Molecular BioSystems

Cite this: *Mol. BioSyst.*, 2011, **7**, 2702–2710

www.rsc.org/molecularbiosystems

PAPER

In silico pharmacology suggests ginger extracts may reduce stroke risks†

Tung-Ti Chang,[ab] Kuan-Chung Chen,[a] Kai-Wei Chang,[a] Hsin-Yi Chen,[c] Fuu-Jen Tsai,[cd] Mao-Feng Sun[ab] and Calvin Yu-Chian Chen*[acef]

Received 8th June 2011, Accepted 29th June 2011
DOI: 10.1039/c1mb05228d

Aberrations in cyclic adenosine monophosphate (cAMP) signaling cascade has been linked to the allergic responses that associate with the risks of stroke or cardiovascular diseases. Phosphodiesterase 4D (PDE4D) has been shown to be highly involved in cAMP regulation and is hence implied to be a potential drug target in stroke prevention. To identify potential PDE4D inhibitors from traditional Chinese medicine (TCM), we employed machine learning modeling techniques to screen a comprehensive TCM database. The multiple linear regression (MLR) and support vector machine (SVM) models constructed have correlation coefficients of 0.8234 and 0.7854 respectively. Three candidates from the ginger family were identified based on the prediction models. Molecular dynamics simulation further validated the binding stabilities of each candidate in comparison to the control inhibitor L-454560. The intermolecular distances suggested that the candidates could hinder PDE4D from binding to cAMP. Furthermore, the HypoGen validation suggested that top2, top3, and the control L-454560 mapped with the predicted pharmacophores. The results suggested that the 3 compounds identified from the ginger family were capable in inhibiting cAMP binding and hydrolysis by PDE4D. We further identified and characterized the ligand binding properties that are associated with the inhibition of PDE4D.

Introduction

Stroke is one of the leading causes of death worldwide. However, available medications remain extremely limited. For example, the commercialized stroke drug, Actilyse®, is only effective within 3 hours of the stroke onset,[1,2] which makes the treatment less applicable in many cases. Other common treatments, including administering aspirin and injecting tissue plasmogen activators, are blood thinning agents that have been reported causing internal bleeding as a side effect. Hence, developing additional effective treatments with minimal side effects is of crucial importance.

By looking into the signaling cascades of the cyclic adenosine monophosphate (cAMP) pathway, deregulation of such second messenger has been linked with many diseases or disorders, including inflammation,[3] cancer,[4] and mental disorders.[5] Recent genome-wide association study (GWAS) on stroke-inducing factors has identified phosphodiesterase 4D (PDE4D), which has a key role in regulating cAMP concentration.[6] GWAS replications on different ethnic groups also suggested relationships between PDE4D and stroke.[7] Worrall and Mychaleckyj's study indicated the indirect effect of PDE4D on stroke or cardiovascular biomarkers. Moreover, PDE4D is responsible for at least 80% of PDE enzyme activity, which predominate cAMP metabolism in inflammatory and injury responses that could have an implicit effect on stroke risks.[8] For this reason, the effects of PDE4D enzymatic activities support the relationship between PDE4D protein and stroke. Hence, the development of novel PDE4D inhibitors could lower the stroke risk that was induced by abnormally upregulated PDE4D.

We focused our attention in discovering PDE4D inhibitors from Traditional Chinese medicine (TCM), which is a popular East Asia medical study that has been practiced over thousands of years. Scientific investigations on this class of medicine have further identified multiple therapeutic values, from inflammation, neurological diseases, to cancer.[9,10] TCM has also been demonstrated to be a resourceful databank for *in silico* drug discovery. Many potential drugs or precursor

[a] *Laboratory of Computational and Systems Biology, School of Chinese Medicine, China Medical University, Taichung, 40402, Taiwan*
[b] *Department of Acupuncture, China Medical University Hospital, Taichung, Taiwan*
[c] *Department of Bioinformatics, Asia University, Taichung, 41354, Taiwan*
[d] *Department of Medical Genetics, China Medical University, Taichung, 40402, Taiwan*
[e] *Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA*
[f] *Computational and Systems Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: ycc@mail.cmu.edu.tw; ycc929@MIT.EDU; Tel: +1-617-353-7123*
† Electronic supplementary information (ESI) available. See DOI: 10.1039/c1mb05228d

compounds have, therefore, been identified.[11–17] To determine the potential PDE4D inhibitors, we employed the most comprehensive 3D small molecule TCM database, TCM Database@Taiwan,[18] for screening. In the past, we had established the PDE4D researches about screening potential candidates based on docking for stroke[17,19] and Alzheimer's disease.[20] In those studies, difficulties in matching the descriptors of top candidate descriptors for prediction models with training set descriptors were experienced. In this study, we screened potential candidates based on the prediction models. Two quantitative structure–activity relationship (QSAR) models, support vector machine (SVM) and multiple linear regressions (MLR) models, were applied to screen the TCM database and predict their activities. The docking simulation was performed to identify docking poses of the common candidates from QSAR models. Molecular dynamics (MD) simulations followed by HypoGen modeling were used to validate the screening results.

Materials and methods

Fig. 1 summarized the general *in silico* drug identification protocol used in this study.

Data collection

The crystal structure of PDE4D was obtained from RCSB Protein Data Bank (PDB ID: 3G4G).[21] For establishing QSAR prediction models, 62 compounds from three different studies were obtained.[22–24] More than 20 000 TCM 3D small molecules were obtained from TCM Database@Taiwan.[18] Ionization states of the ionizing functional groups were adjusted using Accelrys Discovery Studio 2.5 (DS 2.5).[25]

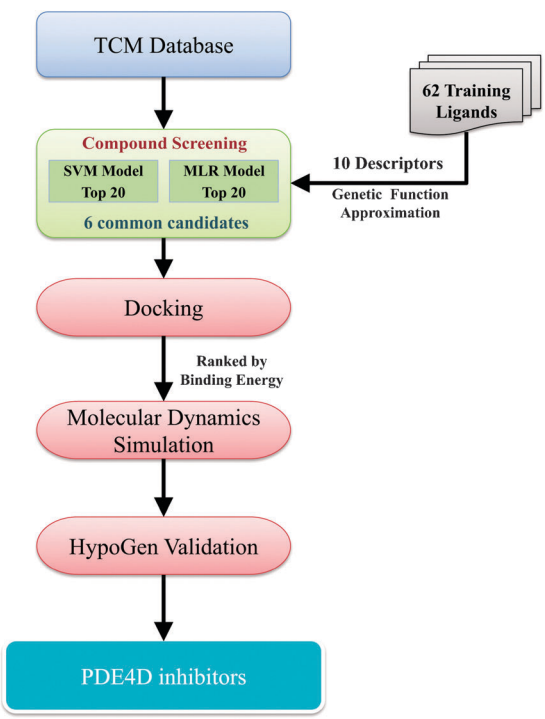

Fig. 1 Schematic flowchart of the screening protocols.

