals through the CAESAR QSAR models implemented on VEGA platform. Moreover,

the represented molecular similarity training sets and 2 × 2 confusion matrix to analyze reliability and AD indices were documented. Overall, predictions showed that the proposed ligands were structurally found in the domain with non-toxic class, as described by AD and predicted reliability class index (Table 3 (section-I)). Confusion matrix is represented

MK-2. R1: 60118340, R2: 60118338, R3: 60117736, R4: 60118473,
R6: 60118322, F10: PYP23 (positive control)

Table 3 Adverse toxicity and anticancer properties of potential lead molecules

VEGA CAESAR—models predictions (Section-I)										CancerIN program predictions (Section-II)							
PubChem CID	Mut.		Car.		Dev.		Skin.		Pre. Pos.	Query Rank ^d	NSCID	PubChemSID	logGI ₅₀	RFW-TC	Kernel MinMax	Cdrug score	p-value ^d
	Ass.	Pre.	Ass.	Pre.	Ass.	Pre.	Ass.	Pre.									
50118,40	1 ^c	1	0 ^b	0	0.325	1 ^b	1	0 ^c	0	4	733252	48430003	− 5.805	0.412	0.364	0.15	0.3622
50118338	1 ^b	1	0 ^a	0	0.255	1 ^c	1	1 ^c	1	5	682768	519468	− 5.018	0.341	0.383	0.131	0.4382
50117736	1 ^b	1	0 ^b	0	0.325	1 ^b	1	1 ^c	1	2	400085	471347	− 5.442	0.499	0.353	0.176	0.2801
50118473	1 ^c	1	0 ^b	0	0.325	1 ^b	1	0 ^c	0	3	733252	48430003	− 5.805	0.454	0.345	0.157	0.3377
50118322	1 ^b	1	0 ^a	0	0.33	1 ^c	1	1 ^c	1	1	684481	520276	− 5.546	0.5	0.359	0.18	0.2694

Mut. mutagenicity, Car. carcinogenicity, Dev. developmental toxicity, Skin. skin sensitization, Ass. assessment, Pre. predicted activity, Pos. predicted position, 0 non-toxic, 1 toxic, logGI₅₀ mean scores from various cell line experimental data

^aGood reliable (the predicted compound is inside the applicability domain of the model)

^bModerately reliable (the predicted compound is could be inside the applicability domain of the model)

^cLow reliable (the predicted compound is outside the applicability domain of the model)

^dLow p value consists of compounds ranked as best

by high SE (1.0000, 1.0000, 1.0000 and 0.9444), ACC (0.5800, 0.2917, 1.0000, and 0.7083) and low SP (0), MCC (0) scores for each end point. Accordingly, all end point analysis sophisticatedly supported the QSAR-models representation space identified for PYPs. Thus, the five chemicals considered can be used as promising non-toxic inhibitors in future in vivo and in vitro studies.

Anticancer activities of lead scaffolds

Anticancer properties of selected molecules were defined by QSAR-based fingerprint alignment approach using the robust CancerIN program. Anticancer potency scores of compounds were foreseen by machine learning methods that assessed through different programming algorithms against experimental data from different cancer cell lines. The program displays logGI₅₀, RFW-TC, kernel-MinMax, Cdrug, and p value scores for considering compounds. These results revealed that 60118322, 60117736, 60118473, 60118340 and 60118338 are ranked based on the highest effectiveness scores with NCI and PubChem matched profiles from assorted cell lines data (Table 3 (section-II)). Among the ranking systems, compounds have – 10.1, – 10.3, – 10.2, – 10.6 and – 10.5 binding energies; however, the anticancer potency is differently found in the tool output. This signifies that matched profiles with PYP23 consisting ligands are more prospective than others. Hence, these candidates may be extensively used as future anticancer agents toward moderating the p38 and p58 pathways associated diseases.

Pharmacokinetics and ADMET evaluation

The physicochemical properties and pharmacological profiles including ADMET features were envisaged by robust SwissADME and admetSAR programs individually. The SwissADME predictions showed that five ligands obeyed Lipinski 'Rule of 5' principles. Pharmacological profiles showed that all leads had high rates of GI absorption, BBB permeant, Pgp-substrate, CYP gene family inhibition like metabolic activities with consistent Log *K_p* (cm/s) skin permeation action (Table 4 (section-I)).

