RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

RSC Advances

ARTICLE

Probing the binding mechanism of novel dual NF-κB/ AP-1 inhibitors by 3D-QSAR, docking and molecular dynamics simulations

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Shaojie Ma,[a,d] Shepei Tan,[a] Danqing Fang,[b*] Rong Zhang,[a] Shengfu Zhou,[a] Wenjuan Wu,[a*] and Kangcheng Zheng[c]

Nuclear factor-κB (NF-κB) and activator protein-1 (AP-1) are promising targets for a number of immunoinflammatory diseases, including asthma, psoriasis, rheumatoid arthritis, and transplant rejection. In this work, based on a dataset consisted of 127 pyrimidine/quinazoline-based derivatives as dual NF-κB/AP-1 inhibitors, an integrated computational protocol including the three-dimensional quantitative structure-activity relationship (3D-QSAR), molecular docking and molecular dynamics (MD) simulations was performed to explore the influence of the structural features on the NF-κB and AP-1 inhibitory activities and design derivatives with improved potency. The obtained CoMFA (comparative molecular field analysis) model exhibited satisfactory internal and external predictability. The most probable binding sites of two receptors had been identified by docking and MD simulations, showing that they just locate on the joint regions between NF-κB (or AP-1) and DNA, where inhibitors can effectively prevent free NF-κB (or AP-1) from binding to DNA. At the same time, the key residues/deoxynucleotides for achieving strong binding were also revealed by docking studies, and the detailed dynamic binding process and binding modes of the inhibitors with different activities were determined by MD simulations. The binding free energies are in good agreement with the experimental bioactivities. The decomposition of binding free energies by MM-GBSA suggests that the hydrophobic interactions play an important role for the binding of compounds to NF-κB and AP-1. The results presented here can provide significant insight into the development of novel potential dual NF-κB /AP-1 inhibitors.

1. Introduction

NF-κB and AP-1 are DNA-binding proteins and two important transcriptions factors for the regulation expression in many cytokines[1, 2]. Many studies have shown that the over-activation of NF-κB (or AP-1) is associated with a number of inflammatory and immune diseases, including ashthma, psoriasis, rheumatoid arthritis and transplant rejection [3]. In activated T cells, NF-κB and AP-1 orchestrate the expression of many genes. NF-κB regulates proinflammatory cytokines interleukin-1 (IL-1), IL-6, IL-8, and tumor necrosis factor-α (TNF-α), while AP-1 regulates the production of the cytokines IL-2, IL-3, granulocyte-macrophage colony stimulating factor (GM-CSF) and matrix metalloproteases (MMP)[4-6]. Therefore, the inhibition towards NF-κB and AP-1 mediating transcriptional activation in T cells is very important to treat inflammatory diseases, and NF-κB and AP-1 have become potential targets for the development of novel anti-inflammatory drugs[7, 8].

As for, a variety of NF-κB or AP-1 inhibitors with different scaffolds have been developed[9, 10], but very few compounds are known to inhibit both NF-κB and AP-1. Recently, Palanki and co-workers have synthesized a series of pyrimidine/quinazoline-based derivatives and assessed their activities [11-14]. They discovered that these compounds potently inhibited NF-κB and AP-1 proteins at low nanomolar concentrations, demonstrating the great potential of developing pyrimidine /quinazoline-based derivatives as a novel class of dual NF-κB /AP-1 inhibitors for inflammatory diseases therapy. Moreover, There is now abundant evidence that dual or multi-targeted inhibitors, usually inhibiting different cell pathways or compensatory mechanisms, might be more effective than selective inhibitors, so the search for dual NF-κB/AP-1 inhibitors is a very attractive work[15]. Although some processes have been made in experimental researches, so far the theoretical studies on the structural factors of these compounds which affect their anti-inflammatory activities as

[a] Department of Physical Chemistry, College of Pharmacy, Guangdong Pharmaceutical University, Guangzhou 510006, PR China.E-mail: wuwenjuan83@126.com; Tel: +86 2039352119
[b] Department of Cardiothoracic Surgery, Affiliated Second Hospital of Guangzhou Medical University, Guangzhou 510260, PR China
[c] School of Chemistry and Chemical Engineering,SunYat-Sen University, Guangzhou 510275, PR China
[d] Department of Pharmacy, Kangda college of Nanjing Medical University, Lianyungang 222000, PR China
† Footnotes relating to the title and/or authors should appear here.
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

well as the inhibitory mechanisms of these compounds toward NF-κB and AP-1 kinases remain largely unknown.

In recent years, 3D-QSAR, molecular docking and molecular dynamics (MD) simulations have been widely and successfully applied to guide the design of lead compounds[16, 17]. The 3D-QSAR models can help to find the key structural features affecting the activities and understand the nonbonding interaction characteristics between the drug molecule and target, because they are vivid and robust[18-20]. Meanwhile, molecular docking is an approach to predict the possible orientations of ligand in the active site of receptor, and to study the protein-ligand interactions[21-23]. In addition, MD simulation is a useful methodology providing vivid pictures to depict the fluctuations and conformational changes of molecules, and allowing further investigating the interaction mechanism of a protein complex with a ligand at the atomic level[24]. Therefore, a combined 3D-QSAR, molecular docking and MD simulation study can offer the deep insight into understanding the structural features of ligand-receptor interactions.

In this paper, a novel series of pyrimidine/quinazoline-based derivatives acting as dual NF-κB/AP-1 inhibitors were selected to perform an integrated computational protocol by using molecular docking, MD simulations and 3D-QSAR methods. The purpose is to develop a rational predictive model and to investigate the interaction details between these compounds and NF-κB (and AP-1). The optimum 3D-QSAR CoMFA models were established, and the key structural features contributing to the inhibitory activities were also identified. Then, the orientations and the probable binding modes of these compounds interacting with both NF-κB and AP-1 were located by docking and MD simulations. We hope the obtained results could guide rational design of novel and more efficacious dual NF-κB/AP-1 inhibitors and offer some references for experiment work.