QSAR modeling

The non-linear SVM and the linear MLR machine learning algorithms were employed to construct QSAR models. The SVM model was established using LibSVM.[26] The MLR model was built using MATLAB.[27] The regression model of SMV (SVR)[28] was established using gridregression.py program in the libsvm-2.91 package.[26]

The prediction modeling protocol built from the last study[17] was applied. 62 compound structures and activity properties gathered from Card *et al.*,[22] Aspiotis *et al.*,[23] and Naganuma *et al.*[24] were obtained and randomly selected for training (51 compounds) and testing (11 compounds) the QSAR models. A genetic function approximation (GFA)[29] module from DS2.5 was employed to identify representative molecular descriptors. The square correlation coefficient (R^2) of pIC_{50} values was assessed to determine the accuracy of each model. The descriptors identified using GFA were Dipole mag, Dipole Y, Jurs FNSA, Jurs PPSA3, Jurs RASA, Jurs RNCG,

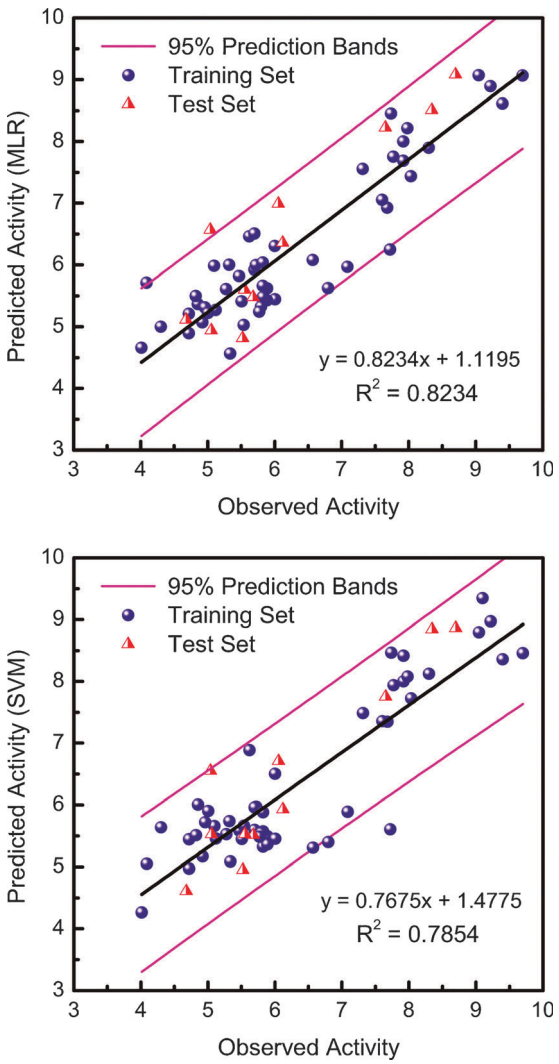

Fig. 2 Comparative plots of MLR (top) and SVM (bottom) models. Correlation trend (black line) and 95% confidence regions (enclosed by magenta lines) are shown. Training set (blue dots) and testing set (red triangle) data are presented.

Jurs TASA, Jurs WPSA3, shadow X length, and shadown Y length. Five-fold cross validation was performed for both MLR and SVM models.

Table 1 Docking results and predicted pIC_{50} for top TCM candidates and L-454560

Name	Predicted pIC_{50}		Binding energy
	MLR	SVM	
(5*S*)-7-(4-Hydroxy-phenyl)-5-methoxy-1-phenylheptan-3-one	7.56	6.62	−140.02
(4*E*)-7-(4-Hydroxyphenyl)-1-phenylhept-4-en-3-one	7.72	6.34	−133.80
Demethoxy-[6]-shogaol	7.52	6.34	−128.50
L-454560[a]	8.42	9.04	−127.55
Fipronil	8.28	6.26	−101.78
Dracorhodin perchlorate	7.83	6.59	−80.98
α-Asarone	9.14	6.35	−54.24

[a] Control.

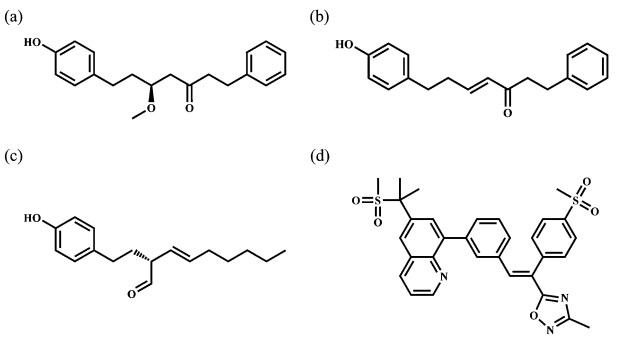

Fig. 3 The molecular structures of (a) top1, (b) top2, (c) top3, and the control (d) L-454560.

Docking

CHARMm force field[30] was applied to the receptor as well as each ligand in the TCM database. The DS2.5 Flexible Docking module[31] was employed for virtual docking, which screens for the ligand compatibilities to the cAMP binding site defined by the co-crystallized compound on the 3G4G structure. The PDE4 inhibitor, L-454560,[32] was used as control. Binding Energy (BE) of each protein–ligand binding pose was evaluated using the Calculate Binding Energies module[33] of DS 2.5.