The ametSAR program provided name profiles and molecular descriptors of ADMET assets, which showed that the ligands occupied absorbance values with BBB, HIA, Caco-2 permeability, P-gp substrate and inhibitor, renal organic cation transporter activities; and they have distribution efficiency with mitochondria through sub-cellular localization of the cell. Further, higher rates of metabolism activates adopted by CYP gene family resolved as similar to SwissADME program. The toxicity potentials of these compounds show low toxic effects and CYP gene inhibition efficiencies were observed from fingerprint profiles, reliable scores and generated p value, and are shown in Table 4

Table 4 Physicochemical, pharmacokinetics, medicinal chemistry and SAR-based ADMET prospects of the preferred compounds

CID	60118340	60118338	60117736	60118473	60118322
Canonical SMILES	<chem>Fc1cc(ccc1c1nccc(c1)c1cc2c([nH]1)CCNC2=O)C(=O)NC1CCCCC1</chem>	<chem>O=C1NCCCc2c1cc([nH]2)c1ccnc(c1)C1=Cc2cccc-c2CCC1</chem>	<chem>Fc1cccc1c1nccc(c1)c1cc2c([nH]1)CC(NC2=O)c1ccc-c1</chem>	<chem>O=C(c1ccc(cc1F)c1nccc(c1)c1cc2c([nH]1)CCNC2=O)NC1C-CCCC1</chem>	<chem>Fc1cccc1c1nccc(c1)c1cc2c([nH]1)C(CNC2=O)c1ccc-c1</chem>
<i>SwissADME predictions (Section-I)</i>					
Physicochemical properties					
MW (g/mol)	432.49	355.43	383.42	432.49	383.42
HBA	4	2	3	4	3
HBD	3	2	2	3	2
XLOGP3	3.5	3.49	3.94	3.5	3.87
TPSA (Å ²)	86.88	57.78	57.78	86.88	57.78
Fraction Csp3	0.32	0.22	0.08	0.32	0.08
MR	124.1	111.62	113.67	124.1	113.67
Pharmacokinetics properties					
GIA	High	High	High	High	High
BBB permeant	No	Yes	Yes	No	Yes
Pgp—substrate	Yes	Yes	Yes	Yes	Yes
CYP450 1A2 inhibitor	Yes	Yes	Yes	Yes	Yes
CYP450 2C19 inhibitor	Yes	No	Yes	Yes	Yes
CYP450 2C9 inhibitor	No	No	No	No	Yes
CYP450 2D6 inhibitor	Yes	Yes	Yes	Yes	Yes
CYP450 3A4 inhibitor	Yes	Yes	Yes	Yes	Yes
Log Kp (cm/s)	− 6.45	− 5.99	− 5.84	− 6.45	− 5.89
Bioactivity and medicinal chemistry features					
Bioavailability Score	0.55	0.55	0.55	0.55	0.55
Synthetic Accessibility	3.37	3.39	3.57	3.37	3.6
<i>admetSAR predictions provided name profiles and ADMET descriptors (Section-II)</i>					
Absorption					
BBB permeant	0.8999	0.9472	0.9697	0.8999	0.9821
HIA	0.9971	1.0000	0.9958	0.9971	1.0000
Caco-2 Permeability	0.6990	0.5526	0.6247	0.6990	0.6446
P-gp Substrate	0.5359	0.6045	0.6461	0.5359	0.7012
P-gp Inhibitor	0.6420	0.5565	0.5927	0.6420	0.6301
Renal organic cation transporter	0.6525	0.5071	0.7544	0.6525	0.7291
Distribution					
Subcellular localization	0.5483	0.5988	0.7050	0.5483	0.7511
Metabolism					
CYP450 2C9 Substrate	0.8239	0.7803	0.7887	0.8239	0.8489
CYP450 2D6 Substrate	0.7923	0.7753	0.8479	0.7923	0.8317

Table 4 (continued)

