MedChemComm

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: Y. P. Quan, L. P. Cheng, T. C. Wang, W. Pang, F. H. Wu and J. W. Huang, *Med. Chem. Commun.*, 2017, DOI: 10.1039/C7MD00416H.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/medchemcomm

View Article Online
DOI: 10.1039/C7MD00416H

MedChemComm

RESEARCH ARTICLE

Molecular modeling study, synthesis and biological evaluation of Combretastatin A-4 analogues as anticancer agents and tubulin inhibitors

Yang Ping Quan[†], Li Ping Cheng[†,*], Tian Chi Wang, Wan Pang*, Fan Hong Wu*, Jin Wen Huang

As the major structural component of microtubules, tubulin is an interesting target for the development of anticancer agents. In this study, 64 tubulin polymerization inhibitors of five-membered heterocycle-based combretastatin A-4 analogues were studied by combination of molecular modeling techniques including 3D-QSAR, molecular docking, and molecular dynamics (MD) simulation. The CoMFA (comparative molecular field analysis) and CoMSIA (comparative molecular similarity indices analysis) models were established with desirable statistical parameters and excellent predictive ability. 20 ns MD simulations were successfully performed to confirm the detailed binding mode and validate the rationality of docking results. Combined the binding free energy calculations and 3D-QSAR results, some new heterocycle-based combretastatin A-4 analogues were designed. Three of them were synthesized and biologically evaluated. Compound **13a** displayed potent antiproliferative activity (IC_{50} value of 1.31 μM against HepG2 cells, IC_{50} value of 1.37 μM against A549 cells) and inhibition of tubulin polymerization activity (IC_{50} value of 0.86 μM). Compound **13b** also presented a good activity against HepG2 cells (IC_{50} value of 4.75 μM). The experimental results demonstrated that the built models were effective for the development of novel anticancer agents and tubulin inhibitors.

Introduction

Microtubules are cytoskeletal filaments composed of α, β-tubulin heterodimer assemblies. The cellular microtubule system is essential in many important cellular processes such as cell division, formation, and maintenance of cell shape, regulation of motility, cell signaling, secretion and intracellular transport.[1-3] Microtubules are generally recognized as an attractive target for the development of potential new anti-cancer agents.[4-5] Among the tubulin targeting agents, combretastatin A-4 (CA-4), derived from the bark of the South African tree *Combretum caffrum*,[6] is one of the well-known natural tubulin-binding molecules. This cis-stilbene strongly inhibits the polymerization of tubulin by binding to the colchicine site on β-tubulin, prevent the polymerization of tubulin into microtubules, and displays exceptional cytotoxicity towards a variety of cancer cell lines.[7] CA-4 has been entered into clinical trials, but unfortunately its efficacy in vivo is limited due to the drug's poor solubility.[8] A more water-soluble disodium phosphate of CA-4 (CA-4P) (see Fig. **1**) is currently under evaluation, and has shown promising results in human clinical trials.[9]

Due to the relatively simple chemical structure and high affinity for the colchicine binding site, a great number of CA-4 analogues had been synthesized and evaluated in SAR studies.[10-19] Previous SAR studies on combretastatins indicate that both the 3',4',5'-trimethoxy substitution pattern on the A-ring and the cis ethylene linkage between the two aryl rings lead to the optimal activity[10-11] (see Fig. **1**). However, the cis-configuration of CA-4 is prone to isomerize more stable and essentially inactive trans-isomer thermodynamically. To retain the cis-configuration of CA-4 required for bioactivity, a series of derivatives were obtained by incorporating the olefinic double bond into some five-membered aromatic heterocyclic rings.[12-19] For example, Wang and co-workers[12] utilized 1,2-substituted five-membered aromatic heterocycles such as imidazole, oxazole, and pyrazole to mimic the cis double bonds in CA-4. The obtained compounds have potent anti-tubulin and cytotoxic activity. Romagnoli and co-workers confirmed that a thiazole,[13,14] triazole,[15] tetrazole,[16] thiophene,[17-18] or furan[19] ring could replace the cis double bonds to maintain the potent cytotoxicity. Although great progress has been made in experimental research, so far theoretical studies on the mechanism of these compounds toward tubulin as well as the structural factors responsible for the cytotoxicity remain limited. Brown et al.[20] had performed a molecular modeling study for colchicine binding site based on a training set of 23 compounds using Comparative Molecular Field Analysis (CoMFA) method. The built model can accurately predict the IC_{50} for tubulin polymerization with an R^2 of 0.88 (n=6) and those of [^{3}H] colchicine displacement with an R^2 of 0.80 (n=7). Liao et al.[21] reported a theoretical study on the binding conformations and QSAR of CA-4 analogs as inhibitors

*School of Chemical and Environmental Engineering, Shanghai Institute of Technology, Shanghai 201418, China
*Corresponding authors. Telephone: +86-21-6087-3250. email: chengliping@sit.edu.cn; pangwan@sit.edu.cn; wfh@sit.edu.cn
†These two authors (Yang Ping Quan and Li Ping Cheng) contributed equally to this work and should be considered as co-first authors.

View Article Online
DOI: 10.1039/C7MD00416H

toward tubulin based on docking and CoMFA methods. In this study, some key structural factors of the compounds responsible for cytotoxicity were reasonably presented. The results of these studies provide evidences for further designing and synthesizing tubulin inhibitors and conducting structure optimization. However, on the one hand, the total number of samples in their training sets is so small (less than 35) that the reliability of the built models seems to be questionable. On the other hand, although molecular docking had been performed in some theoretical studies to explore the interactions between the inhibitors and receptor proteins, the reliability of docking results should be further verified by molecular dynamics (MD) simulation. Up to now, few research groups had comprehensively studied the interaction mechanism of CA-4 analogs with tubulin receptors by MD simulation. To obtain more understanding of the structural and chemical properties required for tubulin inhibitory activity, more 3D-QSAR and other computational exploration analyses for comparatively large number of newly synthesized and tested tubulin inhibitors are necessary. In the present work, 64 five-membered heterocycle-based CA-4 analogues reported as potent tubulin polymerization inhibitors were collected as dataset. A series of studies were performed by combination of molecular modeling techniques including 3D-QSAR, molecular docking and MD simulation. The results of CoMFA and CoMSIA studies will not only explain the conformation or spatial orientation of CA-4 analogs but also provide useful information for the design of potent tubulin polymerization inhibitors. Molecular docking and MD simulations were performed to validate the 3D-QSAR models and quantify the interactions of tubulin-ligand. Some heterocycle-based CA-4 analogues were designed based on the built 3D-QSAR contour maps. Finally, the syntheses and biological evaluation of some novel heterocycle-based CA-4 derivatives were further carried out. The cytotoxic effects of the synthesized compounds against six human cancer cell lines, A549, HCT-116, HeLa, HepG2, MGC-803, MCF-7, and the inhibitory activity on tubulin polymerization was studied.

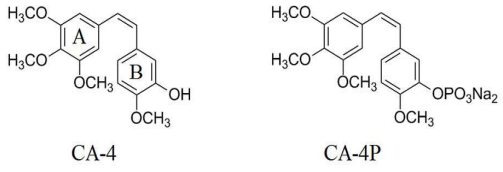


Figure 1. The structures of CA-4 and CA-4P.

Results and Discussion

Data sets

To construct a global predictive combretastatin model for the inhibition of tubulin polymerization, various five-membered heterocycle-based combretastatin A-4 analogues were collected from some different studies[12, 15-18, 22-28] as many as possible in the present study. However, we encountered a problem that the data from different sources were hard to be compared and analyzed. Furthermore, the activity values of CA-4 were not the same in different studies. To eliminate data errors from different sources, a normalization method of processing data was used in the present study. Details are as follows: All data were normalized to an experimental value for combretastatin A-4 (**1**) from the paper in which the data was taken. The active value of CA-4 is set to 1, corrected IC_{50} (IC_{50}c) = IC_{50} (compound) / IC_{50} (CA-4). If one compound has higher activity than CA-4, its IC_{50}c are less than 1, and vice versa. Then the IC_{50}c values are converted to the negative logarithm of this values (pIC_{50}c value), which are used as dependent variables in the 3D-QSAR models. The chemical structures are divided into skeleton templates **a**, **b**, **c**, **d**, and pIC_{50}c values are summarized in *Supplementary Material* (Table **1S**). The whole dataset was randomly divided into a training set of 51 compounds (79.7%) and a test set of 13 compounds (20.3%), respectively. The selection of training and test sets was done by considering structural diversity and the range of biological activities of test compounds similar to that of the training set. The model helped us comprehend the structure-activity relationships of this class of compounds and facilitate the design of novel inhibitors with good activities. In this study, we used the ligand-based alignment rule. The alignment of training set compounds is shown in Fig. **2**.

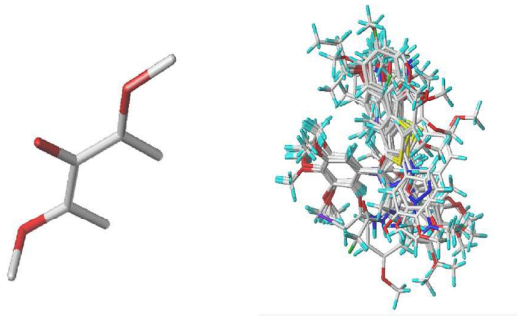


Figure 2. Alignment of the training set and compound **13** used as a template for alignment, with the common substructure shown in bold (left). Molecules are colored in white for common C, red for O, yellow for S, cyan for H, green for F and Cl, respectively.