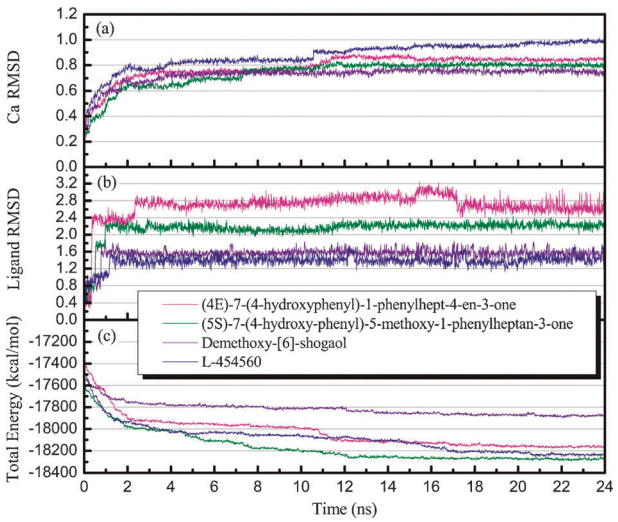

Fig. 5 MD dynamics trajectories of (a) Cα RMSD, (b) ligand RMSD, and (c) total energy of PDE4D in complex with each candidate and L-454560.

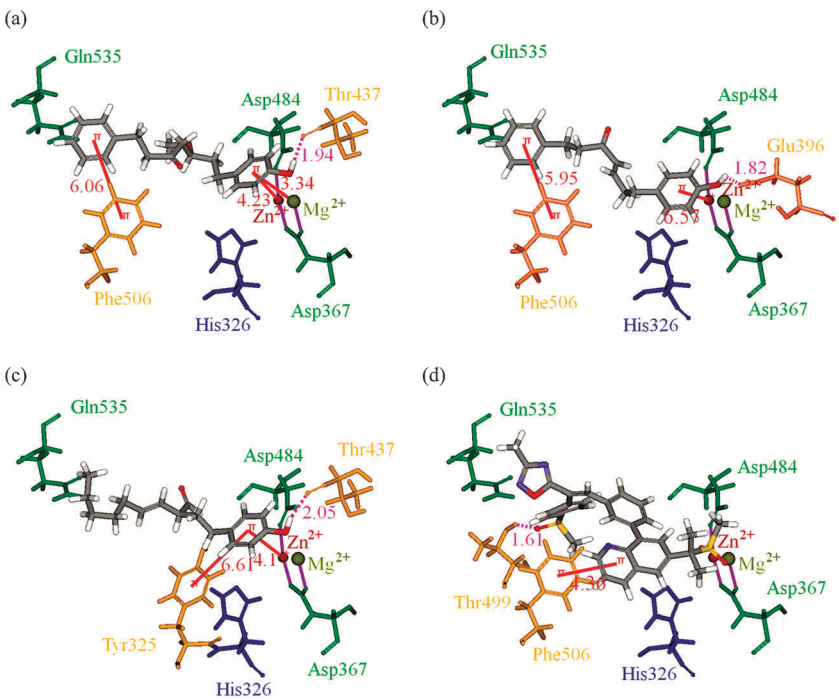

Fig. 4 Docking poses of PDE4D with (a) top1, (b) top2, (c) top3, and (d) L-454560 in the cAMP binding site. Residues of the binding site (green), active site (blue), and other residues (brown) are shown. Electrostatic interactions (magenta line), pi–pi interactions (red line), hydrogen bond interactions (pink dashed line) are presented.

Molecular dynamics simulation

DS2.5 Simulation package[25] was employed for MD simulation CHARMm force field.[30] A 7 Å solvation shell of TIP3P water was created and balanced with sodium cations for the simulation system. For minimization, 6000 cycles of Steepest Descent[34] with restricted protein followed by 6000 cycles of Conjugate Gradient[35] with flexible protein were performed. Time steps were set to 0.002 ps throughout the simulation. All bonds involving hydrogen atoms were fixed with the SHAKE algorithm. The long-range electrostatics were treated by the PME method. Each simulation system was heated from 51 K to 310 K in a 50 ps interval, then followed by a 200 ps equilibration phase. For the production step, constant temperature ensemble (NVT) with 0.4 ps of temperature coupling decay time. MD simulation was performed for 24 ns.

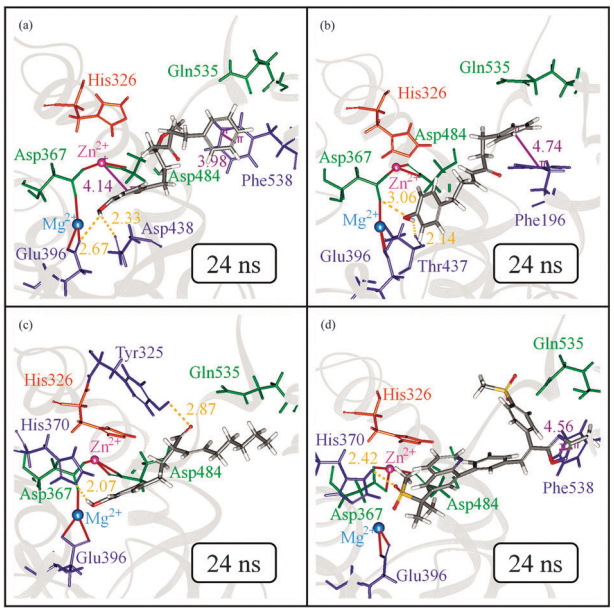

Fig. 6 The docking poses for PDE4D with (a) top1, (b) top2, (c) top3, and (d) L-454560 at 24 ns of MD. Residues of the binding site (green), active site (orange), and other residues (blue) are shown. Electrostatic interactions (red line), pi–pi interactions (magenta line), hydrogen bond (yellow dashed line) are presented.

Pharmacophore modeling

Training set data used to train the machine learning QSAR models were modified to low-energy conformations using a FAST generation module in DS2.5. Common ligand features were identified by the HipHop algorithm.[36] The HypoGen algorithm[37] then construct the 3D-QSAR models based on the common pharmacophore features identified.