2. Materials and methods

2.1. Date set

A set of 127 pyrimidine/quinazoline-based derivatives with a wide spectrum of anti-inflammatory activities against dual NF-kB and AP-1 receptors[11-14] were collected to perform this study. The general structural formulae of the studied compounds and template molecules **109** and **50** are displayed in Fig. 1. The total set of these derivatives was divided into a training set (88 compounds) for 3D-QSAR model generation and a test set (39 compounds, labeled with an asterisk) for model validation (ESITable S1). The test compounds were selected manually considering the structural diversity and wide range of activities in the data set[25]. All original IC_{50} values were converted to pIC_{50} ($-\log IC_{50}$) values and used as dependent variables in the 3D-QSAR study.

The 3D-structures of pyrimidine/quinazoline-based derivatives were constructed by the sketch molecule module in Sybyl 6.9 software. Structural energy minimization was carried out using the Powell gradient algorithm and the Tripos force field with a convergence criterion of 0.001 kcal•mol^{-1}•Å^{-1} and a maximum of 1000 iterations MMFF94 charges were assigned to each inhibitor[26]. The minimized structure was used as the initial conformation for molecular docking.

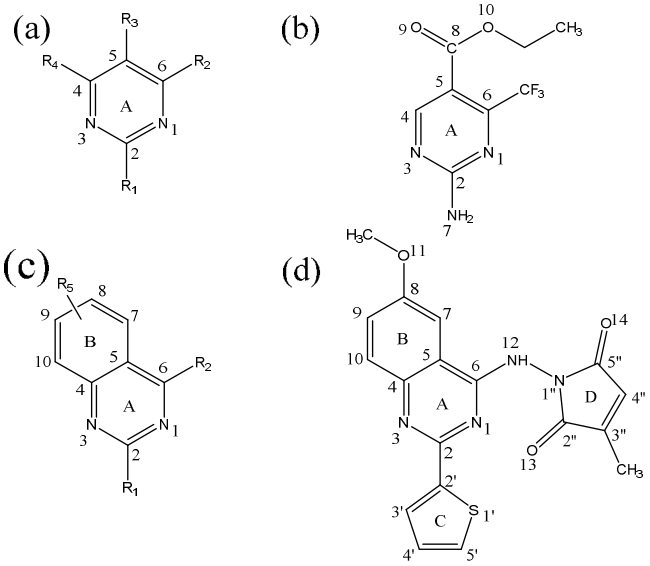


Fig.1 General structural formula and numbering of pyrimidine(a, compounds **1-104)** and quinzolinederivatives (c, compounds **105-127)**, and template molecules (b, compound **50** and d, compound **109**).

2.2. Molecular docking

To determine the probable binding conformations and orientations of the studied derivatives interacting with both NF-κB and AP-1 kinases, docking studies were carried out using the Dock 6.0 program[27].

The X-ray crystal structures of NF-κB (1NFK) and AP-1 (2H7H) were obtained from the Protein Data Bank and used to dock. At the beginning of docking, all the water and subunits were removed, hydrogen atoms and AMBERFF99 charges were added to the protein, and only hydrogen positions were energy-minimized in 10000 cycles with Powell method in SYBYL 6.9. Next, the surface of protein was calculated with DMS program. To obtain possible binding sites, some spheres are generated and selected by Sphgen module of the DOCK 6.0 program. At last, all compounds were flexibly docked into the binding sites where the protein is considered as rigidity. The box size, the grid space, energy cutoff distance, and max orientation were set as 8 Å, 0.3 Å, 12 Å and 10,000, respectively [28].

2.3. 3D-QSAR studies

CoMFA studies were performed using the QSAR option of Sybyl 6.9 software Models of steric and electrostatic fields were based on both Lennard-Jones and Coulombic potentials. The steric and electrostatic fields were calculated at each grid point using a sp3 carbon probe atom with a charge of +1.0, a van der Waals radius of 0.152 nm, a grid spacing of 0.2 nm.

The truncation for both the steric and the electrostatic energies was set to 30 kcal/mol[29, 30].

The 3D-QSAR equations were generated using the partial least square (PLS) statistical method. PLS algorithm with the leave-one-out (*LOO*) cross-validation method was exploited to yield the highest cross-validation correlation coefficient (q^2) and the optimum number of components *N*. The non-cross-validation methods were appraised by the conventional correlation coefficient R^2, standard error of estimates (*SEE*), and F value. To further assess the robustness and statistical validity of the derived models, bootstrapping analysis for 100 runs was also performed[31, 32].

To assess the predictive abilities of 3D-QSAR models generated from the training set, the biological activities of 39 compounds in the external test set were predicted. The predictive power of the models is judged based on the predictive correlation coefficient (R^2_{pred}) calculated by the following equation: R^2_{pred}= (*SD-PRESS*)/*SD* , where *SD* is the sum of the squared derivations between the actual activities of the test set compounds and the mean activity of the training set compounds, and *PRESS* is the sum of the squared derivations between the actual and predicted activities of the test set compounds[33].

2.4. Molecular dynamics simulations

To confirm the docking results, MD simulations were performed with AMBER 9.0 software package[34]. The docked complexes of 1NFK and 2H7H with two compounds (highly active compound **109** and lowly active compound **50**) were used as the initial structures for MD simulations. The electrostatic potentials (ESP) of the ligands were calculated at B3LYP/6-31G (d) level in the Gaussian 09 program, and the partial atomic charges for the ligand atoms were assigned using the RESP protocol implemented in the Antechamber module[35]. The FF03 AMBER force field and the general AMBER force field (gaff) were used to describe the proteins and ligands, respectively. Each complex was neutralized by adding sodium ions and solvated in a truncated octahedron box of water molecules with a margin distance of 12 Å [36, 37].