CID	60118340	60118338	60117736	60118473	60118322
Canonical SMILES	<chem>Fc1cc(ccc1c1nccc(c1)c1cc2c([nH]1)CCNC2=O)C(=O)NC1CCCCC1</chem>	<chem>O=C1NCCc2c1cc([nH]2)c1ccnc(c1)C1=Cc2cccc-c2CCCC1</chem>	<chem>Fc1cccc1c1nccc(c1)c1cc2c([nH]1)CC(NC2=O)c1ccccc1</chem>	<chem>O=C(c1cccc(cc1F)c1nccc(c1)c1cc2c([nH]1)CCNC2=O)NC1CCCCC1</chem>	<chem>Fc1cccc1c1nccc(c1)c1cc2c([nH]1)C(CNC2=O)c1ccccc1</chem>
CYP450 3A4 Substrate	0.5533	0.5953	0.5460	0.5533	0.5000
CYP450 1A2 Inhibitor	0.6350	0.6481	0.7997	0.6350	0.8003
CYP450 2C9 Inhibitor	0.6184	0.5000	0.7038	0.6184	0.7044
CYP450 2D6 Inhibitor	0.7725	0.5307	0.6667	0.7725	0.8081
CYP450 2C19 Inhibitor	0.6680	0.6227	0.6705	0.6680	0.5000
CYP450 3A4 Inhibitor	0.6291	0.6690	0.7761	0.6291	0.6081
CYP Inhibitory Promiscuity	0.6790	0.8541	0.8845	0.6790	0.7632
Toxicity					
Human Ether-a-go-go-related gene inhibition	0.9734 0.8613	0.9886 0.7362	0.9956 0.5136	0.9734 0.8613	0.9950 0.7134
AMES Toxicity	0.6501	0.6974	0.7436	0.6501	0.6370
Carcinogens	0.9046	0.9590	0.8966	0.9046	0.9118
Fish toxicity	0.9087	0.5914	0.8653	0.9087	0.7198
Tetrahymena pyriformis toxicity	0.9263	0.6787	0.9213	0.9263	0.9399
Honey bee toxicity	0.8741	0.6848	0.8337	0.8741	0.8118
Biodegradation	1.0000	0.9730	1.0000	1.0000	1.0000
Acute oral toxicity	0.5202	0.5679	0.5355	0.5202	0.5616
Carcinogenicity (Three-class)	0.6449	0.6805	0.5791	0.6449	0.6709

MW molecular weight, *HBA* hydrogen bond acceptors, *HBD* hydrogen bond donors, *XLOGP3* octanol–water partition coefficient, *TPSA* ($\text{\AA}^2$) topological polar surface area, *MR* molar refractivity, *GIA* gastrointestinal absorption, *BBB* permeant: blood–brain barrier permeant, *Pgp* P-glycoprotein, *HIA* human intestinal absorption

(section-II). These results provided the five ligands that exhibit biological processes as known inhibitors. Thus, they could be utilized as core templates to discover novel human MK-2 inhibitors like p38 α MAPK pathway mediators.

Conclusion

In conclusion, for the identification of novel PYP inhibitors against human MK-2, we designed new path of computational systems biology tools. Modeling and multiple screening strategies revealed that streamlined MK-2 model and SAR-based clustering approaches are potential for hopeful drug-like molecule discovery in the early stage of drug discovery. Docking, comparative clustering and

QSAR modeling consequences revealed that the five PYP inhibitors have high binding affinity than known PYP inhibitors with reliable ADMET and non-toxic properties. Therefore, they may possess similar biological behavior as acknowledged inhibitors. Finally, clustering, high-throughput screening and in silico QSAR-based strategies were capable of predicting novel non-toxic, drug-like anticancer agents, which can act as potential inhibitors against cancer biomarkers (MK-2, VEGFR-2 and EGFR) for malignant disorders therapy.

Acknowledgements K. Kranthi Kumar (F1-17.1/2013-14/RGNF-2013-14-SC-AND-40670/(SAIII/Website)) and B. Uma Devi (F1-17.1/2012-13/RGNF-2012-13-SC-AND-32761/(SA-III/WEBSITE)) are thankful to University Grants Commission (UGC) New Delhi, India, for awarding research fellowship, Rajiv Gandhi National Fellowship (RGNF/SRF), to carry out this work.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