Results and discussions

QSAR modeling

The R^2 of MLR and SVM prediction models were 0.78 and 0.89 respectively (Fig. 2). The MLR prediction model with optimized property descriptors was described as the following:

$$\begin{aligned} predicted\_IC_{50} = {} & -23.43 + 0.1383 \times Dipole\_mag + 0.0921 \\ & \times Dipole\_Y - 31.708 \times Jurs\_FNSA3 \\ & - 0.2866 \times Jurs\_PPSA3 + 28.543 \\ & \times Jurs\_RASA + 20.607 \times Jurs\_RNCG \\ & - 0.0330 \times Jurs\_TASA + 0.691 \\ & \times Jurs\_WPSA3 + 0.428 \times Shadow\_Xlength \\ & + 0.528 \times Shadow\_Ylength \end{aligned}$$

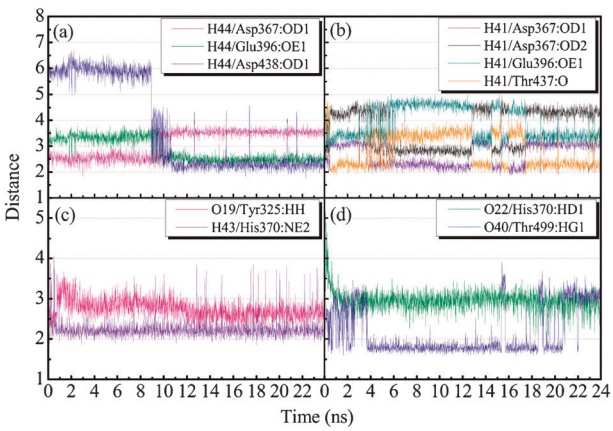

Fig. 7 Distance in Å of hydrogen bonds between PDE4D and (a) top1, (b) top2, (c) top3, and (d) L-454560.

Table 2 H-bonds of PDE4D–ligand complexes for the top 3 compounds and L-454560

Ligand	H-bond	Ligand atom	Amino acid	Max. distance	Average distance	Min. distance	H-bond occupancy[a] (%)
Top 1	1	H44	Asp367:OD1	3.88	3.12	1.83	18.54
	2	H44	Glu396:OE1	4.12	2.83	2.08	34.42
	3	H44	Asp438:OD1	6.69	3.70	1.77	52.79
Top 2	1	H41	Asp367:OD1	3.46	2.69	1.85	43.67
	2	H41	Asp367:OD2	4.96	3.66	2.47	0.08
	3	H41	Glu396:OE1	5.02	3.89	2.50	0.04
	4	H41	Thr437:O	4.75	2.82	1.87	51.67
Top 3	1	O19	Tyr325:HH	3.66	2.75	2.07	13.33
	2	H43	His370:NE2	4.05	2.23	1.85	95.96
L-454560	1	O22	His370:HD1	5.16	2.97	2.02	3.00
	2	O40	Thr499:HG1	3.91	2.15	1.58	72.13

[a] H-bond occupancy cutoff: 2.5 Å.Top 1: (5*S*)-7-(4-hydroxy-phenyl)-5-methoxy-1-phenylheptan-3-one.Top 2: (4*E*)-7-(4-hydroxyphenyl)-1-phenylhept-4-en-3-one.Top 3: Demethoxy-[6]-shogaol.

Most residuals between the prediction and observation qwere within a range of 1.5 suggesting that the prediction models were reliable (Table S1, ESI†, for predicted pIC_{50} of MLR and SVM models). The accuracies of MLR and SVM

Table 3 Residues close to each candidate and control (cutoff: 3.5 Å)

Residue	Val193	Phe196	Ile197	Thr200	Phe201	Tyr325	His326	His330	Asp367	His370	Gly372	Val373
Top 1		+			+	+	+	+	+			
Top 2		+		+	+	+	+		+			
Top 3		+	+	+		+	+		+			
Control	+	+	+	+	+	+	+		+	+	+	+
Residue	Ser374	Asn375	Leu395	Glu396	Thr437	Asp438	Met439	Asp484	Leu485	Asn487	Pro488	Tyr495
Top 1			+	+	+	+	+	+	+	+		
Top 2	+			+	+	+	+		+			
Top 3			+	+	+	+	+	+	+		+	+
Control	+	+		+	+	+	+	+	+	+		+
Residue	Trp498	Thr499	Ile502	Met503	Phe506	Met523	Ser534	Gln535	Gly537	Phe538	Ile542	**Total**
Top 1			+	+	+			+		+	+	20
Top 2			+		+			+		+		16
Top 3			+	+	+	+	+	+		+		22
Control	+	+	+	+	+		+	+	+	+		30

