RSC Advances

This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. This *Accepted Manuscript* will be replaced by the edited, formatted and paginated article as soon as this is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

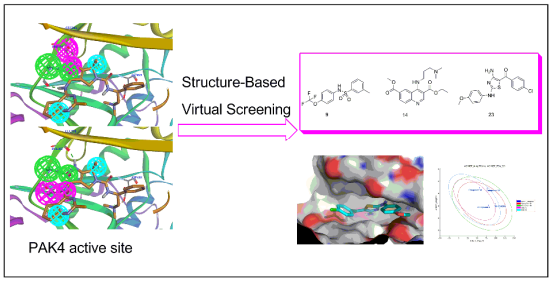
Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Table of contents

Novel PAK4 inhibitors were discovered using structure-based virtual screening approach for further chemical modification.



RSC Advances Accepted Manuscript

RSC Advances

RSCPublishing

PAPER

Cite this: DOI: 10.1039/xoxxoooooo

Received ooth January 2012,
Accepted ooth January 2012

DOI: 10.1039/xoxxoooooo

www.rsc.org/

Structure-based virtual screening and ADME/T-based profiling for low molecular weight chemical starting points as p21-activated kinase 4 inhibitors†

Ruijuan Li, [a] Xiaolin Su, [a] Zheng Chen, [a] Wanxu Huang, [b] Yali Wang, [a] Kaibo Wang, [a] Bin Lin, [a] Jian Wang *[a] and Maosheng Cheng *[a]

Structure-based virtual screening approach to targeting p21-activated kinase 4 (PAK4) was performed to identify good chemical starting points for medicinal chemistry. A pre-filtrated database was screened against two designed PAK4 pharmacophores, and the pharmacophore search hits were docked into a PAK4 crystal structure. Twenty-seven compounds were then selected for in *vitro* PAK4 inhibition assay, and results showed three compounds exhibiting a micro-molar IC_{50} in a dose-response assay. Interactive modes of the three compounds were studied and showed good binding modes in the PAK4 active site. Calculated ADME/T properties of the three hits were also analyzed and showed good drug-like properties. The results of in *vitro* PAK4 inhibition assay, interactive mode study and ADME/T prediction revealed that the three compounds have potential PAK4 inhibitory activities and can be further optimized and developed as lead compounds.

Introduction

P21-activated kinase 4 (PAK4), a serine/threonine protein kinase, was originally identified as an effector protein for Rho-family GTPases.[1] PAK4 plays an important role in regulating cytoskeletal organization, cell proliferation, migration, and nuclear signaling.[2] In addition, PAK4 is over-expressed and amplified in a variety of human cancers;[3] *in vitro* and *in vivo* studies using genetically engineered systems as well as small-molecule tool compounds suggest that PAK4 can be used therapeutically as an oncology target.[4] Thus, identifying PAK4 inhibitors to treat cancer has attracted a great deal of attention in both academia and pharmaceutical industry.

So far, there are several chemical families of PAK4 inhibitors reported such as indolocarbazole-based inhibitor (Staurosporine, 100% PAK4 inhibition at 10 μM),[5] aminopyrazole-based inhibitor (PF-3758309, Ki = 19 nM),[6] tri-substituted purine analogue (CGP74514A, 44% PAK4 inhibition at 10 μM),[5] aminopyrimidine-based inhibitor (FRAX486, IC_{50} = 779 nM) and tri-substituted 1, 3, 5- triazine analogue (KY-04031, IC_{50} = 0.79 ± 0.05 μM).[7] These inhibitors were discovered using high-throughput screening against PAK4. Another PAK4 inhibitor LCH-7749944 (IC_{50} = 14.93 μM), a quinazoline derivative, was rationally designed through a structure-based approach.[8] Most of the available compounds lack satisfactory selectivity for PAK4, physiochemical and pharmacological properties required for druggability. There is still enormous need for novel classes of PAK4 inhibitors rising to these challenges.

With the goals to identify a starting scaffold for design of novel PAK4 inhibitors, we performed a structure-based virtual screening and *in vitro* test against SPECS database utilizing a range of protocols, such as database pre-filtering, pharmacophore screening, molecular docking, result postprocessing, and *in vitro* bioassays. Among the twenty-seven compounds selected by our virtual screening approach, fourteen hits were found with PAK4 inhibition more than 50% at a concentration of 100 μM (51.85% hit rate) and three novel hits showed micro-molar IC_{50} values in dose-response assay. The binding modes of PAK4 with the three hits were further analyzed. Additionally, the three hits were predicted for their ADME/T (absorption, distribution, metabolism, excretion, and toxicity) profiles and exhibited good drug-like properties. This study reports novel and low molecular weight starting points that can be further developed as more potent PAK4 inhibitors.

Results and discussion

Structure-based virtual screening

Computational technique like structure-based virtual screening — developed primarily in the pharmaceutical or biotechnology industries — has been successfully applied for the generation of hit

[a] Key Laboratory of Structure-Based Drug Design and Discovery of Ministry of Education, Shenyang Pharmaceutical University, Shenyang 110016, China. E-Mails: jianwang@syphu.edu.cn (Jian Wang); mscheng@syphu.edu.cn (Maosheng Cheng); Tel.: +86-24-23986419

[b] School of Life Science and Bio-pharmaceutics, Shenyang Pharmaceutical University, Shenyang 110016, China.

† Electronic Supplementary Information (ESI) available: Physicochemical properties of the 27 hit compounds in Tables S1. Pharmacophore mapping fit values and docked binding affinities of the 27 hit compounds in Table S2. See DOI: 10.1039/b000000x/

RSC Advances Accepted Manuscript

and lead structure candidates.[9] Thus we performed a cascaded pharmacophore matching and docking searching of commercially available library to find chemical leads.