Subsequently, two-stage energy minimizations were performed to avoid possible steric stress. Firstly, each complex was fixed with a restraint constant of 2.0 kcal•mol^{-1}•Å^{-1} and the water molecules and sodium ions were minimized with steepest descent (SD) method for 2000 steps followed by conjugated gradient (CG) method for 3000 steps. Secondly, the whole relaxed system was optimized by 5000 steps steepest descent minimization and 5000 steps conjugated gradient minimization. Then, each complex was gradually heated from 0 to 300 K in 200 ps with a weak constraint of 1.0 kmol/mol•Å^2 at a constant volume and equilibrated for 500 ps at 300 K and 1 atm. Finally, 10 ns production MD simulation was performed in a NPT (constant composition, T = 300K and P = 1.0 atm) ensemble. During these steps, the particle mesh Ewald (PME) method was used to treat the long-range electrostatic interactions with non-bonded cutoff of 8.0 Å, and the SHAKE algorithm was turned on to constrain all covalent bonds involving hydrogen atoms with 2 fs time step[38, 39]. Coordinated trajectories were recorded every 1 ps and the stability of the complexes was checked from the root mean square deviation (RMSD).

2.5. Binding Free Energy Calculations

To estimate the binding stability of the ligand-protein complexes, the binding free energy calculations of each binding complex were performed by the MM-PBSA procedure in AMBER 9.0 software [40, 41]. For each system, a total of 200 snapshots of the simulated structures extracted from the last 2 ns stable MD trajectory were used for the calculations. The binding free energy (ΔG_{bind}) is calculated as follows:

$$\Delta G_{bind}=G_{complex}-(G_{protein}+G_{ligand})=\Delta G_{MM}+\Delta G_{sol}-T\Delta S \quad (1)$$

where ΔG_{MM} is the molecular mechanics free energy, ΔG_{sol} is the solvation free energy, and $T\Delta S$ is the entropy contribution. The molecular mechanics gas-phase free energy (ΔG_{MM}) contains the van der Waals energy (ΔG_{vdw}) and electrostatic (ΔG_{ele}) energy:

$$\Delta G_{MM}=\Delta G_{vdw}+\Delta G_{ele} \quad (2)$$

The solvation free energy (ΔG_{sol}) is the sum of the electrostatic solvation, including the polar solvation free energy ($\Delta G_{ele,sol}$) and the nonpolar solvation free energy ($\Delta G_{nonpol,sol}$):

$$\Delta G_{sol}=\Delta G_{ele,sol}+\Delta G_{nonpol,sol} \quad (3)$$

The $\Delta G_{ele,sol}$ was determined by solving Poisson Boltzmann (PB) equation with the dielectric constant for solute and solvent set to 4.0 and 80.0, respectively. The $\Delta G_{nonpol,sol}$ was determined by using:

$$\Delta G_{nonpol,sol}=\gamma\times SASA+\beta \quad (4)$$

where *SASA* is the solvent accessible surface area. The solvation parameters of γ and β were set to 0.0072 kcal/mol Å^2, and 0, respectively. As the calculation of entropy term was time-consuming and its value seldom converge, the entropy contribution has been omitted in this study [42].

For discerning the difference of the binding modes of these complexes, the binding free energies were decomposed to each residue using MM-GBSA method. Each inhibitor-residue pair concludes four energy terms: van der Waals contribution (ΔG_{vdw}), electrostatic contribution (ΔG_{ele}), polar solvation contribution ($\Delta G_{ele,sol}$) and nonpolar solvation contribution ($\Delta G_{nonpol,sol}$), which can be summarized as the following equation:

$$\Delta G_{inhibitor\text{-}residue}=\Delta G_{vdw}+\Delta G_{ele}+\Delta G_{ele,sol}+\Delta G_{nonpol,sol} \quad (5)$$

where ΔG_{vdw} and ΔG_{ele} were calculated with sander program in AMBER 9.0. The polar contribution was determined by the generalized Born (GB) model (GB^{OBC}, igb=2). The nonpolar contribution was computed using the solvent accessible surface area (*SASA*) [43].

3. Results and discussion

3.1. Search the favorable binding sites

Prior to docking, it is important to define the binding pockets for NF-κB and AP-1 models. Two docking steps were used to find out possible binding sites of two receptors. Frist, the DNA helix and the subunit of NF-κB (or AP-1) were used as ligand and receptor, respectively, and the spheres within 12 Å were selected as the binding pocket of two proteins. Sequentially, the binding sites defined by the spheres within 6 Å from the

connecting points between DNA and NF-κB (or AP-1) protein. Herein, we selected seven sites 1-7 as the probable binding sites for NF-κB and AP-1 (Fig. 2), respectively.

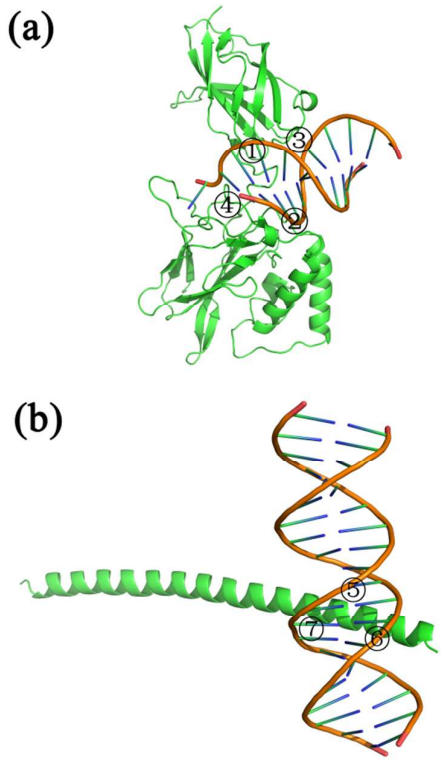
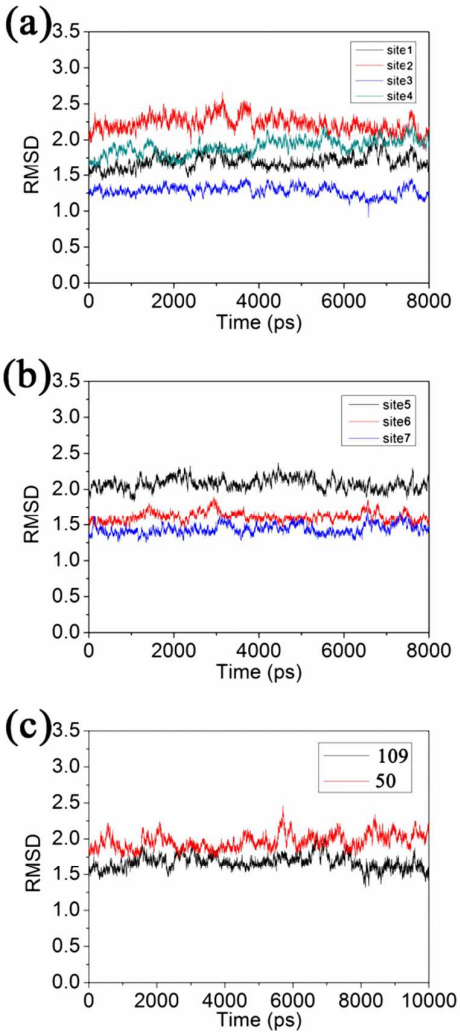