Molecular modeling and statistical analysis

According to the partial least squares (PLS) method,[29] some statistical parameters were used to evaluate and analysis the stand or fall of these models, including q^2, r^2, standard error of estimate (*SEE*) and *F*-statistic values, r^2pred. Table **1** lists the statistical parameter results of the CoMFA and CoMSIA analyses. From Table **1**, the optimal CoMFA and CoMSIA models show good predictive power with the cross-validated q^2 of 0.702 and 0.700. The correlation coefficient r^2, *SEE* and *F*-statistic values are 0.988, 0.054, and 505.339 for CoMFA model and 0.986, 0.060, and 359.657 for CoMSIA model. The values of r^2_{pred} representing the predictive ability

View Article Online
DOI: 10.1039/C7MD00416H

Table 1. Statistical results for the CoMFA and CoMSIA models

Statistical parameters	CoMFA	CoMSIA
q^2	0.702	0.700
NOC	7	8
r^2	0.988	0.986
r^2pred	0.711	0.623
SEE	0.054	0.060
F	505.339	359.657
fraction	-	-
Steric	0.533	0.099
Electrostatic	0.467	0.273
H-acceptor	-	0.171
H-donor	-	0.227
Hydrophobic	-	0.227

for the CoMFA and CoMSIA models are 0.711 and 0.623, respectively. For the CoMFA model, the contributions of the steric and electrostatic fields are 53.3 and 46.7%. The main contribution in the model was steric field. For the optimal CoMSIA model, the steric, electrostatic, hydrophobic, H-Bond donor and acceptor field contribution was 9.9, 27.3, 22.7, 23.0, and 17.1%, respectively. The electrostatic, H-bond donor and hydrophobic fields were found to be the important contributions in the optimal CoMSIA model. The experimental $pIC_{50}c$, predicted $pIC_{50}c$, and the deviation between predicted and experimental values are shown in *Supplementary Material* (Table **2S**). The relationship between actual and predicted $pIC_{50}c$ values of the training and test set is illustrated in Fig. **3A** and **3B** for CoMFA and CoMSIA models, in which most points are evenly distributed along the line Y = X, suggesting that the 3D-QSAR models are of good quality.

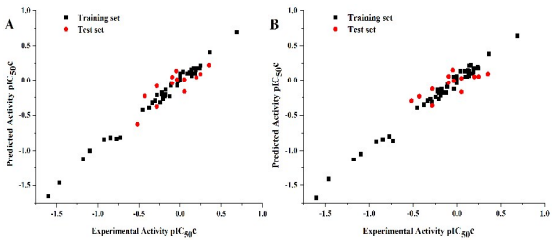


Figure 3. Plots of predicted versus experimental $pIC_{50}c$ values for all the molecules based on CoMFA (**A**) and CoMSIA models (**B**).

CoMFA and CoMSIA contour maps

To visualize the field effects on the target compounds in 3D space, the contour maps generated by CoMFA and CoMSIA were analyzed by superimposing them onto the most active molecule **13**. These contour maps are significant for new drugs design, as they show regions in 3D space where modifications of the molecular fields strongly correlated with variation in biological activity.

The electrostatic field contour maps of CoMFA and CoMSIA are shown in Fig. **4A** and **4B**. The electrostatic field is indicated by blue and red contours, where the blue regions denote that the electropositive groups are favorable to the activity and the red regions indicate that the electronegative groups are favorable to the activity. As shown in Fig. **4A**, it had been recognized that the furan ring and the portion of the benzene ring connected to the C-2 position of furan are encompassed by a large blue colored map. This observation demonstrates that these positions are suitable for substitution with electropositive groups. This is reflected in the inhibitory order of **27** (CH_3) > **30** (CH_2I) > **28** (CH_2Cl), and **33** (Cl) > **34** (F). A large red contour located at the 5-position of benzo[b]furan ring indicates that this position is suitable for substitution with electronegative group. However, in Fig. **4B**, CoMSIA contour map shows that no corresponding red contour presents to the 5-position of benzo[b]furan ring. In Fig. **4B**, a large red contour appears on the 4-position of furan ring. Generally speaking, CoMSIA descriptors are more convincing. It suggests that the substituent at the 5-position of benzo[b]furan may not obtain the desired bioactivity. In addition, two pieces of region of red contour near the trimethoxyphenyl ring indicates that some electronegative groups can be introduced into this region to produce better biological activity.

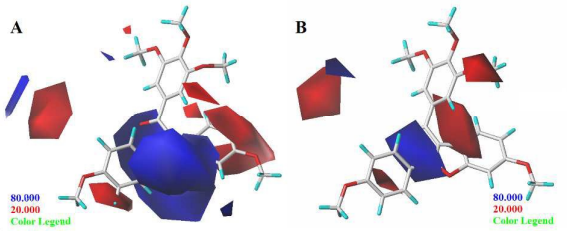


Figure 4. CoMFA (**A**) and CoMSIA (**B**) contour maps. Electrostatic fields: electropositive (blue) and electronegative (red).

As shown in Fig. **5**, the sterically favorable regions are represented in green and unfavorable regions in yellow, respectively. As shown in Fig. **5A**, a large green contour surrounding the C_3~C_6 moiety of the benzene ring which is connected to the C-2 position of furan ring of template molecule **13** indicates the importance of the presence of a bulky group in this region for biological activity. In comparison with molecules of template **a**, **c**, **d**, most molecules in template **b** are substituted by bulky groups at the position of R_7 substituent, which exhibit excellent biological

View Article Online
DOI: 10.1039/C7MD00416H

activity. A yellow contour located at the moiety of benzo[b]furan ring indicates that the presence of bulky substitution in this region will decrease biological activity. For example, the inhibitory order of some compounds is: **26 > 27-30**. As shown in Fig. **5B**, a green contour located at the 6-position of benzo[b]furan ring suggests that compounds with bulky substitution would increase biological activity. For example, the good activity of **52** compared with **48-51** appears to be highly dependent upon the presence of a methoxy substituted H atom at this position. In addition, a large yellow contour surrounding trimethoxyphenyl ring suggests that this position is not suitable for substitution with bulky groups. A large green contour near the -2-(4-methoxy phenyl) suggests that this position is suitable for substitution with bulky groups, which is similar to the result from CoMFA steric contour map.

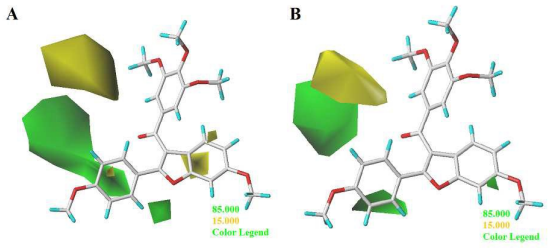


Figure 5. CoMFA (**A**) and CoMSIA (**B**) contour maps. Steric fields: favored (green) and disfavored (yellow).

The CoMSIA hydrophobic contour maps are depicted in Fig. **6A**. Yellow and white contours highlight areas where hydrophobic and hydrophilic properties are favored. As shown in Fig. **6A**, a white polyhedron covering carbonyl group indicates that this position is suitable for substitution with hydrophilic group. Most of the derivatives involved in this study possess hydrophilic groups (such as carbonyl, hydroxy, imino, or five-membered heterocycles containing N or O atoms) at this site, which reveals the importance of the hydrophilic substituent. For instance, the inhibitory order of **42**, **47**, and **1** is **42** > **47** > **1**. One medium-sized region of yellow contour located at the 4-position of benzo[b]furan ring shows the importance of hydrophobic substitute in imparting better biological activity. In addition, the biggest yellow maps are far from the molecular skeleton and they are ignored.

The CoMSIA hydrogen bond donor and accepter contour are shown in Fig. **6B** and **6C**. As shown in Fig. **6B**, the purple contours represent the position where hydrogen bond donor disfavors the biological activity and the cyan contours show that the presence of donor groups in this region should produce better biological activity. A big cyan (H-bond donor favorable) contour map next to the C_4~C_5 moiety of furan suggests that the electronegative H-Bond donors in these regions are favorable for the inhibitory activity. Most of the derivatives involved in this study possess strong electronegative atoms (such as O and N) in the five-membered heterocycle at the furan ring position, which reveals the importance of the H-bond donor substituent. A large cyan contour map covers the meta-position and para-position at the benzene ring of -2-(4-methoxyphenyl). The inhibitory order of **5**, **6**, and **7** (**7** < **6** < **5**) shows that H-Bond donors in this region are beneficial to bioactivity. A purple contour map which was observed above the trimethoxyphenyl indicates that the H-Bond donors in these regions are unfavorable for the inhibitory activity. In addition, other medium-sized cyan contours are far from the molecular, and they are neglected.

As shown in Fig. **6C**, the hydrogen bond acceptor field is represented by magenta and red contours, in which magenta contours denote regions where hydrogen bond acceptor group would be beneficial to the bioactivity, whereas red contours represent hydrogen bond acceptor group would decrease the bioactivity. A magenta contour map covering the connection part and the 4-position of the benzo[b]furan ring and a bit smaller magenta contour near the C-3 position of -2-(4-methoxyphenyl) indicate the H-bond acceptor groups could have a positive influence on the inhibitory activity. Most of the derivatives involved in this study possess groups containing H-bond acceptors (such as carbonyl, hydroxy, imino, or five-membered heterocycles containing N or O atoms) at this site, which reveals the importance of the H-bond acceptor substituents. A red contour map covering the 1,2-position of benzo[b]furan ring implies that H-bond acceptor groups maybe decrease the inhibitory activity.