In this research, two in-house pharmacophore models were both used as query in the initial library screening. The two pharmacophore models were previously constructed by clustering common chemical features from docked poses of nine PAK4 inhibitors and were validated as ideal pharmacophore models that not only properly reflect the structural information inside the binding site of PAK4, but also contain chemical features from more active compounds.[10] To further analyze the pharmacophore features, the two pharmacophore models herein were superimposed on the PAK4 binding site. The high resolution crystal structure of PAK4 with potent ligand (PF-3758309) (PDB code: 2X4Z) was chosen for modeling, and significant result of the cross-docking simulation part previously reported.[10]

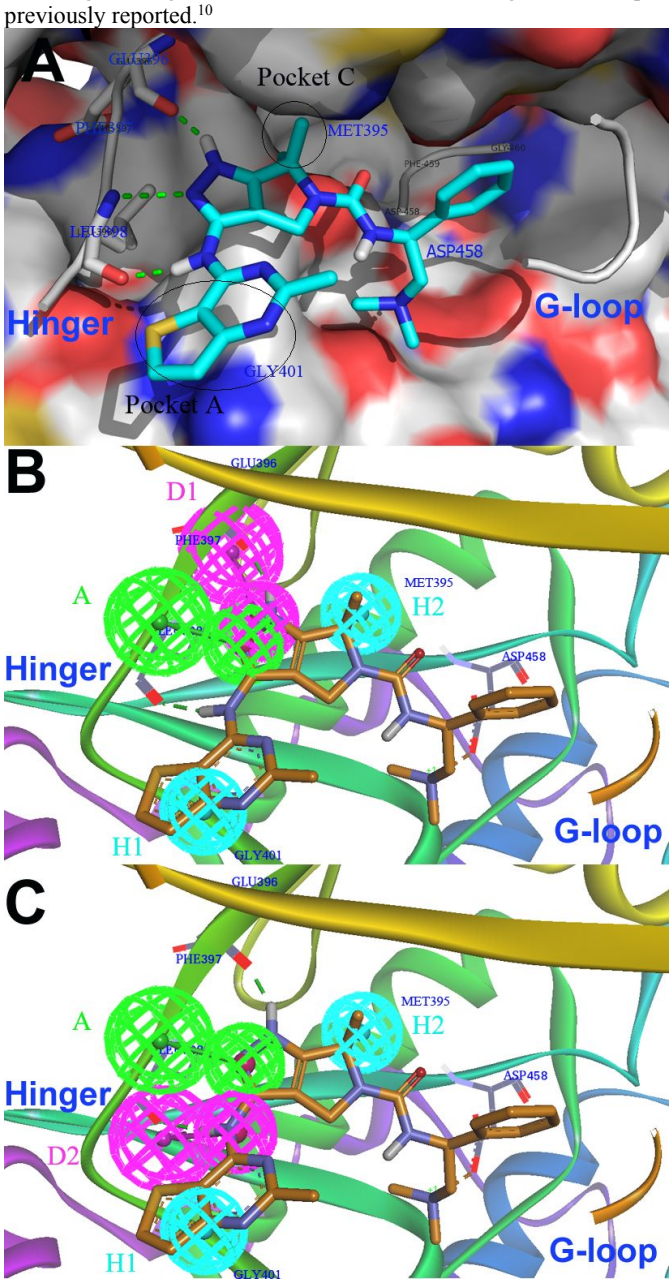


Fig. 1 Pharmacophoric description of ligand in PAK4 active site. (**A**) Crystal structure of PAK4 bound to PF-3758309. Pymol was used to analyze the binding pattern. The compound is shown using cyan sticks. An electrostatic potential surface was added to the active site around the compound. The characteristic partial Hinger and G-loop structures are shown as secondary structures. (**B**) pharmacophore model1; (**C**) pharmacophore model 2. The pharmacophore features are color-coded as follows: the hydrogen bond acceptor is green, the hydrogen bond donor is magenta, and the hydrophobic feature is cyan.

As shown in Fig. 1A, PF-3758309 forms three contacts with the hinger region through H-bond interactions with the pyrrolopyrazole core and the amine of the thienopyrimidine ring; the dimethylamine group forms charge-charge interaction with Asp458; three hydrophobic interactions between PF-3758309 and the PAK4 protein are observed, including (i) thiophene and Gly401 located in pocket A, (ii) gem-dimethyl group of the pyrrolopyrazole core and Met395 contained in pocket C, and (iii) benzyl moiety and G-loop. Among the seven contacts, the charge-charge interaction as well as the hydrophobic interaction between benzyl moiety and G-loop were previously identified as secondary rather than essential inside the binding site of PAK4. Moreover, receptor-based pharmacophore hypothesis is directly derived from the seven interactions between PAK4 and PF-3758309 complex. Since this numerous-featured model is too restrictive, five essential interactive features, including two hydrophobic interactions and three H-bonding interactions, were used as two six-point pharmacophore models for library to find diverse hits. Both models had two hydrophobic groups (H1 and H2) and one H-bond acceptor (A), and the NH of the pyrrolopyrazole core was defined as H-bond donor (D1) in pharmacophore model 1 (Fig. 1B) while the NH linker to the thienopyrimidine ring was defined as H-bond donor (D2) in pharmacophore model 2 (Fig. 1C). The usage of two six-point pharmacophore models in library search can greatly increase structural diversity of the matched compounds rather than a single eight-point pharmacophore model.

A subset of 325 881 commercially available molecules from SPECS was used for virtual screening. In order to remove unreasonable molecules and save computational cost, we applied Lipinski's Rule of Five for filtering unwanted physical and chemical properties, resulting in 159 691 compounds kept in the chemical library. Generating diverse conformations for these compounds increased the number of structures to approximately 9 421 769. We then used this multiple-conformation database to match against the two four-point pharmacophores. Based on the fit values (≥ 2) for a compound matching against either pharmacophore, a database with 3 839 compounds containing the pharmacophore features was generated. It is efficient to apply this hierarchical strategy to reducing the number of "non-hits" passed through the molecular docking stage. The hits from the pharmacophore screening were docked into PAK4 crystal structure and scored by XP Glidescore function. According to the predicted binding energy values (≤ -6), top 204 compounds were picked out, and their docked poses were rigidly matched against the two pharmacophores, to ensure the docked poses were consistent with the original poses generated from pharmacophore match, so as to increase the overall computational accuracy. Finally, twenty-seven compounds were chosen and purchased for *in vitro* bioassay test. The workflow of the virtual screening was shown in Fig. 2.