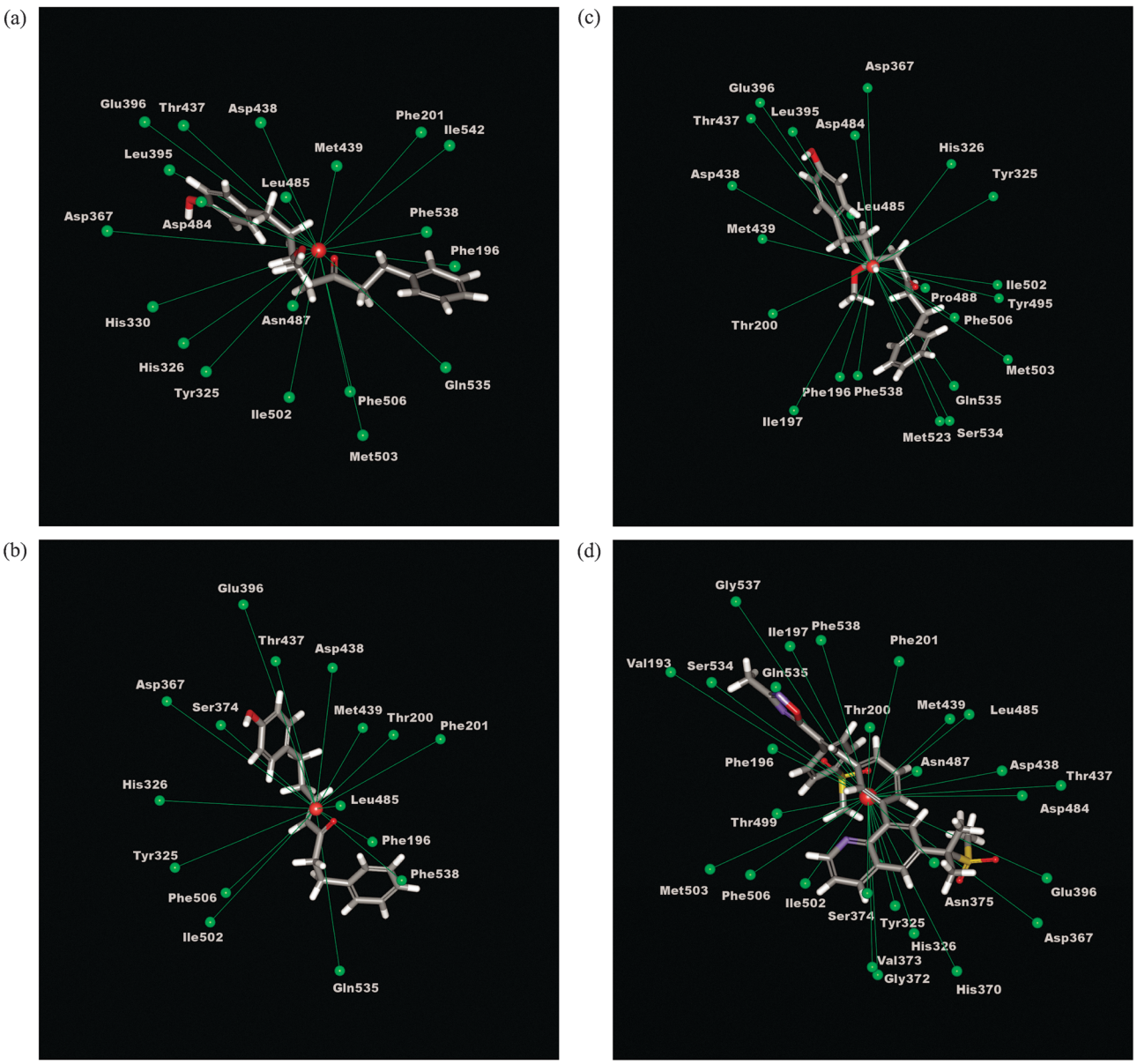

Fig. 8 Relative residue positions compared to (a) top1, (b) top2, (c) top3, and (d) L-454560 based on the center of mass.

models were also validated by predicting pIC_{50} of L-454560, which were 8.42 and 9.04 with respect to the actual value of 8.92 (IC_{50} = 1.2 nM). Both models were employed to screen the TCM database for potential PDE4D inhibitors. The top candidates that are common in both models indicate the potential inhibitor according to these 10 training set descriptors. For the binding affinity, the ligand docking poses were assessed using the docking algorithm and ranked using BE. Six top candidates are shown in Table 1.

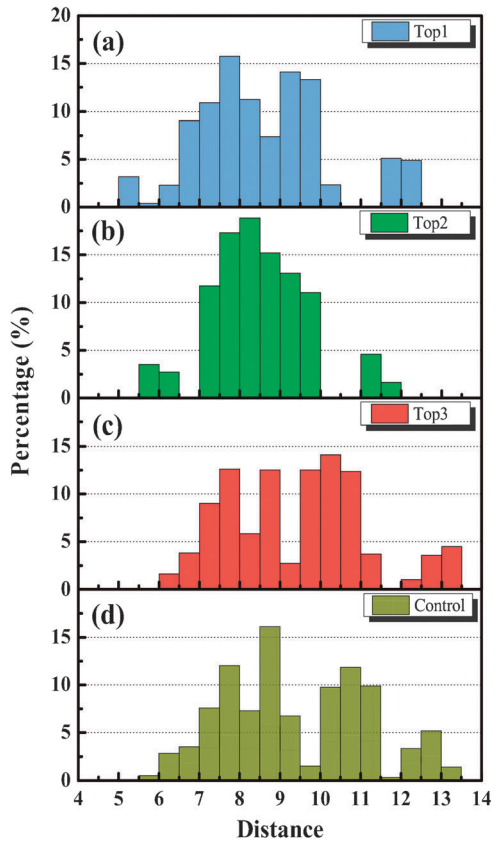

Fig. 9 Distance distribution (in Å) of all residues near to the candidates and the control L-454560.

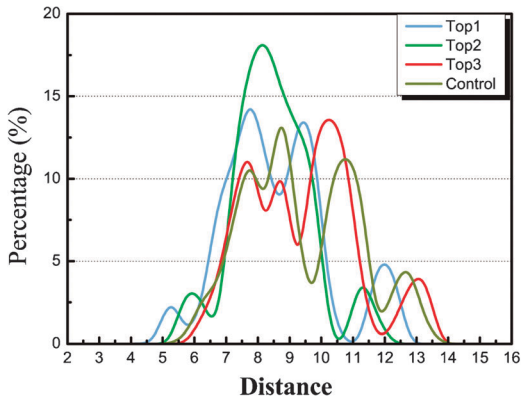

Fig. 10 Relative distance distribution that compares all candidate ligands and the control L-454560.

Docking

The docking results identified 3 candidates, which have higher BE than L-454560 (Table 1). These candidates were (4*E*)-7-(4-hydroxyphenyl)-1-phenylhept-4-en-3-one (top1), (5*S*)-7-(4-hydroxy-phenyl)-5-methoxy-1-phenylheptan-3-one (top2), and demethoxy-[6]-shogaol (top3). The top1 and top2 candidates were isolated from *Alpinia officinarum*. The top3 candidate was isolated from *Zingiber officinale*. Interestingly, all top three compounds came from the ginger family. According to the Compendium of Materia Medica, ginger is documented to promote circulatory functions as well as anti-inflammation and anti-plaque properties.[38] These effects may correlate to the anticipated effects from administering drug-form PDE4D inhibitors for stroke prevention. Hence, the TCM candidates originated from ginger may be some of the active compounds contributing to the observed healing functions of ginger since ancient times.

The structure of each candidate has a phenol group on one end and a relatively less polar group on the other end (Fig. 3). From the docking poses shown in Fig. 4, all phenyl groups have pi–cation interaction with Zn^{2+}. In addition, the hydroxyl group on the phenyl structure formed hydrogen bonds (H-bonds) with nearby residues, Glu396 or Thr437. Comparatively, the control L-454560 formed an ionic interaction with Zn^{2+}. A stabilizing pi–pi interaction is also observed in each binding complex, with either Phe506 on top1, top2, and control, or Tyr325 on top3. Additionally, the phenyl group on the top1 candidate also has pi–cation interaction with the Mg^{2+}. These observations suggested high binding affinities in all candidates and the control.

Molecular dynamics simulation

The stability of each complex was assessed through MD simulation. Root mean square deviation (RMSD) trajectories of each ligand and the corresponding protein (C_α) were calculated individually for each protein–ligand interaction (Fig. 5). Total energy for each binding pose was also calculated. All candidates were stabilized after 18 ns of simulation. L-454560, on the other hand, stabilized after 21 ns. Intriguingly, all ligands, including the control, experienced sudden conformational changes within first 2 ns, implying possible alliteration of the similar binding features. The snapshot of each docking pose at the 24 ns of simulation was presented in Fig. 6. The frequency of H-bond occurrences during MD further suggested the presence of stable H-bonds that held the ligands in close approximation to the key residues Asp367 and Asp484 (Table 2). Additionally, the top1 candidate maintained the pi–cation interactions with the divalent metal cation, Zn^{2+}, which further stabilized the ligand within the docking site (Fig. 6(a)). Intriguingly, the end group of each ligand near the key residue Gln535 formed new interaction patterns in place of the bindings before the simulation. The pi–pi interactions with Phe538 (top1 and control) and Phe196 (top2) as well as the H-bond with Tyr325 (top3) were observed (Fig. 6).