Fig.2 Possible binding sites of the pyrimidine/quinazoline-based derivatives on NF-κB (a) and AP-1 (b) binding to DNA, in which the green represents NF-κB and AP-1 protein, yellow indicates DNA.

In order to find the optimal binding sites of two receptors, docking studies were also carried out on the dataset. All studied compounds were docked into the possible binding sites of NF-κB and AP-1, and the energy scores of these different binding sites are listed in ESI Table S2, where no precise correlations could be found between docking scores and pIC_{50} values. This is not surprising, because the experimental pIC_{50} values are very complicated, depending on not only the binding energy, but also many other factors. The average values of the energy scores are -61.7, -37.1, -45.4, -59.2 and -14.4, -14.6, -21.7 kJ•mol^{-1} from site 1 to 4 for NF-κB and site 5 to 7 for AP-1, respectively. Obviously, the binding sites 1 and 7 have lower average values of the energy scores than sites 2-4 and 5-6, respectively. So, site 1 and site 7 may be the most possible binding sites for NF-κB and AP-1, respectively.

In order to further ensure the rationality of binding sites of two receptors, we selected the above seven binding modes of NF-κB and AP-1 with the highest active compound **109** as initial conformations for following 8.0 ns MD simulations.

First, to argue whether the MD simulations were stable and converged, the RMSD values of protein backbone atoms with respect to the seven starting systems are analyzed and displayed in Fig. 3(a) and (b). The plots show that seven systems reach equilibrium after 0.5 ns. The mean RMSD values were 1.7 Å, 2.2Å, 1.3 Å,1.8 Å, 2.1Å, 1.6 Å and 1.4 Å, respectively, and the relative RMSD fluctuations were very small. These results suggested that the conformations of seven systems were relatively stable throughout the MD simulations. Then, the binding free-energy calculations for the 200 snapshots of the last 1.0 ns MD simulations were also carried out by MM-PBSA method and the corresponding results are -36.68, -16.60, -21.20 and -36.04 kcal/mol from site 1 to 4 for NF-κB, and -29.98, -25.40, and -38.11 kcal/mol for site 5 to 7 of AP-1, respectively. So it is easy to find that the trend of energies for MD simulation is almost the same as that for the docking, and site 1 and site 7 having the lowest binding free energies are the most favorable binding sites for NF-κB and AP-1, respectively. Thus the sites 1 and 7 were selected to further analyses below.



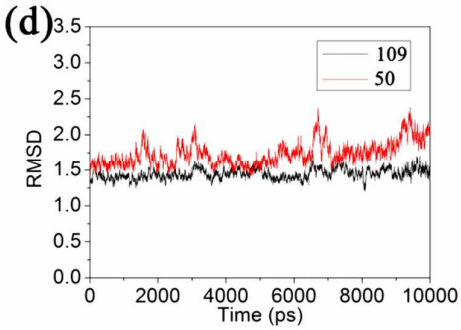


Fig. 3 The rmsd of backbone atoms of the studied complexes during MD simulations.(a) complexes**109**-1NKF in sites 1-4 of NF-κB. (b) complexes**109**-2H7H in sites 5-7 of AP-1.(c) complexes**109**-1NFK and **50**-1NFK. (d) complexes**109**-2H7H and **50**-2H7H.

3.2. Docking Studies

The detailed structures of the most potent dual inhibitor **109** in the binding sites 1 and 7 predicted by the above studies are shown in Fig.4, from which we can find that compound **109** locates the junctions between NF-κB (or AP-1) and the double helix of DNA, and thus it can effectively inhibit NF-κB (or AP-1) from their strongly binding to DNA. Fig. 4(b) shows that compound **109** immerges into a pocket constructed by deoxynucleotide DT-7, DT-8, residues 141-144, 204-208 and 241-244. The ligand can form two H-bonds with DT-7 and Asp206 with corresponding bond lengths of 3.0 Å and 3.3 Å, respectively. Moreover, the electrostatic interaction between the positive His141, Lys144 and the negative S atom of substituent R_1 is very important for stabilizing the ligand. Compound **109** can also make hydrophobic interaction with residues Leu207, Met205, Ser208, Lys241, Ala242 and Pro243. For AP-1 [Fig. 4(d)], the residues nearest to the ligand are Ala13, Ser16, Arg17, Lys20, Leu21 and Arg23, and the deoxynucleotide thymine-207, guanine-208, adenine-209 and guanine-210 can also interact with the ligand. There is one hydrogen bond between the N_1 atom of quinazoline ring and the NH bacкBone of the Lys20. Besides, the electrostatic interaction between the positive Arg17 and the negative S atom on the thienyl ring also contributes to the binding affinity of **109** to AP-1.

Compared with the above docking results for NF-κB and AP-1, we can find that key residues in site 1 (NF-κB) or site 7 (AP-1) (such as Ala243/13, Ser208/16, His144/Arg17, Lys144/20, Leu206/21) are almost the same, demonstrating that the binding conformation and interaction with NF-κB of the target compound **109** are similar to those with AP-1. Therefore, there is no doubt that these compounds acting as dual NF-κB/AP-1 inhibitors.