RSC Advances Accepted Manuscript

This journal is © The Royal Society of Chemistry 2012

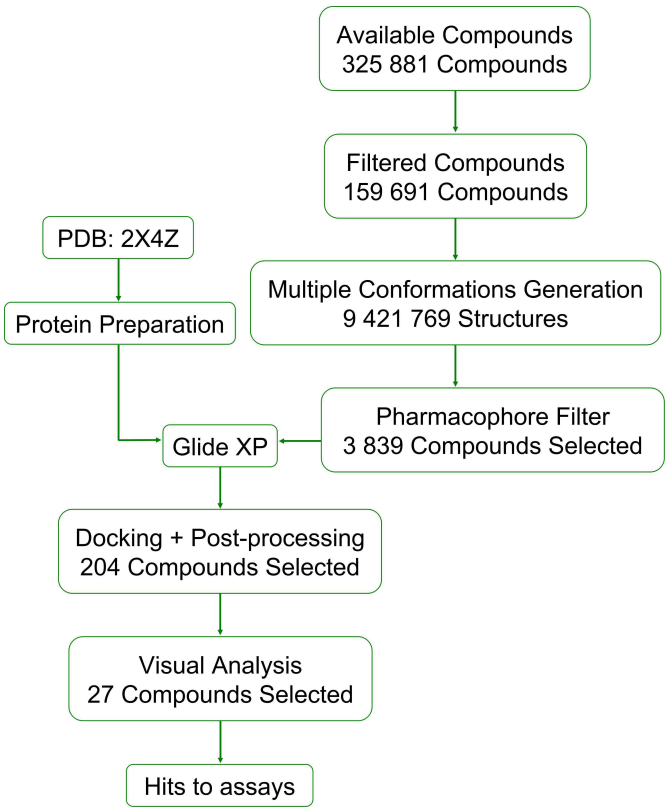


Fig. 2 Workflow of the virtual screening protocol.

The twenty-seven compounds used in the enzyme-based assays are listed in Table S1, and their fit values of pharmacophore match and docked binding affinities are summarized in Table S2. All of the compounds, except **7** and **26**, are low molecular weight between 210 to 410, which are suitable for further optimization.

Bioassay validation

The preliminary screening using the biochemical PAK4 activity assays and twenty-seven compounds is shown in Fig. 3.

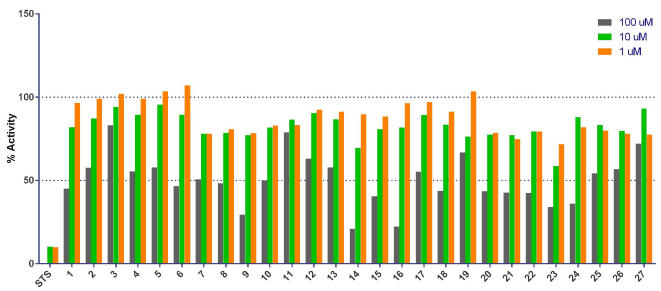


Fig. 3 Result of the preliminary bioassay screening. Biochemical PAK4 activity assays for compounds 1-27 using human PAK4. Staurosporine was used as a reference drug.

Three compounds exhibited less than 30% PAK4 activity; eleven compounds exhibited between 30% and 50% PAK4 activity; and eleven compounds exhibited between 50% and 75% PAK4 activity at 100 µM. Two compounds exhibited between 50% and 75% PAK4 activity at 10 µM. Two compounds exhibited between 50% and 75% PAK4 activity at 1 µM. The 2D structures of the fourteen compounds with less than 50% PAK4 activity at 100 µM are shown in Fig. 4 with the activity data in brackets.

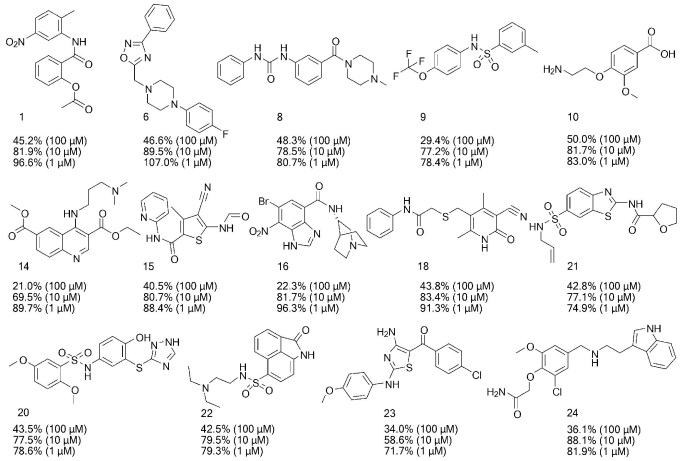
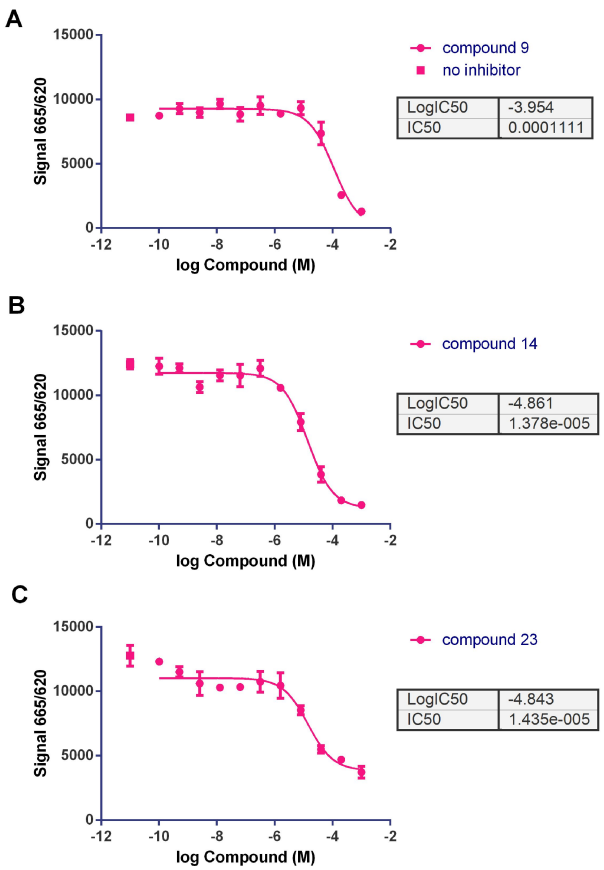


Fig. 4 Selection of compounds from virtual screening. (bracketed figure denotes % PAK4 activity at 100 µM, 10 µM, and 1 µM of compounds).