According to the distance trajectory of the H-bond during MD shown in Fig. 7, the top1 candidate formed H-bonds with Asp367, Glu396, and Asp438 after 10 ns; the top2 candidate

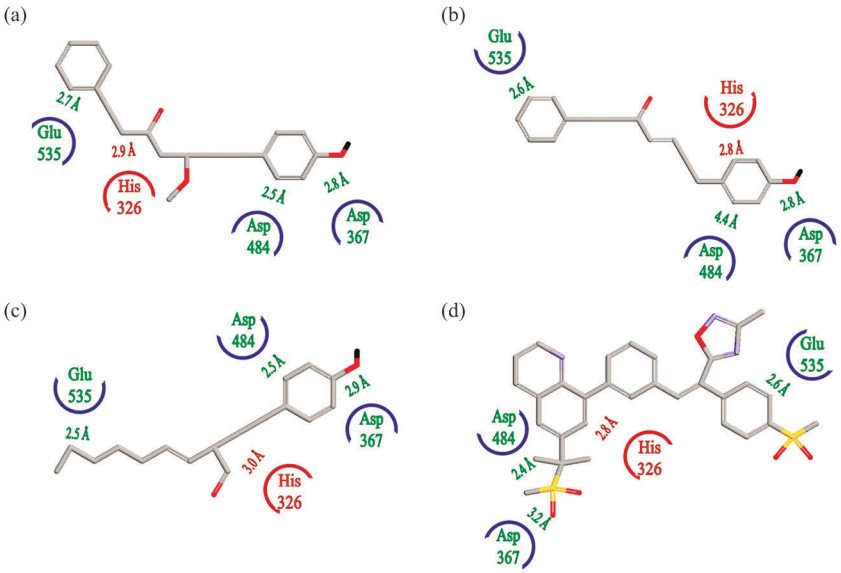

Fig. 11 Approximated distance between key residues and (a) top1, (b) top2, (c) top3, and (d) L-454560 in the stabilized binding conformation. The relative positions of the binding residues (blue) and the active residues (red) are presented.

formed more H-bonds with Asp367, Glu396, and Thr437 when the protein–ligand complex stabilized after 18 ns of simulation. On the other hand, the top3 ligand and L-454560 have H-bonds with His370 (Fig. 7). According to the H-bond trajectory, the control ligand formed a longer H-bond (>3 Å) with Thr499, which also contributes to the binding stabilities. We further determined the positioning of each ligand within the binding site. All nearby residues at a range of 3.5 Å to a ligand were monitored (Table 3). Center of mass was identified for monitoring the distances between ligands and nearby residues during 20–24 ns simulation (Fig. 8). For top1, most distances lay at 5.25 Å, 7.75 Å, 9.25 Å, and 11.75 Å (Fig. 9(a)). For top2, three peaks were observed at 5.75 Å, 8.25 Å, and 11.25 Å, for which the distribution was more concentrated (Fig. 9(b)). The top3 ligand was slightly farther from the nearby residues, where most distances were distributed at 7.75 Å, 8.75 Å, 10.25 Å, and 13.25 Å (Fig. 9(c)). Similar distance distribution peaks were observed for L-454560, at 7.75 Å, 8.75 Å, 10.75 Å, and 12.75 Å (Fig. 9(d)).

As summarized in Fig. 10, the distances between the ligand and nearby residues were in the range of 5 Å to 14 Å, where several distances existed more frequently than others. All peaks presented indicated that most distances were below 10 Å. Short intermolecular distances implied that all candidates had strong protein–ligand interactions that prohibit the natural substrate cAMP from binding. Furthermore, top1 had the shortest intermolecular distance to the binding site whereas top3 was slightly distant from the receptor (Fig. 10). As shown in Fig. 11, all ligands had relatively stable poses within the binding site that is in close approximation to the key binding residues Asp367, Asp484, and Gln535, as well as the activity residue His326 though no direct bonding was observed. This suggested all candidates were "locked" within the binding site and prevent the PDE4D from interacting with its natural substrate cAMP.

Pharmacophore modeling

We further built a HypoGen model for validating the candidates. Variations in modeling features (Table 4) and CatScramble plot (Fig. 12) were assessed to identify a validation model that best represents the pharmacophores of a potential ligand. The most fitting HypoGen model, which had an *R* value of 0.870, comprised with 2 H-bond receptor components and 3 hydrophobic regions that most of which were 3 Å apart from each other (Table 4, Fig. 13(a)). By mapping each TCM candidate to the HypoGen model, the top2, top3, and control compounds showed satisfactory features that matched with the given pharmacophore profiles (Fig. 13(b)–(d)). The ring structures and the hydrophobic chains on the ligands mapped well with the hydrophobic regions defined by the model. In addition, carbonyl groups on top2 and top3, as well as the oxygen atoms on the control mapped to the H-bond acceptors. On the other hand, the top1 compound did not fit with the pharmacophore arrangements due to an additional ether group (R–O–R′) within the hydrophobic regions.

Table 4 The results of pharmacophore hypothesis generation

Hypothesis	Total cost	Cost difference	Error cost	RMS	*R* value	Feature
1	295.85	200.58	278.63	1.504	0.870	AAHHH
2	308.53	187.90	291.32	1.634	0.844	AAHHH
3	309.95	186.48	292.74	1.648	0.841	AHHHH
4	310.36	186.07	293.03	1.651	0.841	AAHHH
5	311.06	185.37	293.86	1.659	0.839	AHHHH
6	312.24	184.19	295.10	1.671	0.836	AHHHH
7	312.90	183.53	295.71	1.677	0.835	AAHHH
8	313.16	183.27	295.97	1.679	0.834	AHHHH
9	313.27	183.16	295.83	1.678	0.835	AAHHH
10	315.64	180.79	298.32	1.702	0.830	AHHHH

Null cost: 496.43. Fixed cost: 225.66. Configuration cost: 15.99. Hydrogen bond acceptors (A) and hydrophobic features (H) were identified and analyzed for modeling.