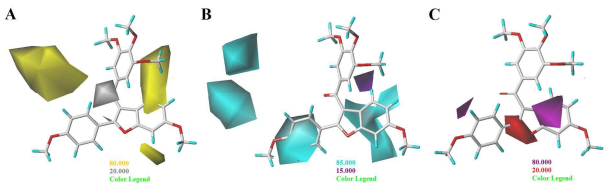


Figure 6. CoMSIA contour maps. (**A**) hydrophobic field: favored (yellow) and disfavored (white). (**B**) Hydrogen bond donor field: favored (cyan) and disfavored (purple). (**C**) Hydrogen bond acceptor field: favored (magenta) and disfavored (red).

Molecular design of new tubulin inhibitors

According to the information derived from contour maps generated by 3D-QSAR models, some information about the template modification were presented for region A, B, C and D on inhibitory activities (Fig. **7**). Based on the summarized structure-activity relationships, six new compounds **13a**-**f** were designed by introduction of F atom or other substituent groups into the regions. The most active molecule **13** was taken as a template molecule. The structures and predicted pIC_{50}c values of **13a-f** based on CoMFA and CoMSIA models are listed in Table **2**. Table **2** shows that the pIC_{50}c values of all designed new molecules are higher than those of CA-4, suggesting their good inhibitory activity.

View Article Online
DOI: 10.1039/C7MD00416H

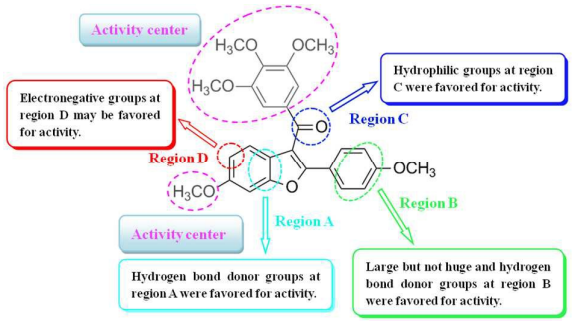
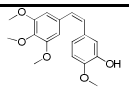
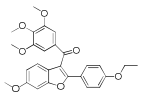
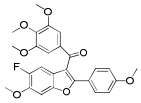
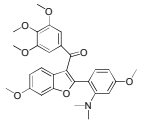
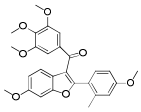
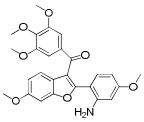
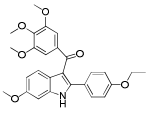


Figure 7. Structure-activity relationships derived from 3D-QSAR studies.

Table 2. Structures and predicted pIC50c values of designed molecules

Compound	Structure	$pIC_{50}c$ (Pred)	
		CoMFA	CoMSIA
CA-4		0.000	0.037
13a		0.351	0.291
13b		0.285	0.308
13c		0.348	0.282
13d		0.425	0.346
13e		0.503	0.206
13f		0.576	0.182

MD simulations analysis

To get a more detailed insight into the stability of the tubulin-inhibitor complex, 20 ns MD simulations and binding free energy calculations were conducted for two representative inhibitors, **CA-4** and **13**, as well the new designed molecules, **13a-f**. All MD simulations were performed using the AMBER 12 package. The overall convergence of MD simulations and system equilibration were monitored by root-mean-square deviation (RMSD) of tubulin-inhibitor backbones atoms (C, C_α, N, and O). The RMSD value (Å) with respect to simulation time (ns) is shown in Fig. **1S** (See *Supplementary Material*). It can be seen that all complexes are mainly stable after 6 ns simulation time run. In addition, the initial structure of **13a**-tubulin (PDB ID: 3UT5) complex and the equilibrium structure after 20 ns MD simulation superimposed in the Fig. **8**. Fig. **8** shows that there seems to be no significant difference between initial structure (in green) and the lowest energy structure (in magenta) except for a slight drift and rotation of bonds. It can be inferred that the binding pocket and the conformation of the ligand are stable, and the docking results are reliable.

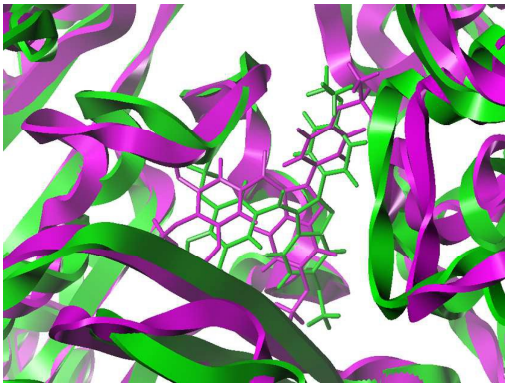


Figure 8. The superimposed the initial structure (green ribbon) and the equilibrium structure (magenta ribbon) of **13a**-3UT5 complexes, initial conformation (green), equilibrium conformation (magenta).

Based on the MD simulations, the binding free energy calculations for seven inhibitors of the last 2 ns trajectory were carried out and are listed in Table **3**. The contributions of different energy components are also reported in Table **3**. The MM/GBSA and MM/PBSA were both used in free energy calculations. For all inhibitors, the ΔG_{gas} and the ΔG_{solv} is negative and positive, respectively, indicating that the former is favorable energetic contribution and the latter is unfavorable energetic contribution to the total binding free energy. It can also be seen that the major contribution to the overall binding free energy is the van der Waals energy of ΔG_{gas}. The binding free energy for the inhibitors **CA-4**, **13**, **13a**, **13b**, **13c**, **13d**, **13e**, and **13f** via MM/GBSA was predicted to be -34.32, -57.59, -57.52, -54.41, -55.78, -50.77, -55.56, and -40.52 kcal·mol^{-1} and via MM/PBSA was predicted to be -21.54, -44.02, -43.45, -42.03, -40.99, -39.79, -36.23, and -27.82 kcal·mol^{-1}, respectively. It is obviously that the binding free energies of **13a-13f** are far more negative than that of **CA-4** based on both MM/GBSA and MM/PBSA method. Among the six new designed compounds, **13a** has the most negative binding free energy, implying that

View Article Online
DOI: 10.1039/C7MD00416H

it perhaps possesses the best inhibitory activity.

Table 3. Binding free energy (ΔGbind, kcal·mol-1) (last 2ns) for ligand-tubulin complexes along with the different energy components

No.	VDW[a]	EEL[b]	ΔG_{gas}	ΔG_{GB}[c]	ΔG_{SA}	ΔG_{solv} (GB)	ΔG_{bind} (GB)	ΔG_{PB}	ΔG_{SA}	ΔG_{solv} (PB)	ΔG_{bind} (PB)
CA4	-46.64	-18.64	-65.28	36.96	-6.01	30.95	-34.32	48.02	-4.29	43.74	-21.54
13	-69.28	-22.54	-92.41	43.02	-8.20	34.82	-57.59	53.83	-5.43	48.40	-44.02
13a	-72.63	-20.17	-92.80	44.16	-8.88	35.28	-57.52	55.18	-5.82	49.35	-43.45
13b	-66.73	-17.49	-84.22	37.87	-8.07	29.81	-54.41	47.47	-5.28	42.19	-42.03
13c	-68.71	-19.20	-87.91	40.06	-7.93	32.13	-55.78	52.75	-5.83	46.92	-40.99
13d	-64.86	-17.65	-82.51	39.97	-8.22	31.74	-50.77	48.49	-5.77	42.72	-39.79
13e	-72.79	-16.77	-89.56	42.44	-8.44	34.00	-55.56	58.79	-5.46	53.33	-36.23
13f	-56.86	-11.73	-68.59	35.09	-7.02	28.07	-40.52	46.23	-5.46	40.77	-27.82

[a]Vander Waals energy; [b]EEL: electrostatic energy; [c]The units of all interaction energy are kcal mol^{-1}

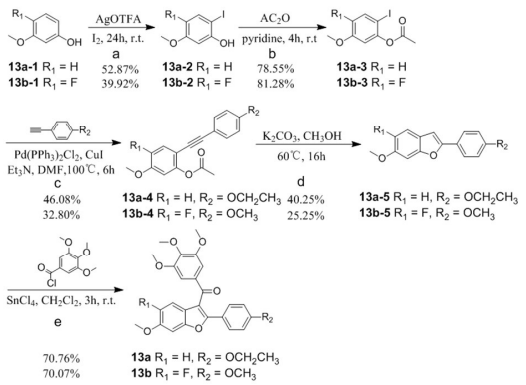
Chemistry

According to the results of binding free energy calculations and the known synthesized CA-4 analogues, three designed compounds **13a**, **13b**, **13f** were firstly selected to be synthesized. The benzo[b]furan derivatives **13a**, **13b** described in this paper were prepared according to the synthetic route (**Scheme-1**),[30] and the route of indole derivative **13f** was shown in **Scheme-2.**[31, 32] The structures of the synthesized intermediates and new benzo[b]furan or indole derivatives were confirmed by ^{1}H NMR, ^{13}H NMR, ^{19}F NMR and HRMS. Iodination of 3-methoxy phenol (**13a-1**) or 4-fluorine-5-methoxy phenol (**13b-1**) was carried out using iodine in presence of catalytic quantity of silver trifluoroacetate (AgOTFA) in chloroform at room temperature for 24 h to obtain the iodide compounds **13a-2** or **13b-2**.[33] Acetylation of iodo-phenol was performed in presence of acetic anhydride in pyridine to afford the corresponding acetate derivatives **13a-3**[34], **13b-3**. Sonogashira coupling of **13a-3** or **13b-3** with 1-ethynyl-4-ethoxybenzene or 1-ethynyl-4-methoxybenzene in presence of $Pd(PPh_3)_2Cl_2$/CuI catalyst load followed by the intramolecular cyclisation in presence of potassium carbonate resulted in the formation of benzo[b]furan. Friedel Crafts procedure was used in presence of 3, 4, 5-trimethoxybenzoyl chloride followed by the addition of **13a-5** or **13b-5** and tin (IV) chloride to afford **13a** or **13b**.