References

- Anderson DR, Meyers MJ, Vernier WF et al (2007) Pyrrolopyridine Inhibitors of mitogen-activated protein kinase-activated protein kinase 2 (MK-2). *J Med Chem* 50:2647–2654
- Anton R, Bauer SM, Peter RWEF, Keck et al (2014) A p38 substrate-specific MK2-EGFP translocation assay for identification and validation of new p38 inhibitors in living cells: a comprising alternative for acquisition of cellular p38 inhibition Data. *PLoS One*. <https://doi.org/10.1371/journal.pone.0095641>
- Bommu UD, Konidala KK, Pabbaraju N, Yeguvapalli S (2017) Ligand-based virtual screening, molecular docking, QSAR and pharmacophore analysis of quercetin-associated potential novel analogs against epidermal growth factor receptor. *J Recept Signal Transduct Res* 37:600–610
- Cargnello M, Roux PP (2011) Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. *Microbiol Mol Biol Rev* 75:50–83
- Chen K, Huang J, Gong W et al (2007) Toll-like receptors in inflammation, infection and cancer. *Int Immunopharmacol* 7:1271–1285
- Cheng F, Li W, Zhou Y et al (2012) admetSAR: a comprehensive source and free tool for assessment of chemical ADMET properties. *J Chem Inf Model* 52:3099–3105
- Chikanza IC (2004) Corticosteroid resistance in rheumatoid arthritis: molecular and cellular perspectives. *Rheumatology* 43:1337–1345
- Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* 2:1511–1519
- Daina A, Michielin O, Zoete V (2017) SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep* 7:42717
- Dallakyan S, Olson AJ (2014) Small-molecule library screening by docking with PyRx. *Methods Mol Biol* 1263:243–250
- Dobbelstein M, Moll U (2014) Targeting tumour-supportive cellular machineries in anticancer drug development. *Nat Rev Drug Discov* 13:179–196
- Fiore M, Forli S, Manetti F (2015) Targeting mitogen-activated protein kinase-activated protein kinase 2 (MAPKAPK2, MK2): medicinal chemistry efforts to lead small molecule inhibitors to clinical trials. *J Med Chem* 59:3609–3634
- Gurgis FMS, Ziazariis W, Munoz L (2013) Mitogen-activated protein kinase-activated protein kinase 2 in neuroinflammation, heat shock protein 27 phosphorylation, and cell cycle: role and targeting. *Mol Pharmacol* 85(2):345–356
- Höpker K, Hagmann H, Khurshid S et al (2012) AATF/Che-1 acts as a phosphorylation-dependent molecular modulator to repress p53-driven apoptosis. *EMBO J* 31:3961–3975
- Kaplan W (2001) Swiss-PDB viewer (deep view). *Brief Bioinform* 2:195–197
- Kim S, Han L, Yu B et al (2015a) PubChem structure–activity relationship (SAR) clusters. *J Chem Inform*. <https://doi.org/10.1186/s13321-015-0070-x>
- Kim S, Thiessen PA, Bolton EE et al (2015b) PubChem substance and compound databases. *Nucleic Acids Res*. <https://doi.org/10.1093/nar/gkv951>
- Kumar KK, Devi BU, Neeraja P (2017) Integration of in silico approaches to determination of endocrine-disrupting perfluorinated chemicals binding potency with steroidogenic acute regulatory protein. *Biochem Biophys Res Commun* 491:1007–1014
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26:283–291
- Leelananda SP, Lindert S (2016) Computational methods in drug discovery. *Beilstein J Org Chem* 12:2694–2718
- Lüthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. *Nature* 356:83–85
- Manke IA, Nguyen A, Lim D et al (2005) MAPKAP kinase-2 is a cell cycle checkpoint kinase that regulates the G2/M transition and S phase progression in response to UV irradiation. *Mol Cell* 17:37–48
- Mourey RJ, Burnette BL, Brustkern SJ et al (2010) A benzothiophene inhibitor of mitogen-activated protein kinase-activated protein kinase 2 inhibits tumor necrosis factor production and has oral anti-inflammatory efficacy in acute and chronic models of inflammation. *J Pharmacol Exp Ther* 333:797–807
- Phatak SS, Stephan CC, Cavasotto CN (2009) High-throughput and in silico screenings in drug discovery. *Expert Opin Drug Discov* 4:947–959
- Putz M, Duda-Seiman C, Duda-Seiman D et al (2016) Chemical structure-biological activity models for pharmacophores' 3D-interactions. *Int J Mol Sci* 17:1087
- Šali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779–815
- Sancar A, Lindsey-Boltz LA, Gaddameedhi S et al (2014) Circadian clock, cancer, and chemotherapy. *Biochemistry* 54:110–123
- Singh H, Kumar R, Singh S et al (2016) Prediction of anticancer molecules using hybrid model developed on molecules screened against NCI-60 cancer cell lines. *BMC Cancer*. <https://doi.org/10.1186/s12885-016-2082-y>
- Sliwoski G, Kothiwale S, Meiler J, Lowe EW (2013) Computational methods in drug discovery. *Pharmacol Rev* 66:334–395
- Tremple N, Dave-Coll N, Nebreda AR (2013) SnapShot: p38 MAPK substrates. *Cell*. <https://doi.org/10.1016/j.cell.2013.01.047>
- Tusar M, Minovski N, Fjodorova N, Novic M (2012) In silico assessment of adverse effects of a large set of 6-fluoroquinolones obtained from a study of tuberculosis chemotherapy. *Curr Drug Saf* 7:313–320
- Xu L, Chen S, Bergan RC (2006) MAPKAPK2 and HSP27 are downstream effectors of p38 MAP kinase-mediated matrix metalloproteinase type 2 activation and cell invasion in human prostate cancer. *Oncogene* 25:2987–2998