3.3. CoMFA Statistical results

CoMFA analysis was performed with steric and electrostatic fields, and its statistical parameters were listed in Table 1. This optimal CoMFA model has high non-cross validation R^2 (0.92), q^2 (0.72) and F value (160.2), as well as small SEE (0.28),

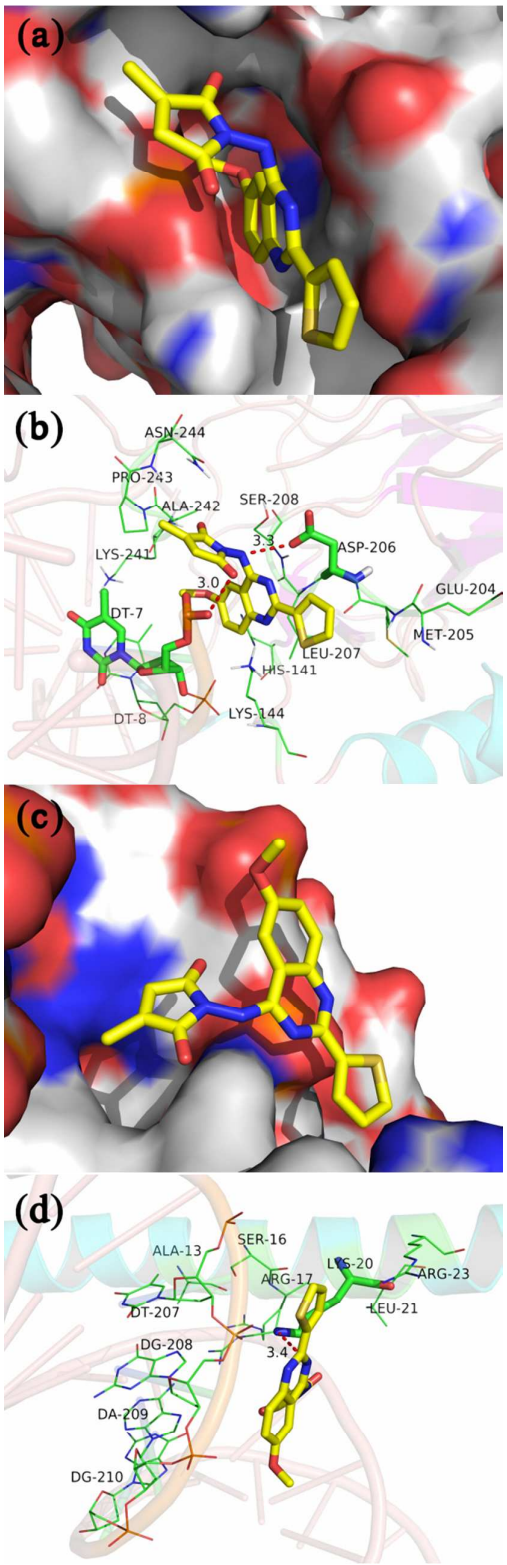


Fig.4 Docking structure of the highly potent compound **109**andcorrespondingsurfaceofNF-κB (a) and AP-1 (c). Interactions between the NF-κB (b) and AP-1 (d) active binding

site and compound **109**. Hydrogen bonds are depicted as dotted lines.

indicating that the established CoMFA model is reliable and predictive for these compounds. Moreover, the R^2_{bs} of 0.95 and SD_{bs} of 0.01 obtained from bootstrapping analysis (100 runs) further verify the statistical validity and robustness of the established CoMFA model. The contributions of the steric and electrostatic fields are 43.5% and 56.5%, respectively, indicating that both the steric field and electrostatic fields have all important influences on the ligand-receptor interactions and that the electrostatic field is more preponderant than the steric field. So both the volume and the polarity of the compound have much impact on its inhibitory activity towards NF-κB and AP-1.

Table 1 Statistical results of the CoMFA model

Statistical parameters	CoMFA
R^2	0.92
N	5
q^2	0.72
SEE	0.28
F	160.2
R^2_{bs}	0.95
SD_{bs}	0.01
R^2_{pred}	0.71

Note: N is the optimal number of components, q^2 is the square of leave-one-out (*LOO*) cross-validation coefficient, R^2 is the square of non-cross-validation coefficient, *SEE* is the standard error of estimation, F is the F-test value, R^2_{bs} is themeanR^2 of bootstrapping analysis (100 runs), SD_{bs} is the mean standard deviation by bootstrapping analysis.

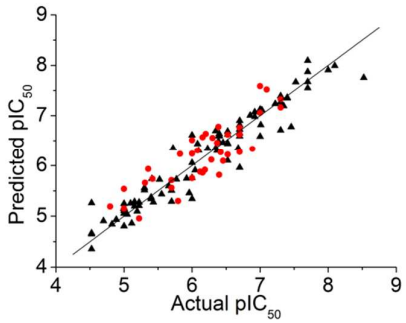


Fig.5 Plots of the predicted versus actual values using the training set (triangle) and test set (dot) based on the CoMFA models.

Furthermore, the test set of 39 inhibitors was used to verify the efficacy of the CoMFA model, where a predictive coefficient R^2_{pred} of 0.71 was achieved, further demonstrating that our CoMFA model has a satisfactory predictive ability. The predicted pIC_{50} values and the residual values of compounds for the CoMFA model are listed in SEI Table S1. The plot of the predicted pIC_{50} values by CoMFA versus the actual ones is depicted in Fig. 5, where most points are evenly distributed along the line $Y = X$, implying that the 3D-QSAR CoMFA model is of good quality.