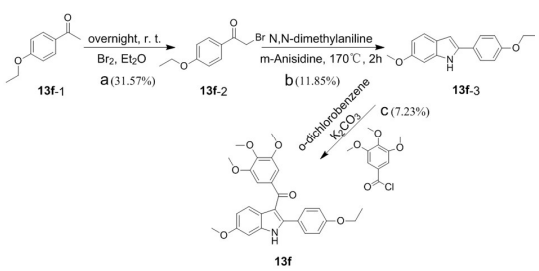
Biological assay

In vitro antiproliferative assay

The benzo[b]furan derivatives (**13a**, **13b**) and indole derivative (**13f**) were evaluated for their antiproliferative activity in vitro against six human cancer cells (A549, lung carcinoma; HCT-116, colon carcinoma; HeLa, cervical carcinoma; HepG2, liver carcinoma; MGC-803, gastric carcinoma; MCF-7, mammary carcinoma) using CCK-8 assays. **CA-4** and **CA-4P** served



Scheme-1 Synthesis of benzo[b]furan derivatives **13a**, **13b**



Scheme-2 Synthesis of indole derivative **13f**

View Article Online
DOI: 10.1039/C7MD00416H

as positive control compounds. The IC_{50} values of the compounds are summarized in Table **4**.

As shown in Table **4**, among the evaluated compounds, compound **13a** exhibits the most potent antiproliferative activity, with IC_{50} values of 1.37, 12.55, 8.99, 1.31, 9.97, and 0.91 μM against A549, HCT-116, HeLa, HepG2, MGC-803, and MCF-7 cells, respectively. It is thus more potent than the reference **CA-4** against the A549, HeLa, and HepG2 cells. Specifically, **13a** is 4.4-fold more active against A549 cells and 12.2-fold more active against HepG2 cells compared with **CA-4**, while its activity against HeLa, HCT-116, and MCF-7 cells is comparable to that of **CA-4**. For compound **13a**, the antiproliferative activities against the six human cancer cells are all superior to those of compound **13b**. It shows an enhancement in antiproliferative activity by more than 5-fold as compared with **13b** except for HeLa and HepG2 cells, while **13b** is more potent than **CA-4** against the HepG2 cells. Compound **13f** exhibits no antiproliferative activities against the cell lines at all. Except for the activity against MCF-7 cells of **13a** and **CA-4**, all the other activities are lower than **CA-4P**.

Table 4. In vitro antiproliferative activities of synthetic compounds

Com.	$IC_{50} \pm SD$ (μM)					
	A549	HCT-116	HeLa	HepG2	MGC-803	MCF-7
13a	1.37 ± 0.14	12.55 ± 0.60	8.99 ± 1.11	1.31 ± 0.17	9.97 ± 0.81	0.91 ± 0.16
13b	6.87 ± 0.34	78.23 ± 3.83	19.1 ± 1.59	4.75 ± 0.20	47.26 ± 1.17	6.49 ± 0.55
13f	-	-	-	-	-	-
CA-4	5.99 ± 0.46	9.00 ± 0.85	11.33 ± 1.03	16.04 ± 0.50	1.76 ± 0.85	0.67 ± 0.13
CA-4P	0.40 ± 0.17	0.24 ± 0.05	0.56 ± 0.12	0.27 ± 0.04	0.47 ± 0.14	0.96 ± 0.22

In vitro tubulin polymerization assay and docking analysis

To investigate whether the designed **CA-4** analogues interact with tubulin, compounds **13a**, **13b**, **13f** were selected to evaluate their inhibitory effects on tubulin polymerization.[35] The corresponding results are shown in Table **5**. For comparison, **CA-4** and **CA-4P** were evaluated as the reference compounds. The result shows that compound **13a** can strongly inhibit tubulin polymerization with IC_{50} value of 0.86 μM, which is comparable with that of **CA-4** (IC_{50}, 0.88 μM) and superior to **CA-4P** (IC_{50}, 4.79 μM). It is a pity that the compound **13b** has no inhibition activity for tubulin polymerization, which exhibits appreciable activity in vitro antiproliferative assays. **13f** has no obvious inhibitory activity for tubulin polymerization either.

Table 5. Inhibition of tubulin polymerization

Compounds	$IC_{50} \pm SD$ (μM)
13a	0.86 ± 0.07
13b	> 200
13f	> 200
CA-4	0.88 ± 0.22
CA-4P	4.79 ± 0.40

To better understand the discrepancy of activity between **13a**, **13b**, and **13f** for tubulin polymerization, a series of molecular docking simulations were carried out. Fig. **9** shows the 3D binding modes of the representative **13a**, **13b**, **13f,** and **CA-4** at the colchicine binding site of tubulin (PDB ID: 3UT5). As shown in Fig. **9**, compound **13a** and **13b** could be perfectly docked into the active site of tubulin. For example, they bound with the active site in a very similar pattern to the binding manner of **CA-4**. The trimethoxyphenyl group takes up a same position as the corresponding moiety of **CA-4**, which is in proximity to the Cys241. The carbonyl groups of **13a** and **13b** keep a cis-configuration for activity like the cis-double bonds of **CA-4**. The 6-methoxy-benzo[b]furan moiety seems to be deeply placed in a tight hydrophobic pocket defined by Met259, Asn258, Ala180, Lys352, and Asn350 and the ethoxyphenyl moiety is located in the hydrophilic region (Fig. **10**). The benzo[b]furan ring of **13a** and **13b** establishes CH-π interactions with Lys352 in Fig. **10**, respectively. In Fig. **10A**, one H atom of 6-methoxy of **13a** forms CH···O non classical H bond with carbonyl of Asn350, and the other two H atoms form CH···O non classical H bonds with carbonyl of Asn258, respectively. Besides, the $-CH_2-$ of 2-(4-ethoxyphenyl) of **13a** have formed two CH···O non classical H bonds with carbonyl of Asn249. However, in Fig. **10B**, two H atoms of 6-methoxy of **13b** form CH···O non classical H bonds with carbonyl of

View Article Online
DOI: 10.1039/C7MD00416H

Asn350 and Asn258, respectively. The rest of H atom of 6-methoxy of **13b** forms a intramolecular non classical H bond with 5-F atom of **13b**, and **13b** does not generate interaction with Asn249 like that of **13a**. These may cause **13b** to lose the inhibition of tubulin polymerization activity. In addition, Fig. **9** and Fig. **11** show that the conformation of **13f** seems to be out of the colchicine binding site. Furthermore, the docking poses of the indole moiety and the ethoxyphenyl moiety of **13f** have an opposite orientation compared with that of **13** (Fig. **11**), **13a** and **13b** (Fig. **9**), which may cause the loss of bioactivity. The conformation of **13** is similar to that of **13a** and **13b**.

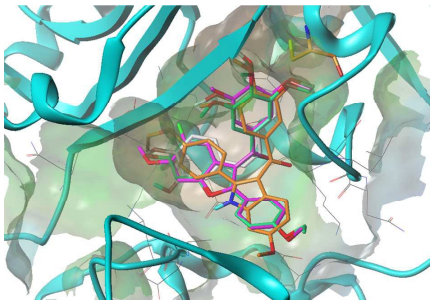


Figure 9. Proposed binding modes of **13a** (in green), **13b** (in magenta), and **13f** (in orange) in the colchicine site of tubulin. **CA-4** as reference compound is shown in white.

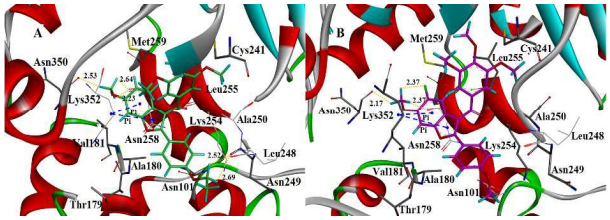


Figure 10. The detailed binding poses of **13a (A)** and **13b (B)**. Yellow dashed lines represent non classical hydrogen bonds. Blue dashed lines represent CH-π interactions.

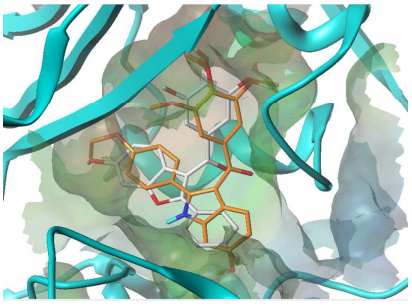


Figure 11. Proposed binding modes of **13** (in white) and **13f** (in orange).