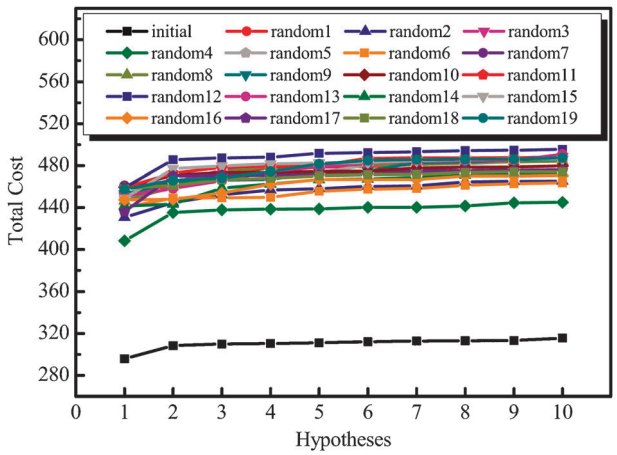

Fig. 12 CatScramble validation plot that compares total costs between initial and randomized HypoGen hypothesis.

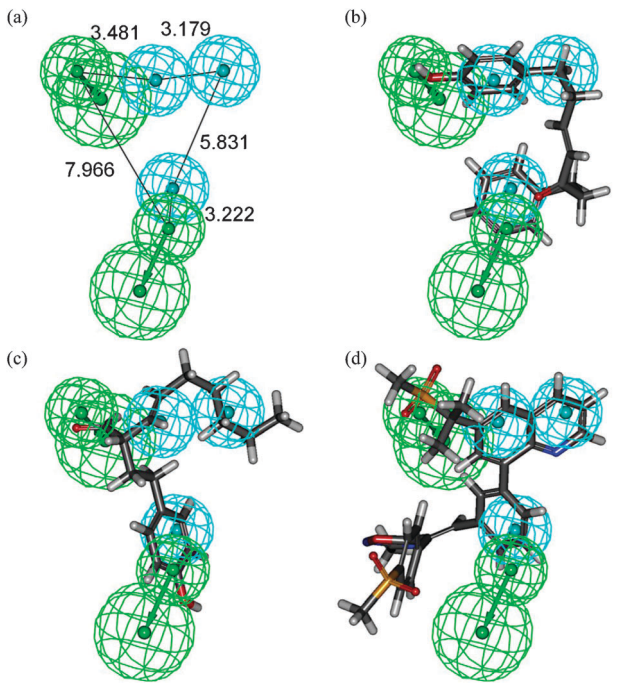

Fig. 13 HypoGen results of (a) pharmacophores map with labeled distances, (b) mapping with top2, (c) mapping with top3, and (d) mapping with L-454560. Hydrophobic region (blue sphere), hydrogen bond acceptor region (green sphere), and hydrogen bond partner direction (green arrow) are presented.

Pharmacophore results complement the findings in MD simulation. As shown in Fig. 13, the phenol group in control, top 2, and top 3 compounds formed H-bonds with Asp367, matching the H-bond acceptor feature in the pharmacophore model. Top 3 also formed an additional H-bond with Tyr325, matching the second H-bond acceptor feature in pharmacophore modeling.

Conclusion

By screening the TCM database with MLR and SVM models followed by docking algorithm, we identified three potential PDE4D inhibitors from the ginger family. These ligands had similar molecular structures. BE evaluations on all 3 candidates further suggested potentially higher binding compatibilities to the cAMP dock site in PDE4D. The MD simulation further identified stable binding conformation of each ligand, though differed from the initial docking conformations. Analysis of the intermolecular bonding suggested interference of cAMP binding and hydrolysis activities of PDE4D. HypoGen validation determined that the top1 candidate may not fit with the desired pharmacophore regions. However, top1 may still be considered as a potent PDE4D inhibitor based on other validations. Hence, we propose that all top1, top2 and top3 compounds from the ginger family could be further analyzed for their therapeutic potentials in lowering stroke risks.

Acknowledgements

The research was supported by grants from the National Science Council of Taiwan (NSC 99-2221-E-039-013), Committee on Chinese Medicine and Pharmacy (CCMP100-RD-030), China Medical University and Asia University (CMU98-TCM, CMU99-TCM, CMU99-S-02, CMU99-ASIA-25, CMU99-ASIA-26 CMU99-ASIA-27 CMU99-ASIA-28). This study is also supported in part by Taiwan Department of Health Clinical Trial and Research Center of Excellence (DOH100-TD-B-111-004) and Taiwan Department of Health Cancer Research Center of Excellence (DOH100-TD-C-111-005). We are grateful to the Asia University cloud-computing facilities.

References

1 Anonymous, *N. Engl. J. Med.*, 1995, **333**, 1581–1587.
2 Actilyse, License and Approval., http://www.actilyse.com/com/Main/introduction/licence_approval/index.htm, accessed June, 24, 2011.
3 S. Salinthone, V. Yadav, R. V. Schillace, D. N. Bourdette and D. W. Carr, *PLoS One*, 2010, **5**, e13058.
4 N. Dumaz, Y. Light and R. Marais, *Mol. Cell. Biol.*, 2002, **22**, 3717–3728.
5 M. Kvajo, H. McKellar and J. A. Gogos, *Curr. Top. Behav. Neurosci.*, 2010, **4**, 629–656.
6 S. Gretarsdottir, G. Thorleifsson, S. T. Reynisdottir, A. Manolescu, S. Jonsdottir, T. Jonsdottir, T. Gudmundsdottir, S. M. Bjarnadottir, O. B. Einarsson, H. M. Gudjonsdottir, M. Hawkins, G. Gudmundsson, H. Gudmundsdottir, H. Andrason, A. S. Gudmundsdottir, M. Sigurdardottir, T. T. Chou, J. Nahmias, S. Goss, S. Sveinbjornsdottir, E. M. Valdimarsson, F. Jakobsson, U. Agnarsson, V. Gudnason, G. Thorgeirsson, J. Fingerle, M. Gurney, D. Gudbjartsson, M. L. Frigge, A. Kong, K. Stefansson and J. R. Gulcher, *Nat. Genet.*, 2003, **35**, 131–138.
7 C. Fong, D. C. Ko, M. Wasnick, M. Radey, S. I. Miller and M. Brittnacher, *Bioinformatics*, 2010, **26**, 560–564.
8 B. B. Worrall and J. C. Mychaleckyj, *Stroke*, 2006, **37**, 1955–1957.
9 Y. M. Leung, Y. H. Tsou, C. S. Kuo, S. Y. Lin, P. Y. Wu, M. J. Hour and Y. H. Kuo, *Phytomedicine*, 2010, **18**, 46–51.
10 M. D. Yang, K. C. Lai, T. Y. Lai, S. C. Hsu, C. L. Kuo, C. S. Yu, M. L. Lin, J. S. Yang, H. M. Kuo, S. H. Wu and J. G. Chung, *Anticancer Res.*, 2010, **30**, 2135–2143.
11 T. T. Chang, H. J. Huang, K. J. Lee, H. W. Yu, H. Y. Chen, F. J. Tsai, M. F. Sun and C. Y. C. Chen, *J. Biomol. Struct. Dyn.*, 2010, **28**, 309–321.
12 H. J. Huang, K. J. Lee, H. W. Yu, H. Y. Chen, F. J. Tsai and C. Y. C. Chen, *J. Biomol. Struct. Dyn.*, 2010, **28**, 187–200.