These starting points differ from most published PAK4 inhibitors. The activity cutoff for the assay is based on the data, and any compound with less than 30% PAK4 activity was defined as active. Compounds with between 30% and 50% PAK4 activity were designated as moderately active, and compounds with between 50% and 75% PAK4 activity were designed as weakly active; the remainder were considered inactive. Using these criteria, three representative compounds were further tested using a dose-response experiment. The IC_{50} results are shown in Fig. 5.



RSC Advances Accepted Manuscript

Fig. 5 Inhibitory curves of the three potential compounds toward PAK4.

The IC_{50} values for compounds **9**, **14**, and **23** are 111 μM, 14 μM, and 14 μM, respectively, indicating that these low molecular weight compounds are good starting points for further structural modification. Their interactive modes with PAK4 are clarified in detail as below.

Interactions between the compounds and PAK4

The binding modes of compounds **9**, **14** and **23** to PAK4 were analyzed using PyMol and Discovery Studio Visualizer 4.0. The results are shown in Fig. 6. In general, all of the compounds adopted extended conformations to perfectly fit the PAK4 active site shape and interacted with the surrounding residues in PAK4.

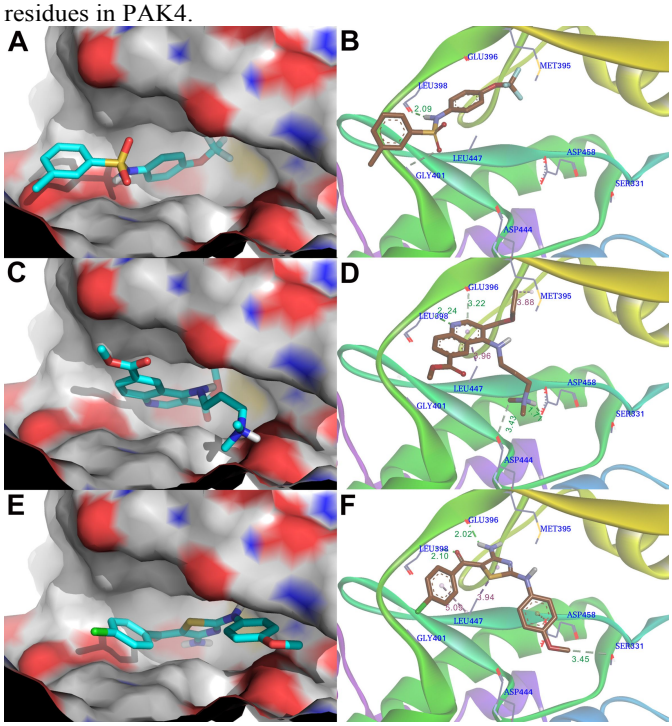


Fig. 6 Binding patterns of the three potential compounds with PAK4. PAK4 pocket is shown as a surface that is color-coded based on hydrophobicity. The secondary structure of PAK4 is shown as a color rainbow. The classical H-bond is shown using green dotted lines, a non-classic H-bond is shown using pale-cyan dotted lines, and hydrophobic interactions are shown using pink dotted lines. The distances of the observed interactions are specified. (**A**) Compound **9** binds in the PAK4 pocket. (**B**) The interactive mode between compound **9** and PAK4 active site. (**C**) Compound **14** binds in the PAK4 pocket. (**D**) The interactive mode between compound **14** and PAK4 active site. (**E**) Compound **23** binds in the PAK4 pocket. (**F**) The interactive mode between compound **23** and PAK4 active site.

The binding mode for the PAK4 active site and inhibitor **9** is shown in Figure 6A and 6B. The 4-trifluoromethoxyl aminobenzene extends into the inner portion of the PAK4 active site, and the remainder of the scaffold extends to the outer portion of the PAK4 active site. The sulfamide NH donates a classic H-bond to Leu398, and the sulfamide O accepts a non-classic carbon H-bond from Gly401.

The binding mode for compound **14** inside the PAK4 active site is shown in Figure 6C and 6D. The basic amine scaffold and methyl formate are attached to the 6-position of the quinoline ring, which extends from the PAK4 pocket; and the remaining scaffold is adopted into the inner active pocket formed by the surrounding residues in PAK4. The nitrogen atom in the quinoline ring accepts a classic H-bond from the Leu398 hydrogen atom. The hydrogen atom attached to the 2-position carbon of the quinoline ring donates a non-classic carbon H-bond to Glu396. The basic nitrogen is the terminal electron donor, and the hydrogen attached thereon donates a classic H-bond to Asp458. The hydrogen of the methyl group attached to the basic amine donates anon-classic carbon H-bond to Asp444. The quinoline ring forms a π-alkyl hydrophobic interaction with Leu447. The carbon atom of ethoxymethyl on the 3-position of the quinoline ring forms an alkyl-alkyl hydrophobic interaction with Met395, which is inside the inner portion of the PAK4 active site.

For compound **23**,the interactive mode inside the PAK4 active site is shown in Figure 6E and 6F. The compound **23** structure extends into the inner portion of the PAK4 active site, except for the 4-chlorophenyl group. The oxygen atom of the carboxyl group attached at the 5-position of the thiazole ring accepts a classic H-bond from Leu398. A hydrogen from the amino group at the thiazole ring 4-position donates a classic H-bond to Glu396. The hydrogen of the methoxyl attached at the aminobenzene 4-position donates a non-classic carbon H-bond to Ser331. The 4-methoxyl aminobenzene phenyl ring, as a π donor, forms a non-classic H-bond with Asp458. Both the 4-chlorophenyl group and thiazole ring form a π-alkyl hydrophobic interaction with Leu447.