Conclusion

In the present work, a series of five-membered heterocycle-based CA-4 analogues were studied by a combination of computer-aided drug design techniques, i.e. 3D-QSAR study, molecular docking, MD simulations, and binding free energy calculation. The established 3D-QSAR models show significant statistical quality and excellent predictive ability. According to the information derived from contour maps generated by 3D-QSAR models, some information about the template modification were presented for region A, B, C and D on inhibitory activities, six new inhibitors (**13a-13f**) were designed and predicted. Molecular docking for synthetic compounds was carried out. 20 ns MD simulations had been successfully run for representative compounds and their binding free energies were calculated. The total binding free energies of all designed compounds are much more negative than that of CA-4. Based on the theoretical results and the known synthesized scheme, **13a**, **13b**, and **13f** are synthesized, and biologically evaluated in cellular and tubulin inhibition assays. **13a** displays potent antiproliferative activity in cellular assays (IC_{50} value of 1.31 μM against HepG2 cells, IC_{50} value of 1.37 μM against A549 cells) and inhibition of tubulin polymerization activity (IC_{50} value of 0.86 μM). **13b** also exhibits a decent antiproliferative activity with IC_{50} value 4.75 μM against HepG2 cells.

Experiment Section

Data set and alignment

A total of 64 five-membered heterocycle-based combretastatin A-4 analogues were collected from different research groups for constructing 3D-QSAR models. Three dimensional structures of the molecules and subsequent 3D-QSAR studies were performed in SYBYL-X 2.0 (Tripos, Inc., USA). All compounds were constructed using the Sketch Molecule module and then minimized by the Powell gradient algorithm. The energy gradient termination criterion and the maximum iterations were set to 0.005 kcal/(mol·Å) and 100000, respectively. The energy minimization was performed using Tripos force field with Gasteiger-Huckel charges. The treated molecules were divided into a training set (51 compounds) and a validated test set (13 compounds). To obtain a good alignment, the database alignment method was performed using the most active compound as a template.

CoMFA and CoMSIA models

The CoMFA and CoMFA models were developed for the same training set and test set. The CoMFA model was carried on steric and electrostatic fields. The CoMSIA model was carried out using steric, electrostatic, hydrophobic, hydrogen bond donor, and hydrogen bond accepter fields. 3D contour maps of CoMFA and CoMSIA models were graphed using the field type "Stdev*Coeff". Some statistical parameters were used to evaluate the built of models. Partial least squares analysis and leave one

View Article Online
DOI: 10.1039/C7MD00416H

out validation were used to obtain the cross-validation correlation coefficient (q^2) and an optimal number of components (*NOC)*, as well the square correlation coefficient (r^2), F-statistic values (*F*), and standard error of estimate (*SEE*). These statistical parameters described the predictive ability of the models. To further check the predictive ability of the models, the predicted correlation coefficient (r^2_{pred}) was also calculated. For a predictive QSAR model r^2_{pred} value should be higher than 0.5.[36, 37]

Molecular docking

Molecular docking was carried out using the Surflex-Dock module in Sybyl-X 2.0. The crystal structure of tubulin (PDB ID: 3UT5) was downloaded from RSCB Protein Data Bank.[38] The complex with colchicine forming by Chains A and B of the tubulin heterodimer were considered for docking. The ligands were docked into the colchicine binding site by an empirical scoring fuction and a patented search engine in Surflex-Dock. Before the docking process, the colchicine ligand containing in the tubulin complex was extracted and water molecules were removed. Subsequently, protein was prepared by using Biopolymer module implemented in Sybyl. The polar hydrogen atoms were added. AMBER7 FF99 charges were assigned to protein atoms.[39] Other parameters were taken by default. Then a protomol was generated based on original ligand. The molecule was flexible docked into the active pocket. The Suflex-Dock total scores and corresponding conformations were employed for further studies.

Molecular Dynamics (MD) simulations

MD simulations were carried out by Amber 12.0 package. The general amber force field (gaff) and the ff99SB force field were used for the ligands and the protein, respectively. The counter-ions, Na^+ or Cl^-, were added to neutralize the the unbalanced charges in the complexes.[40] Each system was added 10 Å out of the solute with an octahedral TIP3P water box to reduce computational demand.

All the solvated simulations follow the procedure of minimization, heating, density, equilibration and production. Initially to begin with, the minimization of the entire system was performed by two steps. In the first step, the atom position of all solute species was minimized with 500 cycles of each steepest descent and conjugate gradient methods by restraining the protein-ligand complex with a force constant of 10 kcal·mol^{-1}·Å^{-2}. In the second step, the whole system was minimized without any restraints by 1000 cycles of steeps descent and 5000 cycles of conjugate gradient.[41] Then the minimized system was gradually heated under NVT ensemble from 0 to 300 K over a period of 20 ps, followed by density equilibration under NPT ensemble at 300 K over 100 ps with weak restraint of 2 kcal·mol^{-1}·Å^{-2} of constant pressure (1 atm). In this progress the SHAKE algorithm was applied to all bonds involving hydrogen atoms, and the langevin dynamics was used for temperature control.[42] Finally, a 20 ns production run was performed under NPT (1 atm) ensemble with cut-off distance of 10 Å at 300 K. Periodic boundary condition was used for the whole simulation run in order to encounter bulk effect. The integration time step was set to 2 fs and coordinate trajectories were recorded in every 2 ps in the whole simulation. Visualization and analysis were carried out by VMD. The figure of root mean square deviation (RMSD) and cluster analysis were performed by Xmgrace program.

Binding free energy calculations

The Molecular Mechanics/Generalized Born Surface Area (MM/GBSA)[43] and the Molecular Mechanics/Poisson Boltzmann Surface Area (MM/PBSA)[44] methodologies were carried out for binding free energy calculations of systems based on the minimized structures or the MD trajectories of protein-ligand complexes. The calculations were performed on 50 snapshots (frequency 20) from the last 2 ns trajectories (1000 frames). All other parameters were kept at their default values. Here the change in the conformational entropy upon ligands binding was not considered because of the expensive computational cost and low prediction accuracy. The corresponding calculation equations are listed as follows:

$$\Delta G_{bind} = \Delta H - T\Delta S \approx \Delta G_{gas} + \Delta G_{sol} - T\Delta S;$$

$$\Delta G_{gas} = \Delta E_{ele} + \Delta E_{vdw}; \Delta G_{sol} = \Delta G_{PB/GB} + \Delta G_{SA}$$

Where ΔG_{gas} is the gas-phase interaction energy between protein and ligand, it includes ΔE_{ele} (electrostatic), and ΔE_{vdw} (van der Waals) energies. ΔG_{sol} is the sum of electrostatic solvation energy (polar contribution), $\Delta G_{PB/GB}$, and the nonelectrostatic solvation component (nonpolar contribution), ΔG_{SA}.

Chemistry

Commercially available chemicals were used without further purification unless otherwise specified. Thin layer chromatography (TLC) analysis were carried out on pre-coated plates silica gel G F254 and spots were visualized under UV light (254 nm). Chromatography separations were performed on silica gel (300-400 mesh) flash columns by GENERAL-REAGENT® and the eluting solvents are indicated in the procedures. ^{1}H NMR, ^{13}C NMR and ^{19}F NMR spectra were recorded on a Bruker AVANCE III at 500 MHz, 125 MHz or 470 MHz in chloroform-d using TMS (δ = 0.0 ppm) as an internal standard or (Methyl Sulfoxide)-d_6. Chemical shifts are reported in δ parts per million (ppm) with TMS and coupling constants (*J*) in Hz. HRMS were recorded on a solariX 70 FT-MS spectrometer (Bruker) using anhydrous dichloromethane as solvent. Melting points were measured using a WRS-1C melting point apparatus (Shanghai ShenGuang Instrument Co., Ltd., Shanghai China). The purity of the targeted compounds were measured by Agilent 1260 infinity HPLC with 75% CH_3CN + 25% H_2O as mobile phase.

View Article Online
DOI: 10.1039/C7MD00416H

General procedure for the preparation of 13a, 13b

General procedure for the synthesis of compounds 13a-2, 13b-2

I_2 (5.1 g, 20.0 mmol) was dissolved in $CHCl_3$ (150 mL) with stirring over 1.5 h. A 250 mL tinfoil wrapped round-buttomed flask was charged with **13a-1**/**13b-1** (20.0 mmol) and AgOTFA (4.5 g, 20.0 mmol). The solution was added slowly into the tinfoil wrapped round-bottomed flask over 3 h, and the reaction was stirred at room temperature for 24 h. The mixture was filtered saturated $Na_2S_2O_3$, saturated $NaHCO_3$, washed with brine, dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure. The residue was purified via flash chromatography using CH_2Cl_2 to give compound **13a-2**, **13b-2** as a white solid.

General procedure for the synthesis of compounds 13a-3, 13b-3

To a solution of compound **13a-2**/**13b-2** (7.46 mmol) in dichloromethane (40 mL) was added pyridine (1.2 g, 14.92 mmol) followed by drop wise addition of acetic anhydride (1.54 g, 14.92 mmol) at 0 °C. The reaction mixture was stirred for 4 h at ambient temperature and poured into CH_2Cl_2 washed with saturated $NaHCO_3$ and brine solution. The organic layer was dried over anhydrous Na_2SO_4, filtered and evaporated under reduced pressure. The purification was performed by flash chromatography over 300-400 silica gel and eluted with pet-ether/EtOAc to obtain compounds **13a-3** as a colorless oily liquid, **13b-3** as a white crystal.

General procedure for the synthesis of compounds 13a-4, 13b-4

To a solution of compound **13a-3**/**13b-3** (5.0 mmol) in DMF (10 mL) was added 1-ethynyl-4-ethoxybenzene (950 mg, 6.5 mmol) or 1-ethynyl-4-methxoybenzene (860 mg, 6.5 mmol), triethylamine (1.02 g, 10.0 mmol) followed by CuI (50 mg, 0.25mmol) and $Pd(PPh_3)_2Cl_2$ (175 mg, 0.25 mmol) under N_2. The reaction mixture was heated at 100 °C for 6 h.