3.4. CoMFA contour maps analyses

The 3D-QSAR models can be displayed as vivid 3D contour maps, which can provide a more exhaustive interpretation of the biological activity and the related molecular region information, and thus may also be helpful to detect the key residues determining the activities of the studied compounds. Fig. 6(a) shows the sterically favorable (green) and disfavorable (yellow) regions of the template compound **109**. There is a big yellow contour, a small green and a small yellow contours near substituent R_1 and embedding 5'- and 4'- positions of the thienyl ring, suggesting that a moderate-sized substituent R_1 with its 5'- or 4'- atoms being medium-size atom can improve the activity. The docking results shows that the substituent R_1 locates on proximity of the side chain of residue Glu204 and Thr202 in NF-κB (or Arg23 and DT207 in AP-1), suggesting that overlarge groups have unfavorable steric hindrance for the receptor. This is consistent with the fact that compounds **18**, **32**, and **105** with 5'- or 4'- thienyl as substituent R_1 exhibits higher activity than corresponding compounds **21** and **23**, **30** and **34**, and **106** and **107** with 2-benzo-thienyl, phenyl, cyclopropyl, CH_3CH_2 or CF_3 at the same location. Comparing derivative **49** with **48** and **50**, their activity discrepancies can be explained by these contours. A large green polyhedron is embedded in the pyrrole ring of substituent R_2, demonstrating that bulky substituent R_2 would enhance the activity, which is consistent with the fact that there is a big carve in docking. Therefore, compounds **61**, **60** and **59** have an order for the activity of **61**>**60**>**59**, with the corresponding R_2 substituent -ethyl, -methyl, -H, respectively. Comparing derivative **62** with **48**, their activity discrepancies can be explained by this green contour. Four yellow contours and a green contour are found near the C_8- and C_9-positions of ring-**B**, suggesting that moderate-sized substituents at these positions will benefit the activity. This can be explained by the fact that compounds **110** and **127** with OCH_3 or $N(CH_3)_2$ linking to C_9-position have high activities than corresponding compounds **117** and **105**, **126** and **107** with N-morphinoyl, 1-piperidyl, or H at the same location. Compared compound **109** with **105**, as well as **120** with **107**, their activity discrepancies can be also explained by those contours.

The electrostatic contour map of CoMFA is displayed in Fig. 6(b). A big red and a small blue contours near the 1'- position and 3'- position of ring-**C**, suggests that introducing a high electronegative group to the 1'- position or electropositive group to 3'-position of ring-**C** may improve the activity. This may be attributed to electrostatic interactions between the electronegative S atom of ring-**C** and NH_3^+ group of Lys144 in NF-κB (or Arg17 in AP-1).

Most of the excellent derivatives (**105**, **108-116**) all possess an election-withdrawing S atom at the 1'-position, meanwhile, those with high electronegative N or O atoms at the 3'-position of ring-**C** (**75-77**, **50**, **59**, **27** and **10**) are most inactive compounds. Comparing derivative **1** with **2**, as well as **26** with **27**, their activity discrepancies can be explained by this red or blue contour. For substituent R_2, two blue contours

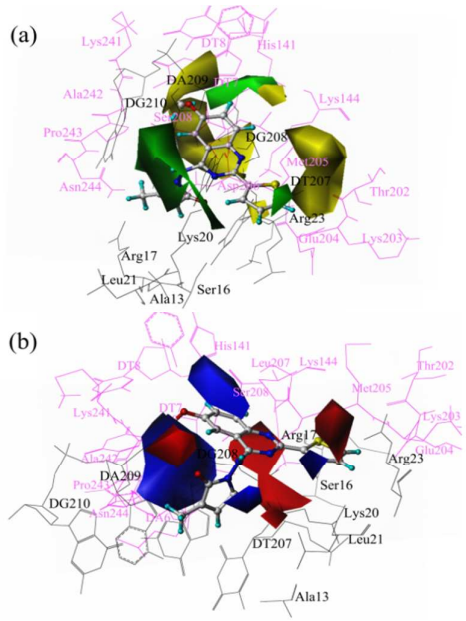


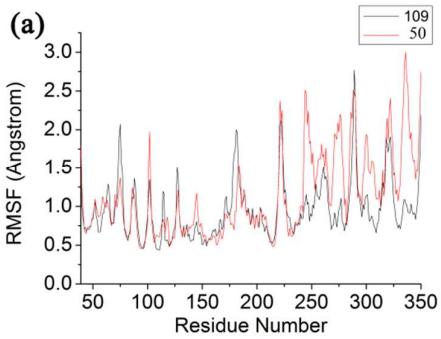
Fig.6 CoMFA contour maps of the highly active compound **109**, in which the purple and black represent NF-κB and AP-1 protein, respectively. (a) Steric contour map, (b) Electrostatic contour map.

embedding in the 2''-, 5''-positions of ring **D** and near the terminal H atoms of -CH_3 at 3''-substituent of ring **D** and C_7-, C_8- positions of ring **B**, implies that electropositive groups at these positions can improve the activity. Therefore, compounds **67**, **69**, **70** have an order for the activity of **69**>**67**>**70**, with the corresponding 3''-substitent of ring **D** -CH_3, -H and -Cl, respectively. Comparing compound **67** with **72** as well as **101** with **98**, their activity discrepancies can be explained by these blue contours. Meanwhile, it can be easily found that four red contours are near the N_{12}, O_{13} and O_{14} atoms, which mean some electronegative groups at these positions are favorable. It indicates that electron-withdrawing groups or atoms linking to ring **D** would increase the activity. Compounds **50**, **59**, **75**-**77** bearing the electropositive C or H atom at 12-position, with -CF_3 or H as substituent R_2 are the most inactive compounds. In addition, there is a bigger blue contour in the vicinity of C_9- and C_{10}-substituents of ring-**B**, suggesting the importance of electropositive atoms or groups on this region. Most of the excellent derivatives (**108**-**113**, **119** and **120**) all possess -OCH_3 linking to ring-**B**, in which -CH_3 with electropositivity just falls into the blue areas. Similarly, the fact that compounds **108** and **109** are more active than compounds **114** and **115** can also be interpreted. Compounds **122**, **123** and **124** have an order for the activity of **122**>**123**>**124**, with the corresponding C_8-substituent -SMe, -SOMe, -SO_2Me, respectively.

From the above, we can deduce that the medium-sized substituent R_1 with electronegative group on 1′-position is favorable to the activity. In Figure 4a and c, it can be apparently observed that the terminal of R_1 is close to the surface of protein and the negative charge can be replenished with the positive charge around S atom. As for R_2, it is also shown that bulky and electronegative groups linking to 2′- and 5′- positions of ring D would increase the activity. In comparison with Figure 4a and c, there are large binding pockets accommodating R_2 in two receptors. Thus, we can conclude that the structural features of binding sites for two enzymes are consistent with the results of 3D-QSAR. We can further validate that sites 1 and 7 may be the optimum binding sites for NF-kB and AP-1, respectively.