Based on the above elucidation, it is evident that H-bonds and hydrophobic contacts are essential for interactions between compounds and PAK4. The binding modes for these compounds and PAK4 were beneficial to discern the active site regions and useful as templates for us to further develop more potent PAK4 inhibitors.

Theoretical evaluation of ADME/T properties

Pharmacokinetic properties were calculated for compounds **9**, **14**, **23**, and PF-3758309. The pharmacokinetic analysis results are shown in Fig. 7.

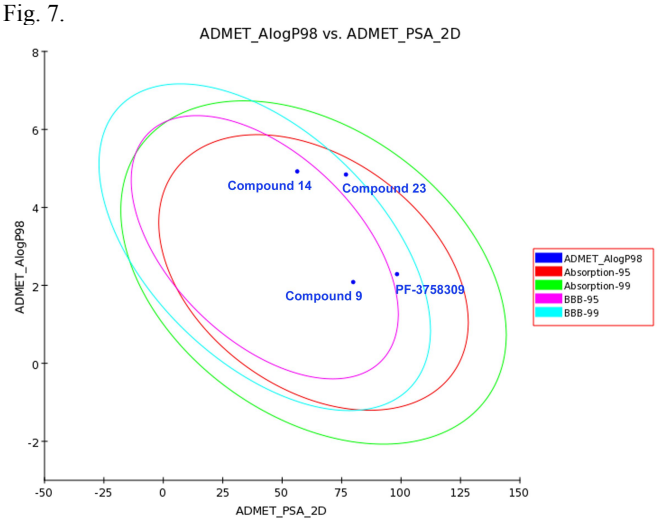


Fig. 7 Plot of PSA versus AlogP for the candidate compounds. The 95% and 99% confidence limit ellipses that correspond to the blood-brain barrier and intestinal absorption models.

Abbreviations: ADME/T, absorption, distribution, metabolism, excretion and toxicity; AlogP, the logarithm of the partition coefficient between octanol and water; PSA, polar surface area; 2D, two-dimensional; BBB, blood brain barrier.

The biplot figure shows two analogous 95% and 99% confidence ellipses for the blood-brain barrier (BBB) penetration and human

RSC Advances Accepted Manuscript

This journal is © The Royal Society of Chemistry 2012

intestinal absorption (HIA) models, respectively. All four compounds are in the 99% confidence range for BBB penetration and HIA, indicating that the computational ADME/T descriptor is a fairly reliable forecast.

The detailed results for the pharmacokinetic properties and toxicity analyses using the three selected compounds and PF-3758309 are shown in Table 1.

Table 1. In silico ADME/T prediction of the three potential compounds compared with the potent inhibitor PF-3758309.

ADME/T parameters	9	14	23	PF-3758309
AlogP98[a]	4.923	2.085	4.843	2.288
PSA[b]	56.341	79.885	76.842	98.315
Aqueous Solubility[c]	1	3	1	2
HIA[d]	0	0	0	0
PPB[e]	highly bound	highly bound	highly bound	highly bound
BBB penetration[f]	1	3	1	3
CYP450 2D6 binding[g]	0	1	0	0
Hepatotoxicity	toxic	nontoxic	toxic	toxic
DTP[h]	Toxicity	Toxicity	Toxicity	Toxicity
FDA[i] **rodent carcinogenicity**	Non-carcinogenic	Non-carcinogenic	Non-carcinogenic	Non-carcinogenic
Ames mutagenicity	Mutagenic	Mutagenic	Non-mutagenic	Non-mutagenic
Aerobic biodegradability	Non-degradable	Degradable	Non-degradable	Non-degradable
Skin sensitization	Strong-Sensitizer	Moderate-Sensitizer	Strong-Sensitizer	Strong-Sensitizer
Skin irritating	Strong-irritant	Mild-irritant	Non-irritant	Non-irritant

[a]Alog P98 (atom-based log P) ($\leqslant$ -2.0 or $\geqslant$ 7.0: very low absorption).
[b]PSA (polar surface area) (>150: very low absorption).
[c]Level of aqueous solubility predicted: 0 (extremely low), 1 (very low, but possible), 2 (low), 3 (good), 4 (optimal), 5 (too soluble), 6 (warning: molecules with one or more unknown Alog P calculations).
[d]HIA (human intestinal absorption), level of human intestinal absorption prediction: 0 (good), 1 (moderate), 2 (poor), 3 (very poor).
[e]PPB, plasma protein binding.
[f]BBB (blood brain barrier), level blood brain barrier penetration prediction: 0 (very high penetrate), 1 (high), 2 (medium), 3 (low), 4 (undefined).
[g]Prediction cytochrome P4502D6 enzyme inhibition (0: non-inhibitor; 1: inhibitor). hDTP, development toxicity potential.
[i]FDA, food and drug administration.

RSC Advances Accepted Manuscript

The AlogP was predicted to determine the compounds' hydrophilicity. A high AlogP was associated with poor absorption or permeation; thus, the value should be less than 5. The PSA is another key property linked to drug bioavailability; the passively absorbed molecules with PSA < 140 have high oral bioavailability. The results suggest that the four compounds are within these limits. All of the compounds, except **14**, show promising solubility levels for aqueous solubility. Likewise, all of the compounds can be efficiently absorbed in the human intestine and highly bound to plasma protein. Encouragingly, **9** and **23** do not exhibit high BBB penetration, and they do not inhibit cytochrome P450; thus, these two compounds can readily undergo oxidation and hydroxylation during the first phase of metabolism. Compound **14,** which includes a low BBB level and cytochrome P450 inhibition, may not be rapidly eliminated *in vivo* and prolong the action time. Although *in silico* predictions demonstrate that all of the compounds, except **14**, are hepatotoxic, further structural modifications may improve the compounds' toxicity profiles. For the toxicity risk, compound **23** is consistent with the reference molecule PF-3758309. Both compound **23** and PF-3758309 are non-carcinogenic, non-mutagenic, non-degradable, and non-irritating. Computational pharmacokinetic and toxicology studies on the active compounds **9**, **14**, and **23**, compared with PF-3758309, suggest that compound **23** can be used as good starting point for further development and designing new derivatives.