The reaction was monitored by TLC and then slowly warmed to an ambient temperature, extracted with EtOAc and water several times. The combined organic layer was dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure. The purification was carried out by flash chromatography over 300-400 silica gel and eluted with pet-ether/EtOAc to afford compound **13a-4**, **13b-4** as a yellow solid.

General procedure for the synthesis of compounds 13a-5, 13b-5

To a solution of compound **13a-4**/**13b-4** (1.6 mmol) in methanol (10 mL) was added potassium carbonate (553 mg, 4.0 mmol) and heated to 60 °C for 16 h, reflux. The excess methanol was evaporated under reduced pressure, diluted with water, and extracted with EtOAc, dried over anhydrous Na_2SO_4 and filtered, evaporated under reduced pressure. The purification of **13a-5** was performed by flash chromatography over 300-400 silica gel (pet ether : EtOAc = 30 : 1, v/v) to obtain a white solid. The purification of **13b-5** was eluted with n-hexane, EtOAc to afford a white solid.

General procedure for the synthesis of compounds 13a, 13b

To a solution of compound **13a-5**/**13b-5** (50 mg) in dichloromethane (2 mL) was added 3, 4, 5-trimethoxybenzoyl chloride (1.2 eq) followed by anhydrous $SnCl_4$ (1.0 eq) drop wise at 0 °C. The reaction mixture was stirred at room temperature for 3 h and then quenched with ice and stirred for 1 h. The organic layer was separated and the aqueous layer was extracted with CH_2Cl_2 several times. The combined organic extracts was dried over anhydrous Na_2SO_4, filtered and evaporated under reduced pressure. The purification was carried out by flash chromatography over 300-400 silica gel and eluted with pet-ether/EtOAc to afford compound **13a**, **13b** as a light yellow-green solid.

2-iodo-5-methoxyphenol (13a-2)

White solid, yield: 52.87%. ^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.49 (d, *J* = 9.0 Hz, 1H), 6.60 (d, *J* = 2.5 Hz, 1H), 6.34 (dd, *J* = 6.0, 2.5 Hz, 1H), 5.31 (s, 1H), 3.77 (s, 3H).

2-iodo-5-methoxyphenyl acetate (13a-3)

Colorless oily liquid, yield: 78.55%. ^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.66 (d, *J* = 9.0 Hz, 1H), 6.69 (d, *J* = 3.0 Hz, 1H), 6.60 (dd, *J* = 6.0, 2.5 Hz, 1H), 3.77 (s, 3H), 2.36 (s, 3H).

2-((4-ethoxyphenyl)ethynyl)-5-methoxyphenyl acetate (13a-4)

Yellow solid, yield: 46.08%, mp 121.4-121.7 °C. ^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.46 (d, *J* = 8.5 Hz, 1H), 7.40 (d, *J* = 8.5 Hz, 2H), 6.85 (d, *J* = 8.5 Hz, 2H), 6.78 (dd, *J* = 6.0, 2.5 Hz, 1H), 6.67 (d, *J* = 2.5 Hz, 1H), 4.05 (q, *J* = 7.0 Hz, 2H), 3.82 (s, 3H), 2.36 (s, 3H), 1.42 (t, *J* =7.0 Hz, 3H). ^{13}C NMR (125 MHz, $CDCl_3$) δ = 168.77, 160.28, 159.02, 152.58, 133.40, 132.84, 115.23, 114.54, 112.01, 109.89, 108.26, 92.87, 82.84, 63.53, 55.56, 20.89, 14.76. ESI-HRMS (m/z): calculated for $C_{19}H_{18}O_4$ $(M+H)^+$, 311.12386, found: 311.12381.

2-(4-ethoxyphenyl)-6-methoxybenzofuran (13a-5)

White solid, yield: 40.25%, mp 123.9-125.2 °C. ^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.73 (d, *J* = 9.0 Hz, 2H), 7.41 (d, *J* = 8.0 Hz, 1H), 7.06 (s, 1H), 6.96 (d, *J* =

View Article Online
DOI: 10.1039/C7MD00416H

9.0 Hz, 2H), 6.86 (dd, *J* = 6.0, 2.5 Hz, 1H), 6.80 (s, 1H), 4.09 (q, *J* = 7.0 Hz, 2H), 3.87 (s, 3H), 1.45 (t, *J* = 7.0 Hz, 3H). ^{13}C NMR (125 MHz, $CDCl_3$) δ = 159.03, 157.70, 155.68, 155.39, 125.93, 123.47, 122.88, 120.64, 114.77, 111.67, 99.42, 95.94, 63.55, 55.75, 14.85. ESI-HRMS (m/z): calculated for $C_{17}H_{16}O_3$ $(M+H)^+$, 268.10994, found: 268.10936.

(2-(4-ethoxyphenyl)-6-methoxybenzofuran-3-yl)(3,4,5-trimethoxyphenyl) methanone (13a)

Purity: 97%, yellow-green solid, yield: 70.76%, mp 135.9-137.4 °C. ^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.55 – 7.53 (m, 3H), 7.12 (s, 2H), 7.08 (d, *J* = 2.0 Hz, 1H), 6.91 (dd, *J* = 6.5, 2.0 Hz, 1H), 6.79 (d, *J* = 9.0 Hz, 2H), 4.00 (q, *J* = 7.0 Hz, 2H), 3.88 (s, 3H), 3.86 (s, 3H), 3.69 (s, 6H), 1.39 (t, *J* = 7.0 Hz, 3H). ^{13}C NMR (125 MHz, $CDCl_3$) δ = 190.90, 159.99, 158.50, 157.34, 154.69, 152.86, 142.39, 132.68, 129.84, 122.21, 121.99, 121.70, 114.54, 114.35, 112.60, 107.42, 95.65, 63.55, 60.93, 56.10, 55.76, 14.66. ESI-HRMS (m/z): calculated for $C_{27}H_{26}O_7$ $(M+H)^+$, 463.17121, found: 463.17670.

4-fluoro-2-iodo-5-methoxyphenol (13b-2)

White solid, yield: 39.92%, mp 82.4-83.4 °C. ^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.31 (d, *J* = 10 Hz, 1H), 6.67 (d, *J* = 7.5 Hz, 1H), 5.13 - 5.12 (m, 1H), 3.85 (s, 3H). ^{19}F NMR (470 MHz, $CDCl_3$) δ = -143.08. ^{13}C NMR (125 MHz, $CDCl_3$) δ = 151.73 (d, *J* = 2.5 Hz), 149.2, 146.82 (d, *J* = 240.6Hz), 123.70 (d, *J* = 22.5 Hz), 100.30, 70.97 (d, *J* = 7.5 Hz), 56.27. ESI-HRMS (m/z): calculated for $C_7H_6FIO_2$ $(M+H)^+$, 268.94301, found: 268.94769.

4-fluoro-2-iodo-5-methoxyphenyl acetate (13b-3)

White crystal, yield: 81.28%, mp 84.6-84.9 °C. ^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.47 (d, *J* = 10 Hz, 1H), 6.74 (d, *J* = 7.5 Hz, 1H), 3.85 (s, 3H), 2.35 (s, 3H). ^{19}F NMR (470 MHz, $CDCl_3$) δ = -135.80. ^{13}C NMR (125 MHz, $CDCl_3$) δ = 168.57, 150.05 (d, *J* = 247.5 Hz), 148.50 (d, *J* = 11.3 Hz), 147.60 (d, *J* = 2.5 Hz), 125.22 (d, *J* = 21.3 Hz), 108.28, 99.99, 56.43, 21.15. ESI-HRMS (m/z): calculated for $C_9H_8FIO_3$ $(M+H)^+$, 310.95357, found: 310.95754.

4-fluoro-5-methoxy-2-((4-methoxyphenyl)ethynyl)phenyl acetate (13b-4)

Yellow solid, yield: 32.80%, mp 112.9-114.2 °C. ^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.43 (d, *J* = 9.0 Hz, 2H), 7.27 (d, *J* = 11.5 Hz, 1H), 6.89 (d, *J* = 8.5 Hz, 2H), 6.75 (d, *J* = 7.5 Hz, 1H), 3.90 (s, 3H), 3.84 (s, 3H), 2.37 (s, 3H). ^{19}F NMR (470 MHz, $CDCl_3$) δ = -137.90. ^{13}C NMR (125 Hz, $CDCl_3$) δ = 168.81, 159.88, 149.6 (d, *J* = 243.8 Hz), 148.24 (d, *J* = 12.5 Hz), 147.99 (d, *J* = 3.8 Hz), 132.94, 119.00 (d, *J* = 21.3 Hz), 114.91, 114.09, 109.59 (d, *J* = 10Hz), 107.86, 93.63, 81.94, 56.38, 55.31, 20.77. ESI-HRMS (m/z): calculated for $C_{18}H_{15}FO_4$ $(M+Na)^+$, 337.09544, found: 337.08508.

5-fluoro-6-methoxy-2-(4-methoxyphenyl)benzofuran (13b-5)

White solid, yield: 25.25%, mp 118.6-120.4 °C. ^{1}H NMR (500 MHz, DMSO) δ = 7.79 (d, *J* = 9.0 Hz, 2H), 7.49 (d, *J* = 7.0 Hz, 1H), 7.45 (d, *J* = 11.0 Hz, 1H), 7.17 (s, 1H), 7.06 (d, *J* = 8.5 Hz, 2H), 3.90 (s, 3H), 3.82 (s, 3H). ^{19}F NMR (470 MHz, DMSO) δ = -140.21. ^{13}C NMR (125 MHz, DMSO) δ = 160.05, 156.03, 150.84, 149.80 (d, *J* = 236.3 Hz), 145.83 (d, *J* = 13.8Hz), 126.29, 122.96, 121.40 (d, *J* = 10.0 Hz), 115.00, 107.03 (d, *J* = 22.5 Hz), 100.68, 97.60, 56.93, 55.74. ESI-HRMS (m/z): calculated for $C_{16}H_{13}FO_3$ $(M+H)^+$, 273.08823, found: 273.09217.