3.5. Molecular dynamics simulations

To further investigate the detailed ligand-acceptor interactions in the binding process, the four docking complexes (compounds **109**-2H7H, **50**-2H7H, **109**-1NFK and **50**-1NFK) were performed for 10 ns MD simulations. The RMSD plots [Fig. 3(c) and (d)] indicate that each system is stable throughout the MD simulations. Moreover, the superimpositions of the average structure of the last 1 ns trajectory and the initial docked structures for four systems are displayed in ESI Figure S1, where the blue line represents the initial structure of the docked complex, and the magenta line represents the average MD simulated structure. It can be found that the docked complexes and the MD average structures are well overlapped at the same binding site with only slight positional derivatives, which further verified the reasonability and rationality of the docking results. Furthermore, analyses of root-mean-square fluctuation (RMSF) versus the residue number for four complexes are displayed in Fig. 7. From Fig. 7, we can see that the protein structures of the two complexes in each diagram share similar RMSF distributions and similar trends of dynamic features. The active site, including His141-Lys144, Met204-Ser208, and Lys241-Asn244 for NF-κB (or Ala13, Ser16, Arg17, and Lys20-Arg23 in AP-1), has larger conformational drift for the **50**-1NFK (or -2H7H) system than that for **109**-1NFK (or -2H7H) system, suggesting that the compound **109** should have a more stable interaction with the receptor than **50**. Overall, these analyses for binding stabilization consist with the experimental activities.



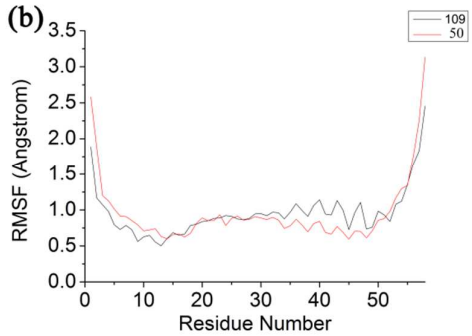


Fig.7 The RMSFs of each residue of the protein for the four systems. (a) complexes**109**-1NFK and **50**-1NFK, (b) complexes **109**-2H7H and **50**-2H7H.

The hydrogen bond interaction is also one of the important forces stabilizing the binding of a ligand to a receptor. For the four systems, the hydrogen bond interactions from MD simulations are listed in Table 2. The hydrogen bond was characterized by distance (< 3.5 Å) and orientation (the angle D-H…A > 120°). From Table 2, we can see that in NF-κB systems, compound **109** (or **50**) can form one strong H-bond with Asp206 (or Tyr57) and another relatively weak H-bond with Lys144 (or Asp206), because of the different structural scaffold of two ligands. While in the **109**-2H7H system, deoxynucleotide DC210 solidly formed a hydrogen bond with compound **109**, with one occasionally formed hydrogen bond was observed between DG208 and compound **109**, instead of the long distance H-bond (3.4 Å) between N_1 and Lys20 in docking study.

Table 2 H-bond analysis from MD

system	donor	acceptor	occupancy (%)	Distance (Å)	Angle (°)
109-1NFK	Asp206@OD2	ligand@N12-H	87.45	2.878	34.18
	ligand@O13	Lys144@NZ-HZ2	11.20	2.933	26.38
50-1NFK	ligand@O9	Tyr57@OH-HH	97.12	3.043	28.86
	Asp206@OD2	ligand@N7-H	21.4	3.009	28.56
109-2H7H	ligand@O14	DC210@N4-H42	67.80	3.092	39.00
	DG208@O2P	ligand @N12-H	7.00	3.040	19.61

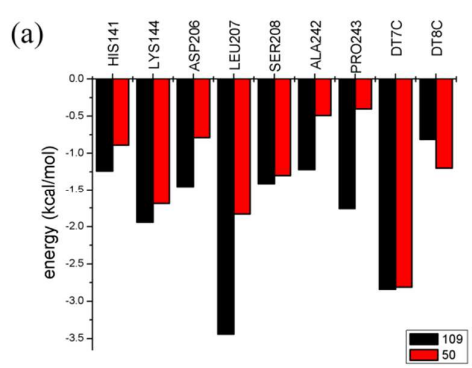
3.6. Binding free energy analysis

The calculated binding free energies of four systems are presented in Table 3. It can be seen that the calculated ΔG_{bind} values for **109**-1NFK and **109**-2H7H systems (-46.05 and -57.59 kcal/mol) were higher than the values of **50**-1NFK and **50**-2H7H systems (-31.17 and -28.10 kcal/mol), suggesting that compound **109** can form stronger binding to the receptor than **50**, which was in accordance with the order of experimental activities. Form Table 3, it can be seen that the van der waals interactions (ΔG_{vdw}) make significant contribution to the binding, and there is great difference of van der waals energies for compounds **109** and **50**. The electrostatic interactions (ΔG_{ele}) and the nonpolar solvation free energy ($\Delta G_{nonpol,sol}$) slightly favor the affinity. Whereas the polar solvation free energy ($\Delta G_{ele,sol}$) opposes the binding strongly. From these results, we can conclude that in the four studied systems, the hydrophobic interactions play a determinant role for stabilizing the ligand in the receptor.

Table 3 The binding free energy of the four systems

Complex	polar contributions		nonpolar contributions		ΔG_{bind}
	ΔG_{ele}	$\Delta G_{ele,sol}$	ΔG_{vdw}	$\Delta G_{nonpol,sol}$	
109-1NFK	-4.70	9.08	-45.11	-5.32	-46.05
50-1NFK	-4.96	7.56	-29.91	-3.86	-31.17
109-2H7H	-7.01	11.94	-56.28	-6.23	-57.59
50-2H7H	-1.14	4.78	-27.57	-4.16	-28.10

To gain further insight into the detailed ligand-receptor interactions, binding free energy was decomposed to ligand-residue pairs using the MM-GBSA approach. From Fig. 8 (a), it can be seen that almost all residues energetically contribute more for the binding of compound **109** than that of compound **50**, especially residues Asp206, Leu207, Ala242 and Pro243. As far as most of latter three residues are nonpolar, it is reasonable to conjecture that there are strong van der waals interactions between inhibitors and these residues. Moreover, the hydrogen bond interactions between Asp206 and inhibitors, which had been validated by docking studies, also played key roles in the binding. Meanwhile, we can also find that the amino acids exhibited more favorable interaction contributions to the binding process of two NF-κB complexes than those of deoxynucleotides. For AP-1 [Fig.8(b)], it is shown that the residues Ala13, Ser16, Arg17, DG208 and DC210 are mainly responsible for the energy difference between **109** and **50**. Obviously, the amino acids and deoxynucleotides have all important influences on the binding free energies for two compounds.