The mode of compound **23** binding to PAK4 in comparison to PF-3758309 was made by superimposing the docked pose of compound **23** onto the native pose of PF-3758309 cocrystallized in PAK4 structure (Fig. 8).

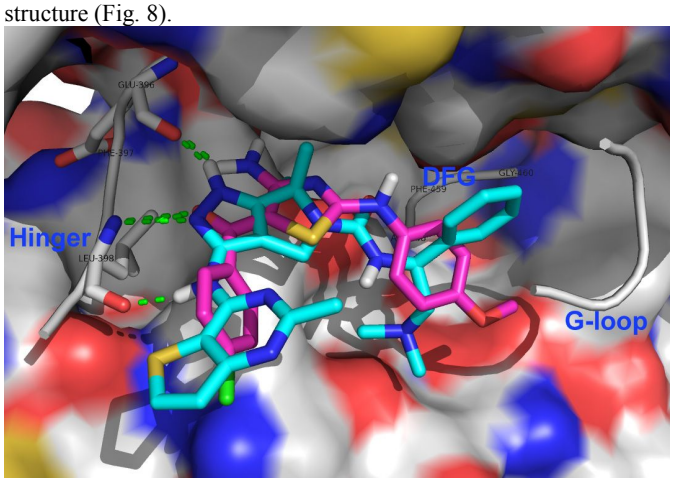


Fig. 8 Alignment between compound 23 and PF-3758309 in PAK4 active site. Compound 23 was shown as carbon in purple and PF-3758309 was shown as carbon in cyan.

Comparing compound **23** with PF-3758309, both the 4-amine group and 5-carbonyl of the thiazole ring of compound **23** mimic the pyrazole ring of PF-3758309. The 4-chlorobenzene group of compound **23** binds similarly to the pyrimidine ring of PF-3758309, occupying the entrance of PAK4 cavity. The 2-(4-methoxybenzenamine) group of thiazole ring of compound **23** extends into the g-loop region of PAK4. In summary, the compound **23** can be used as a profitable template for developing new PAK4 inhibitors.

Conclusions

PAK4 has emerged as an important cancer target, and researchers are considerably interested in developing PAK4 inhibitors as biological markers and leads for developing therapeutics. While initial approaches were based on random screening for non-selective kinase inhibitors, more recent efforts have focused on identifying new chemical scaffolds for inhibitor optimization. Such approaches have been used to identify highly selective and potent inhibitors, which serve as a basis for further inhibitor development with therapeutic applications.

In this study, three compounds with PAK4 inhibition IC_{50} values of 111 μM, 14 μM, and 14 μM were discovered using a structure-based virtual screening approach and *in vitro* bioassay validation. The binding modes and computational ADME/T properties of the three compounds were also analyzed. The three compounds exhibit ADME/T properties similar to PF-3758309, including good human intestinal absorption, high plasma protein binding, AlogP < 5, PSA < 140, and non-carcinogenic properties. The aforementioned results of *in vitro* bioassay, docking study and ADME/T prediction show that the three drug-like compounds are able to inhibit PAK4 activity and can be optimized as lead compounds with satisfactory pharmacokinetic properties.

Experimental Section

Structure-based virtual screening

All the calculations were conducted on Dell PowerEdge R900 workstation under Redhat 6.4 platform. Database preparation and pharmacophore screening were performed with Discovery Studio, version 3.0 (DS 3.0).[11] Molecular Docking was performed with Glide module in schrödinger.[12] The protein-ligand binding affinity was caluculated by XSCORE.[13] The non-bonding interactions between the inhibitors and the receptors were displayed in Discovery Studio Visualizer 4.0.

Database preparation. The publicly available database SPECS distributed by ZINC database version 12 was chosen for virtual screening.[14] The database was initially filtrated using the Lipinski's Rule of Five[15] and the Filter Molecules program in DS 3.0 to remove unreasonable molecules with unwanted physical and chemical properties. For each molecule, a set of physical and biological properties were predicted and used to assess the drug-likeness profile. Molecules with the following properties were removed by default: molecular weight > 500, logP < -4, logP > 8, hydrogen bond donors and acceptors > 12, rotatable bonds > 10, polar surface area > 140, single bond chain length > 6, chiral centers > 4, unconstrained chiral centers > 3, transition metals > 8 rings, and d-hybrids. After filtering, the remaining molecules in the database were utilized to generate multiple conformations using the conformation search and minimization program in DS 3.0. Following parameters were set: maximum conformations = 255, energy threshold = 20 kcal/mol and root-mean-square distance (RMSD) = 0.8 Å. The energy threshold is the value used to discard high-energy conformations. The maximum allowed conformations per molecule was 255 to ensure complete conformational coverage. The final 3D multiple conformations database was subjected to the pharmacophore query.

Pharmacophore screening. The two designed pharmacophores were constructed as previously reported.[10] Then the presence of these pharmacophores was searched in the 3D multiple conformations database of available compounds using Ligand Profiler protocol in DS 3.0. During the screening process, parameter settings were shown below. "Fitting Method" was flexible instead of rigid and "Maximum Omitted Features" option was set to -1 instead of 0. Molecules matching all the features of the pharmacophore model were selected for subsequent molecular docking analysis.

Molecular docking. *Structure preparation:* The crystal structure of PAK4 with PDB code 2X4Z was used for the

RSC Advances Accepted Manuscript

 This journal is © The Royal Society of Chemistry 2012

docking experiments. We processed 2X4Z using the protein preparation wizard in Schrödinger suite. First, missing residues, loop segments, and all hydrogen atoms were added. Second, the complex's protonation states were adjusted for consistency with pH 7.4. Third, hydrogen bond sampling was verified and adjusted by modifying water molecule orientations. Fourth, all hydrogen atoms and the total structure for the 2X4Z complex were evaluated for energy accuracy using an all-atom force field, OPLS_2005;[16] restrained minimization; and heavy atom convergence to a 0.3 Å RMSD. Fifth, the above prepared 2X4Z complex structure was used to generate a receptor grid file. The position of the co-crystallized ligand PF-3758309 was used to determine the active site location (x= 20.61, y= 20.87, and z= 58.45) and size (inner box= 10 Å × 10 Å × 10 Å; outer box= 20 Å × 20 Å × 20 Å). The 2X4Z grid was generated using the OPLS_2005 force field.