(5-fluoro-6-methoxy-2-(4-methoxyphenyl)benzofuran-3-yl)(3,4,5-trimeth oxy phenyl)methanone (13b)

Purity: 98%, yellow-green solid, yield: 70.07%, mp 154.5-154.7 °C. ^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.50 (d, *J* = 8.5 Hz, 2H), 7.38 (d, *J* = 10.5 Hz, 1H), 7.14 (d, *J* = 6.5 Hz, 1H), 7.08 (s, 2H), 6.79 (d, *J* = 8.5 Hz, 2H), 3.94 (s, 3H), 3.85 (s, 3H), 3.76 (s, 3H), 3.68 (s, 6H). ^{19}F NMR (470 MHz, $CDCl_3$) δ = -138.74. ^{13}C NMR (125 MHz, $CDCl_3$) δ = 190.41, 160.76, 158.10, 152.91, 150.58 (d, *J* = 240 Hz), 149.85, 146.81 (d, *J* = 13.8 Hz), 142.60, 132.39, 129.92, 122.14, 120.79 (d, *J* = 8.8 Hz), 114.74 (d, *J* = 3.8 Hz), 113.90, 107.60, 107.42, 96.29, 60.93, 56.61, 56.12, 55.32. ESI-HRMS (m/z): calculated for $C_{26}H_{23}FO_7$ $(M+H)^+$, 467.14614, found: 467.15066.

Procedure for the preparation of 13f

Synthesis of 2-Bromo-1-(4-ethoxyphenyl)ethanone (13f-2)

To a solution of the 1-(4-ethoxyphenyl)ethan-1-one (3.284 g, 20.0 mmol) in Et_2O (30 mL) was added Bromine (3.2 g, 20.0 mmol). The solution was stirred overnight at room temperature (until the color of the solution changed from red to light-yellow). The reaction progress was monitored by TLC. When the starting material disappeared , the reaction was quenched with water, and extracted with EtOAc. The combined organic phases were washed with saturated aqueous $NaHCO_3$ solution, saturated aqueous $Na_2S_2O_3$ solution, brine, and dried over anhydrous Na_2SO_4. After being filtered and concentrated, the residue was purified by flash chromatography over 300-400 silica gel (pet ether : EtOAc = 50 : 1, v/v) to get compound **13f-2** 1.54 g, white crystal, yield 31.57%.

Synthesis of 2-(4-ethoxyphenyl)-6-methoxy-1H-indole (13f-3)

To a boiling mixture of m-anisidine (740 mg, 6.0 mmol) and *N*,*N*-dimethylaniline (3.5 mL) was added compound **2** (730 mg in EtOAc, 3.0 mmol) slowly by syringe. After addition, the mixture was kept at 170 °C for 2 h. The reaction mixture was cooled to room temperature and a dark-colored solid was formed. EtOAc was added along with HCl (2N). The aqueous layer was extracted with EtOAc several times. The combined organic layers were washed with brine, and dried over anhydrous Na_2SO_4. After being filtered and concentrated, the crude compound was purified by flash chromatography over 300-400 silica gel (pet ether : EtOAc = 30 : 1, v/v) to afford compound **13f-3** 95 mg, white solid, yield 11.85%.

View Article Online
DOI: 10.1039/C7MD00416H

Synthesis of (2-(4-ethoxyphenyl)-6-methoxy-1H-indol-3-yl) (3,4,5-trimeth oxyphenyl) methanone (13f)

To a well stirred solution of compound **13f-3** (60 mg, 0.225 mmol), in 1,2-dichlorobenzene (5 mL) was added 3,4,5-trimethoxybenzoyl chloride (104 mg, 0.45 mmol), anhydrous potassium carbonate (83 mg, 0.6 mmol). The reaction mixture was heated to reflux for 12 h. After cooling down to room temperature, solvent was removed by rotary evaporators in vacuo. A dark solid formed which was dissolved in dichloromethane (1 mL), purification through thin layer chromatography (silica gel HSGF254) afforded compound **13f** 10 mg, dark yellow solid, yield 7.23%.

2-Bromo-1-(4-ethoxyphenyl)ethanone (13f-2)

White crystal, yield: 31.57%. ^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.95 (d, *J* = 9.0 Hz, 2H), 6.93 (d, *J* = 8.5 Hz, 2H), 4.39 (s, 2H), 4.10 (q, *J* = 7.0 Hz, 2H), 1.44 (t, *J* = 7.0 Hz, 3H).

2-(4-ethoxyphenyl)-6-methoxy-1H-indole (13f-3)

White solid, yield: 11.85%, mp 129.7-130.5 °C. ^{1}H NMR (500 MHz, $CDCl_3$) δ = 8.30 (s, 1H), 7.55 (d, *J* = 8.5Hz, 2H), 7.11 (t, *J* = 8.0Hz, 1H), 7.01 (d, *J* = 8.0 Hz, 1H), 6.95 (d, *J* = 8.5 Hz, 2H), 6.85 (s, 1H), 6.57 (d, *J* = 7.5Hz, 1H), 4.06 (q, *J* = 7.0 Hz, 2H), 3.99 (s, 3H), 1.45 (t, *J* = 7.0 Hz, 3H). ^{13}C NMR (125 MHz, $CDCl_3$) δ = 158.59, 153.18, 138.01, 136.66, 126.31, 125.12, 122.66, 120.14, 115.05, 104.30, 100.08, 95.99, 63.60, 55.40, 14.84. ESI-HRMS (m/z): calculated for $C_{17}H_{17}NO_2$ $(M+H)^+$, 268.12928, found: 268.13335.

(2-(4-ethoxyphenyl)-6-methoxy-1H-indol-3-yl)(3,4,5-trimethoxyp henyl) methanone (13f)

Purity: 95%, yellow solid, yield: 7.23%, mp n.d. ^{1}H NMR (500 MHz, $CDCl_3$) δ = 10.64 (s, 1H), 7.67 (d, *J* = 8.5 Hz, 2H), 7.63 (d, *J* = 8.5 Hz, 1H), 7.03 (s, 2H), 6.99 (d, *J* = 8.0 Hz, 2H), 6.87 (s, 1H), 6.57 (d, *J* = 8.5 Hz, 1H), 4.13 - 4.09 (m, 2H), 4.07 (s, 3H), 3.96 (s, 3H), 3.92 (s, 6H), 1.45 (t, *J* = 7.0 Hz, 3H). ^{13}C NMR (125 MHz, $CDCl_3$) δ = 195.72, 158.91, 157.93, 152.90, 148.98, 145.59, 138.50, 137.95, 134.78, 130.48, 126.55, 120.43, 115.08, 113.88, 106.87, 99.84, 95.57, 63.62, 60.99, 56.34, 55.70, 14.83. ESI-HRMS (m/z): calculated for $C_{27}H_{27}NO_6$ $(M+H)^+$, 462.19111, found: 462.19010.

Biological assays

Cell proliferation inhibition assays (CCK-8 assay)

The cell viability detection kit, EnoGeneCellTM Counting Kit-8 (CCK-8) was used for cell proliferation inhibition assays.[45, 46] The experiments were carried out on the cells with the ratio of living cells up to 90%. The cells were digested, counted and made into cell suspensions (1*105 cells/mL). 100 μL cell suspensions/well (1*104 cells/well) were seeded in 96-well plates (Costar), cultured in 5% CO_2 incubator (MCO-15AC, SANYO, Japan) at 37 °C for 24 h. 100 μL/well containing corresponding drug mediums was added. The negative control groups, menstruum control groups and positive control groups were set up, 5 wells of each group. The DMSO concentration was kept below 0.01% in cell culture, so it did not effect on cell growth. After being cultured in 5% CO_2 incubator at 37 °C for 72 h, CCK-8 solution was added (10 μL/well) to each well of 96-well plates and incubated for another 4 h. The OD values were measured by microplate reader (Thermo MK3, US) at 450 nm. the IC_{50} values and cancer cells inhibition ratio were obtained. Experiments were repeated two times. All the cell lines used for experiment were obtained from Nanjing OGpharma Co., Ltd., China.

In vitro Tubulin Polymerization Assay

Pig brain microtubule protein was isolated by three cycles of temperature-dependent assembly/disassembly according to Shelanski et al.[47] in 100 mM PIPES (pH 6.5), 1 mM $MgSO_4$, 2 mM EGTA, 1 mM GTP and 1 mM 2-mercaptoethanol. In the first cycle of polymerization, glycerol and phenylmethylsulfonyl fluoride were added to 4 M and 0.2 mM, respectively. Homogeneous tubulin was prepared from microtubule protein by phosphocellulose (P11) chromatography as described 28. The purified proteins were stored in aliquots at -70 °C.

Tubulin protein (4.8 mg/mL) was mixed with different concentrations of compound in PEM buffer (100 mM PIPES, 1 mM $MgCl_2$, and 1 mM EGTA) containing 1 mM GTP and 5 % glycerol. Microtubule polymerization was monitored at 37 °C by light scattering at 340 nm using a SPECTRA MAX 190 (Molecular Device) spectrophotometer. The plateau absorbance values were used for calculations.