13 H. J. Huang, K. J. Lee, H. W. Yu, C. Y. Chen, C. H. Hsu, H. Y. Chen, F. J. Tsai and C. Y. C. Chen, *J. Biomol. Struct. Dyn.*, 2010, **28**, 23–37.
14 C. Y. C. Chen, *J. Mol. Graphics Modell.*, 2009, **28**, 261–269.
15 C. Y. Chen, Y. H. Chang, D. T. Bau, H. J. Huang, F. J. Tsai, C. H. Tsai and C. Y. C. Chen, *J. Biomol. Struct. Dyn.*, 2009, **27**, 171–178.
16 K. C. Chen, K. W. Chang, H. Y. Chen and C. Y. C. Chen, *Mol. BioSyst.*, 2011, in press.
17 K. C. Chen and C. Y. C. Chen, *Soft Matter*, 2011, **7**, 4001–4008.
18 C. Y. C. Chen, *PLoS One*, 2011, **6**, e15939.
19 M. F. Sun, K. C. Chen, K. W. Chang, H. Y. Chen, F. J. Tsai, M. Fisher, J. G. Lin and C. Y. C. Chen, *Mini-Rev. Med. Chem.*, 2011, in press.
20 M. F. Sun, T. T. Chang, K. C. Chen, Y. H. Wong, S. C. Yang, K. W. Chang, T. Y. Tsai, H. Y. Chen, F. J. Tsai, J. G. Lin and C. Y. C. Chen, *Mol. Simul.*, 2011, in press.
21 A. B. Burgin, O. T. Magnusson, J. Singh, P. Witte, B. L. Staker, J. M. Bjornsson, M. Thorsteinsdottir, S. Hrafnsdottir, T. Hagen, A. S. Kiselyov, L. J. Stewart and M. E. Gurney, *Nat. Biotechnol.*, 2010, **28**, 63–70.
22 G. L. Card, L. Blasdel, B. P. England, C. Zhang, Y. Suzuki, S. Gillette, D. Fong, P. N. Ibrahim, D. R. Artis, G. Bollag, M. V. Milburn, S. H. Kim, J. Schlessinger and K. Y. Zhang, *Nat. Biotechnol.*, 2005, **23**, 201–207.
23 R. Aspiotis, D. Deschenes, D. Dube, Y. Girard, Z. Huang, F. Laliberte, S. Liu, R. Papp, D. W. Nicholson and R. N. Young, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 5502–5505.
24 K. Naganuma, A. Omura, N. Maekawara, M. Saitoh, N. Ohkawa, T. Kubota, H. Nagumo, T. Kodama, M. Takemura, Y. Ohtsuka, J. Nakamura, R. Tsujita, K. Kawasaki, H. Yokoi and M. Kawanishi, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 3174–3176.
25 *Accelrys Software 2.5 edn*, Accelrys Software Inc., San Diego, 2009.
26 R. E. Fan, P. H. Chen and C. J. Lin, *J. Mach. Learn. Res.*, 2005, **6**, 1889–1918.
27 A. K. Saxena and P. Prathipati, *SAR QSAR Environ. Res.*, 2003, **14**, 433–445.
28 A. J. Smola and B. Scholkopf, *Stat. Comput.*, 2004, **14**, 199–222.
29 D. Rogers and A. J. Hopfinger, *J. Chem. Inf. Comput. Sci.*, 1994, **34**, 854.
30 B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 1983, **4**, 187–217.
31 J. Koska, V. Z. Spassov, A. J. Maynard, L. Yan, N. Austin, P. K. Flook and C. M. Venkatachalam, *J. Chem. Inf. Model.*, 2008, **48**, 1965–1973.
32 D. Macdonald, A. Mastracchio, H. Perrier, D. Dube, M. Gallant, P. Lacombe, D. Deschenes, B. Roy, J. Scheigetz, K. Bateman, C. Li, L. A. Trimble, S. Day, N. Chauret, D. A. Nicoll-Griffith, J. M. Silva, Z. Huang, F. Laliberte, S. Liu, D. Ethier, D. Pon, E. Muise, L. Boulet, C. C. Chan, A. Styhler, S. Charleson, J. Mancini, P. Masson, D. Claveau, D. Nicholson, M. Turner, R. N. Young and Y. Girard, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 5241–5246.
33 J. Tirado-Rives and W. L. Jorgensen, *J. Med. Chem.*, 2006, **49**, 5880–5884.
34 R. Fletcher, Optimization, Academic Press, New York and London, 1969.
35 R. Fletcher and C. M. Reeves, *Comput. J.*, 1964, **7**, 149–154.
36 O. O. Clement and A. T. Mehl, in *Pharmacophore Perception, Development, and Use in Drug Design*, ed. O. F. Güner, International University Line, La Jolla, CA, 2000, ch. 69–84.
37 H. Li, J. Sutter and R. Hoffman, in *Pharmacophore Perception, Development, and Use in Drug Design*, ed. O. F. Güner, International University Line, La Jolla, CA, 2000, pp. 171–189.
38 S. Li, *Compendium of Materia Medica (Bencao Gangmu)*, Foreign Languages Press, Beijing, China, 1596.