Docking Screening: Glide XP module in Schrödinger suite was used for docking screening since it was successful at reproducing the PAK4 crystal structure binding modes in preliminary experiments. The hits from the pharmacophore search were saved as SD file and prepared at pH 7 protonation state. Then they were imported to Glide XP to dock into the prepared receptor grid with enhanced docking precision. Top ten docked poses for each ligand were retained, with other docking parameters as default settings. The Glide XP scoring function was used to predict binding energy between compounds and PAK4.

Postprocessing of docking result and compound selection. According to the values of predicted binding energy, compounds were ranked. The top ten docked poses for each compound were used for further postprocessing analysis. The postprocessing procedure was as follows. First, all of the docked poses were rigidly matched against the two designed pharmacophores, and only the docked poses that fitted into either of the pharmacophores were accepted. Second, various measures of ligand-receptor interactions in these docked structures were analyzed and clustered manually, to reduce duplicate conformers. These measures included clashes with residues in the active site; interactions with the crucial residues (i.e., Glu396 and Leu398); hydrophobic interactions; and solvent-exposed lipophilic pocket area. After the docked poses were postprocessed, the top-ranked pose for each compound was taken and re-scored in its modeled position by XSCORE. The non-bonding interactions between the inhibitors and the receptors were displayed in Discovery Studio Visualizer 4.0.

Bioassay at the molecular level

After a virtual screening of the SPECS database, twenty-seven commercially available compounds were purchased and assessed. The PAK4 kinase assay was performed using the HTRF®KinEASE™-STK kit (Cisbio Bioassays, France) in a 384-well, low volume microplate (ThermoFisher Scientific, USA.). Purified PAK4 enzyme was purchased from Carna Biosciences (Japan).

The kinase activity was assessed using the conditions determined below. PAK4 0.0256 ng/µl was incubated with substrate S2 at the saturating concentration 1 µM and ATP at the Km concentration 4 µM with or without the compounds in 5 mM $MgCl_2$, 1 mM DTT and one-fold of KinEASE enzymatic buffer at a total volume of 10 µl. The enzymatic reaction was initiated by adding kinase, incubated at room temperature for 40 min, and terminated by adding 10 µl EDTA-containing detection reagents, which were prepared in accordance with the kit instructions. The kinase activity was in a linear range with the protein quantity and incubation time. An HTRF signal was collected through reading the plate using an Infinite® F500 microplate reader (Tecan, Switzerland). Each measurement was repeated at least twice.

The test compounds were prepared in DMSO at stock solutions of 20 mM. To screen the primary compounds, the compounds were diluted using kinase reaction buffer and tested at 100, 10 and 1.0 µM in a 10 µl kinase reaction with DMSO at or below 0.5%.

For the IC_{50} studies, the compounds were serial diluted five-fold using kinase reaction buffer to yield 1000 µM to 0.1 nM in a 10 µl kinase reaction. The kinase activity was assayed as described above. The IC_{50} was determined by fitting the data to a sigmoidal dose-response curve.

ADME/T prediction

ADME/T profile of the screened compounds was assessed using the standard descriptors protocol in the DS 3.0.[11] In the ADME/T module, various parameters (aqueous solubility, blood–brain barrier penetration, CYP2D6 inhibition, hepatotoxicity, human intestinal absorption, and plasma protein binding) were used to quantitatively predict the properties of a set of rules that specify the ADME/T characteristics of the compounds. These selected ADME/T (Absorption, Distribution, Metabolism, Excretion, and Toxicity) predictive properties would play major roles in evaluating the pharmacokinetics and toxicity of possible inhibitors *in vivo*.

Acknowledgements

The authors gratefully acknowledge the National Natural Science Foundation of China (Grant No. 81102379, 81230077, and J1210029), the National High Technology Research and Development Program of China (Grant 2007AA02Z305) for the financial support of this work.

Notes and references

1 A. Abo, J. Qu, M. S. Cammarano, C. Dan, A. Fritsch, V. Baud, B. Belisle and A. Minden, *EMBO J.*, 1998, **17**, 6527-6540.

2 (*a*) C. Dan, A. Kelly, O. Bernard and A. Minden, *J. Biol. Chem.*, 2001, **276**, 32115-32121; (*b*) X. Li and A. Minden, *J. Biol. Chem.*, 2005, **280**, 41192-41200; (*c*) S. Baldassa, A. M. Calogero, G. Colombo, R. Zippel and N. Gnesutta, *J. Cell. Physiol.*, 2010, **224**, 722-733; (*d*) T. Nekrasova and A. Minden, *J. Cell Biochem.*, 2011, **112**, 1795-1806; (*e*) Y. Tian, L. Lei and A. Minden, *Dev. Biol.*, 2011, **353**, 206-216; (*f*) Y. Li, Y. G. Shao, Y. X. Tong, T. Shen, J. Zhang, Y. S. Li, H. Gu and F. Li, *Bba-Mol. Cell. Res.*, 2012, **1823**, 465-475; (*g*) G. Bompard, G. Rabeharivelo, J. Cau, A. Abrieu, C. Delsert and N. Morin, *Oncogene*, 2013, **32**, 910-919.

3 (*a*) D. W. Parsons, T. L. Wang, Y. Samuels, A. Bardelli, J. M. Cummins, L. DeLong, N. Silliman, J. Ptak, S. Szabo, J. K. V. Willson, S. Markowitz, K. Kinzler, B. Vogelstein, C. Lengauer and V. E. Velculescu, *Nature*, 2005, **436**, 792-792; (*b*) S. Chen, T. Auletta, O. Dovirak, C. Hutter, K. Kuntz, S. El-ftesi, J. Kendall, H. Han, D. D. V. Hoff, R. Ashfaq, A. Maitra, C. A. Iacobuzio-Donahue, R. H. Hruban and R. Lucito, *Cancer Biol. Ther.*, 2008, **7**, 1793-1802; (*c*) X. D. Li, Q. Ke, Y. S. Li, F. N. Liu, G. Zhu and F. Li, *Int. J. Biochem. Cell Biol.*, 2010, **42**, 70-79; (*d*) Y. Liu, N. Chen, X. Cui, X. Zheng, L. Deng, S. Price, V. Karantza and A. Minden, *Oncogene*, 2010, **29**, 5883-5894; (*e*) C. M. Wells, A. D. Whale, M. Parsons, J. R. Masters and G. E. Jones, *J. Cell Sci.*, 2010, **123**, 1663-1673; (*f*) H. K. Ahn, J. Jang, J. Lee, P. S. Hoon, J. O. Park, Y. S. Park, H. Y. Lim, K. M.