Acknowledgments

Thanks for financial support given by Natural Science Foundation of Shanghai (No. 15ZR1440400), National Natural Science Foundation of China (No. 21472126 and 21602135), Collaborative Innovation fund (XTCX2016-14), and the Shanghai Municipal Education Commission (Plateau Discipline Construction Program).

Conflict of Interest

The authors declare no competing interest.

References

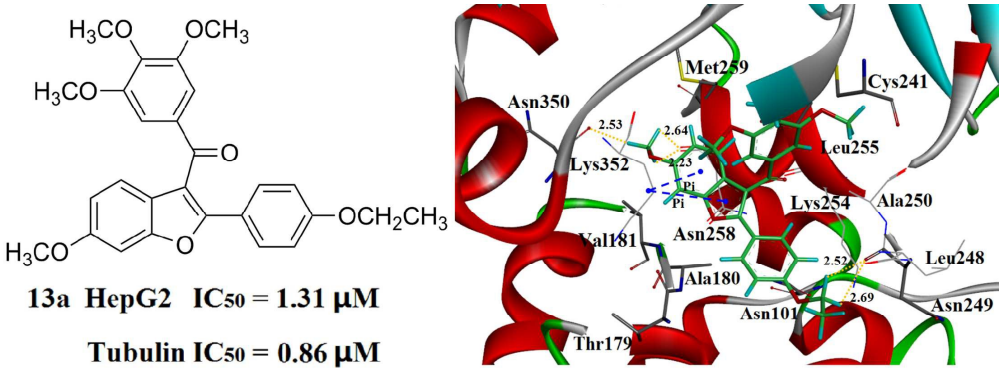
1 K. H. Downing, *Annu. Rev. Cell Dev. Biol.* , 2000, **16**, 89-111.
2 L. A. Amos, *Org. Biomol. Chem.* , 2004, **2**, 2153-2160.
3 S. Honore, E. Pasquier, D. Braguer, *Cell. Mol. Life Sci.* , 2005, **62**, 3039-3056.
4 A. L. Risinger, F. J. Giles, S. L. Mooberry, *Cancer Treat. Rev.* , 2008, **35**, 255-261.
5 D. G. Kingston, *J. Nat. Prod.* , 2009, **72**, 507-515.

View Article Online
DOI: 10.1039/C7MD00416H

6 G. Pettit, S. B. Singh, E. Hamel, C. M. Lin, D. S. Alberts, D. Garcia-Kendal, *Experientia.* , 1989, **45**, 209-211.
7 N. H. Nam, *Curr. Med. Chem.* , 2003, **10**, 1697-1722.
8 K. Ohsumi, R. Nakagawa, Y. Fukuda, T. Hatanaka, Y. Morinaga, Y. Nihei, *J. Med. Chem.* , 1998, **41**, 3022-3032.
9 G. R. Pettit, M. R. Rhodes, *Anti-cancer Drug Des.* , 1998, **13**, 183-191.
10 G. C. Tron, T. Pirali, G. Sorba, F. Pagliai, *J. Med. Chem.* , 2006, **49**, 3033-3044.
11 T. Hatanaka, K. Fujita, K. Ohsumi, *Bioorg. Med. Chem.* Lett. , 1998, **8**, 3371-3374.
12 L. Wang, K. W. Woods, Q. Li, K. J. Barr, R. W. McCroskey, *J. Med. Chem.* , 2002, **45**, 1697-1711.
13 R. Romagnoli, P. G. Baraldi, A. Brancale, A. Ricci, *J. Med. Chem.* , 2011, **54**, 5144-5153.
14 R. Romagnoli, P. G. Baraldi, M. K. Salvador, D. Preti, M. A. Tabrizi, *J. Med. Chem.* , 2012, **55**, 5433-5445.
15 R. Romagnoli, P. G. Baraldi, O. Cruz-Lopez, C. L. Cara, M. D. Carrion, *J. Med. Chem.* , 2010, **53**, 4248-4258.
16 R. Romagnoli, P. G. Baraldi, M. K. Salvador, D. Preti, M. A. Tabrizi, A. Brancale, *J. Med. Chem.* , 2012, **55**, 475-488.
17 R. Romagnoli, P. G. Baraldi, M. K. Salvador, D. Preti, *J. Med. Chem.* , 2013, **56**, 2606-2618.
18 R. Romagnoli, P. G. Baraldi, C. Lopez-Cara, D. Preti, M. A. Tabrizi, J. Balzarini, M.Bassetto, *J. Med. Chem.* , 2013, **56**, 9296-9309.
19 R. Romagnoli, P. G. Baraldi, M. K. Salvador, F. Prencipe, C. Lopez-Cara, S. S. Ortega, A. Brancale, E. Hamel, *J. Med. Chem.* , 2015, **58**, 3209-3222.
20 M. L. Brown, J. M. Rieger, T. L. Macdonald, *Med. Chem.* , 2000, **8**, 1433-1441.
21 S. Y. Liao, J. C. Chen, T. F. Miao, Y. Shen, K. G. Zheng, *J. Enzyme Inhib. Med. Chem.* , 2010, **25**, 421-429.
22 B. L. Flynn, P. Verdier-Pinard, E. Hame, *Org. Lett.* , 2001, **3**, 651-654.
23 R. Romagnoli, P. G. Baraldi, M. G. Pavani, M. A. Tabrizi, D. Preti, F. Fruttarolo, L. Piccagli, K. Jung, E. Hamel, M. Borgatti, R. Gambari, *J. Med. Chem.* , 2006, **49**, 3906-3915.
24 B. L. Flynn, G. S. Gill, D. W. Grobelny, J. H. Chaplin, D. Paul, A. F. Leske, *J. Med. Chem.* , 2011, **54**, 6014-6027.
25 D. J. Kerr, E. Hamel, M. K. Jung, B. L. Flynn, *Bioorg. Med. Chem.* , 2007, **15**, 3290-3298.
26 R. Romagnoli, P. G. Baraldi, T. Sarkar, M. D. Carrion, C. L. Cara, O. Cruz-Lopez, D. Preti, *J. Med. Chem.* , 2008, **51**, 1464-1468.
27 R. Romagnoli, P. G. Baraldi, M. D. Carrion, C. L. Cara, D. Preti, F. Fruttarolo, M. G. Pavani, M. A. Tabrizi, M. Tolomeo, S. Grimaudo, *J. Med. Chem.* , 2007, **50**, 2273–2277.
28 R. Romagnoli, P. G. Baraldi, T. Sarkar, M. D. Carrion, C. L. Cara, O. Cruz-Lopez, C. L. Cara, M. Tolomeo, S. Grimaudo, A. D. Cristina, *Bioorg. Med. Chem.* , 2008, **16**, 8419-8426.
29 L. Ståhle, S. Wold, *J. Chemometr.* , 1987, **1**, 185-196.
30 V. K. Reddy, J. V. Rao, L. B. Reddy, *Asian J. Chem.* , 2015, **27 (6)**, 2245-2248.
31 Patent: Indole-Containing Compounds with Anti-tubulin and Vascular Targeting Activity. , WO2004099139A1.
32 J. Z. Chen, D. L. Liu, N. Butt, *Angewandte Chemie.* , 2013, **52 (44)**, 11632-11636.
33 J. Liu, X. Zhou, C. Wang, *Chem. Commun*. , 2016, **52 (29)**, 5152-5155.
34 K. Hiroya, N. Suzuki, A. Yasuhara, *J. Chem. Soc. , Perkin Trans.* 2000, **1**, 4339-4346.
35 R. Remagnoli, P. G. Baraldi, M. K. Salvador, *J. Med. Chem.* , 2012, **55**, 5433-5445.
36 P. K. Ojha, K. Roy, *Chemometr Intell Lab.* , 2011, **109**, 146-161.
37 A. Tropsha, *Mol Inf.* , 2010, **29**, 476-488.
38 F. M. Ranaivoson, B. Gigant, S. Berritt, M. Joullie, M. Knossow, *Acta Crystallogr Sect D: Biol Crystallogr D.* , 2012, **68**, 927-934.
39 A. Jain, N. Jay, *J. Med. Chem.* , 2003, **46**, 499-511.
40 H. Y. Sun, Y. Y. Li, M. Y. Shen, S. Tian, T. J. Hou, *Phys. Chem. Chem. Phys.* , 2014, **16**, 22035-22045.
41 N. Sanghai, V. Jain, R. Preet, S. Kandekar, *Med. Chem. Comm.* , 2014, **5**, 766-782.
42 S. Miyamoto, P. A. Kollman, *J Comput Chem.* , 1992, **13**, 952-962.
43 P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, *Acc Chem Res.* , 2000, **33**, 889-897.
44 N. Homeyer, H. Gohlke, *Mol Inf.* , 2012, **31**, 114-122.
45 M. Linsalata, M. Notarnicola, V. Tutino, M. BifµLco, *Anticancer Res.* , 2010, **30 (7)**, 2583-2589.
46 M. Sµgiyama, H. Takahashi, K. Hosono, H. Endo, S. Kato, *Int J Oncol.* , 2009, **34 (2)**, 339-344.
47 M. L. Shelanski, F. Gaskin, C. R. Cantor, *Proc Natl Acad Sci U S A.* , 1973, **70**, 765-768.

View Article Online
DOI: 10.1039/C7MD00416H

Graphical abstract



Compound **13a**, more effective than CA-4 against the HepG2 cells and tubulin, and the proposed binding mode for **13a**.