RSC Advances Accepted Manuscript

Kim and W. K. Kang, *Transl. Oncol.*, 2011, **4**, 345-349; (*g*) H. J. Zhang, M. K. Siu, M. C. Yeung, L. L. Jiang, V. C. Mak, H. Y. Ngan, O. G. Wong, H. Q. Zhang and A. N. Cheung, *Carcinogenesis*, 2011, **32**, 765-771; (*h*) A. Minden, *ISRN Oncol.*, 2012, **2012**, 694201; (*i*) M. H. Park, H. S. Lee, C. S. Lee, S. T. You, D. J. Kim, B. H. Park, M. J. Kang, W. D. Heo, E. Y. Shin, M. A. Schwartz and E. G. Kim, *Oncogene*, 2013, **32**, 2475-2482; (*j*) A. D. Whale, A. Dart, M. Holt, G. E. Jones and C. M. Wells, *Oncogene*, 2013, **32**, 2114-2120.

4 (*a*) S. G. Eckhardt, *Ann. Oncol.*, 2010, **21**, 13-13; *(b)* L. Sun, X. Qu and P. Yin, *Drug Future*, 2012, **37**, 203-208; (*c*) A. E. Dart and C. M. Wells, *Eur. J. Cell Biol.*, 2013, **92**, 129-138.

5 J. Eswaran, W. H. Lee, J. E. Debreczeni, P. Filippakopoulos, A. Turnbull, O. Fedorov, S. W. Deacon, J. R. Peterson and S. Knapp, *Structure*, 2007, **15**, 201-213.

6 (*a*) B. W. Murray, C. X. Guo, J. Piraino, J. K. Westwick, C. Zhang, J. Lamerdin, E. Dagostino, D. Knighton, C. M. Loi, M. Zager, E. Kraynov, I. Popoff, J. G. Christensen, R. Martinez, S. E. Kephart, J. Marakovits, S. Karlicek, S. Bergqvist and T. Smeal, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 9446-9451; (*b*) C. X. Guo, I. McAlpine, J. H. Zhang, D. D. Knighton, S. Kephart, M. C. Johnson, H. T. Li, D. Bouzida, A. L. Yang, L. M. Dong, J. Marakovits, J. Tikhe, P. Richardson, L. C. Guo, R. Kania, M. P. Edwards, E. Kraynov, J. Christensen, J. Piraino, J. Lee, E. Dagostino, C. Del-Carmen, Y. L. Deng, T. Smeal and B. W. Murray, *J. Med. Chem.*, 2012, **55**, 4728-4739.

7 (*a*) B. M. Dolan, S. G. Duron, D. A. Campbell, B. Vollrath, B. S. S. Rao, H. Y. Ko, G. G. Lin, A. Govindarajan, S. Y. Choi and S. Tonegawa, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 5671-5676; (*b*) B. J. Ryu, S. Kim, B. Min, K. Y. Kim, J. S. Lee, W. J. Park, H. Lee, S. H. Kim and S. Y. Park, *Cancer Lett.*, 2014, **349**, 45-50.

8 J. Zhang, J. Wang, Q. Q. Guo, Y. Wang, Y. Zhou, H. Z. Peng, M. S. Cheng, D. M. Zhao and F. Li, *Cancer Lett.*, 2012, **317**, 24-32.

9 (*a*) K. H. Bleicher, H. J. Bohm, K. Muller and A. I. Alanine, *Nat. Rev. Drug Discov.*, 2003, **2**, 369-378; (*b*) W. L. Jorgensen, *Science*, 2004, **303**, 1813-1818.

10 R. J. Li, J. Wang, Z. Xu, W. X. Huang, J. Li, S. F. Jin, D. M. Zhao and M. S. Cheng, *Chemmedchem*, 2014, **9**, 1012-1022.

11 N. Y. Accelrys Inc, *Discovery Studio Tutorials, version 3.5*, 2011.

12 (*a*) N. Y. Schrödinger LLC, *Glide, version 5.8*, 2012; (*b*) R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin and D. T. Mainz, *J. Med. Chem.*, 2006, **49**, 6177−619; (*c*) R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, D. E. Shaw, M. Shelley, J. K. Perry, P. Francis and P. S. Shenkin, *J. Med. Chem.,* 2004, **47**, 1739−1749; (*d*) T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard and J. L. Banks, *J. Med. Chem.,* 2004, **47**, 1750−1759.

13 (*a*) R. Wang, L. Lai and S. Wang, *J. Comput. Aided Mol. Des.*, 2002, **16**, 11-26; (*b*) C. Obiol-Pardo and J. Rubio-Martinez, *J. Chem. Inf. Model.*, 2007, **47**, 134-142.

14 (*a*) J. J. Irwin, T. Sterling, M. M. Mysinger, E. S. Bolstad and R. G. Coleman, *J. Chem. Inf. Model.*, 2012, **52**, 1757-1768; (*b*) J. J. Irwin and B. K. Shoichet, *J. Chem. Inf. Model.*, 2005, **45**, 177-182.

15 (*a*) C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Deliver. Rev.*, 1997, **23**, 3; (*b*) C. A. Lipinski, *J. Pharmacol. Toxicol. Methods*, 2000, **44**, 253.

16 (*a*) D. Shivakumar, J. Williams, Y. Wu, W. Damm, J. Shelley and W. Sherman, *J. Chem. Theory Comput.*, 2010, **6**, 1509−1519; (*b*) W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225−11236; (*c*) W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1988, **110**, 1657−1666.

RSC Advances Accepted Manuscript

 This journal is © The Royal Society of Chemistry 2012