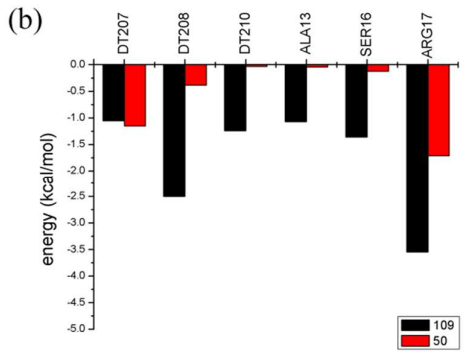


Fig.8 Free energy decomposition plots for the four systems. (a) complexes**109**-1NFK and **50**-1NFK, (b) complexes **109**-2H7H and **50**-2H7H.

The above analyses showed that hydrophobic interactions play a very important role for stabilizing these dual NF-κB/AP-1 inhibitors to the receptors.

4. Conclusions

In the present study, a combined method of the comparative molecular field analysis (CoMFA), molecular docking and MD simulation was performed to study the possible binding modes and the inhibitory mechanism for a series dual NF-κB/AP-1 inhibitors. The CoMFA study gave stable and statistically significant predictive models with high q^2, R^2 and R^2_{pred}, and the structural features influencing the inhibitory activity were discussed in detail. The docking results revealed that the sites 1 and 7 with the lowest average values of energy scores may be the most favorable binding sites for NF-κB and AP-1, respectively. The MD simulation and MM-PBSA calculations confirmed the reasonable binding modes of these complexes at the binding sites and the key interaction features. The calculated binding free energies were in accordance with the experimental activities. The decomposition of binding free energy to each residue revealed that hydrophobic interaction is a predominant factor affecting the binding process. These results could provide a detailed understanding of the binding mechanisms between targets and inhibitors and it can direct the drug molecular design.

Acknowledgements

The authors gratefully acknowledge supports of this research by the Natural Science Foundation of Guangdong (No.S2011010002483, S2011010002964), the Science and Technology planning Project of Guangzhou (No.2013J4100071). We also heartily thank the computation environment support by the Department of Biochemistry, College of Life Sciences, SunYat-Sen University for SYBYL 6.9.

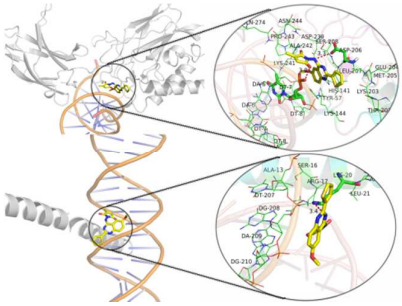
Notes and references

1 H. Kleinert, T. Wallerath, G. Fritz, I. Ihrig - Biedert, F. Rodriguez - Pascual, D. A. Geller and U. Forstermann, *Br. J.Pharmacol.*, 1998, **125**, 193-201.
2 P. Delerive, K. De Bosscher, S. Besnard, W. V. Berghe, J. M. Peters, F. J. Gonzalez, J.-C. Fruchart, A. Tedgui, G. Haegeman and B. Staels, *J. Biol. Chem.*, 1999, **274**, 32048-32054.
3 A. J. Lewis and A. M. Manning, *Curr.Opin. Chem. Biol.*, 1999, **3**, 489-494.
4 P. A. Baeuerle and D. Baltimore, *Cell*, 1996, **87**, 13-20.
5 H. Liu, P. Sidiropoulos, G. Song, L. J. Pagliari, M. J. Birrer, B. Stein, J. Anrather and R. M. Pope, *J.Immunol.*, 2000, **164**, 4277-4285.
6 J. S. Wolf, Z. Chen, G. Dong, J. B. Sunwoo, C. C. Bancroft, D. E. Capo, N. T. Yeh, N. Mukaida and C. Van Waes, *Clin. Cancer Res.*, 2001, **7**, 1812-1820.
7 M. R. Young, H.S. Yang and N. H. Colburn, *Trends Mol. Med.*, 2003, **9**, 36-41.
8 J. K. Kundu and Y.J. Surh, *Mutat. Res.*, 2004, **555**, 65-80.
9 K. Tsuchida, H. Chaki, T. Takakura, H. Kotsubo, T. Tanaka, Y. Aikawa, S. Shiozawa and S. Hirono, *J. Med. Chem.*, 2006, **49**, 80-91.
10 G. Freire, C. Ocampo, N. Ilbawi, A. J. Griffin and M. Gupta, *J. Mol. Cell.Cardio.*, 2007, **43**, 465-478.
11 M. S. Palanki, P. E. Erdman, A. M. Manning, A. Ow, L. J. Ransone, C. Spooner, C. Suto and M. Suto, *Bioorg. Med. Chem. Let.*, 2000, **10**, 1645-1648.
12 M. S. Palanki, P. E. Erdman, L. M. Gayo-Fung, G. I. Shevlin, R. W. Sullivan, M. E. Goldman, L. J. Ransone, B. L. Bennett, A. M. Manning and M. J. Suto, *J. Med. Chem.*, 2000, **43**, 3995-4004.
13 M. S. Palanki, P. E. Erdman, M. Ren, M. Suto, B. L. Bennett, A. Manning, L. Ransone, C. Spooner, S. Desai and A. Ow, *Bioorg. Med. Chem.Lett.*, 2003, **13**, 4077-4080.
14 J. M. Doshi, D. Tian and C. Xing, *J. Med. Chem.*, 2006, **49**, 7731-7739.
15 S. C. Afford, J. Ahmed-choudhury, S. Randhawa, C. Russell,J. Youster, H. A. Crosby, A. Eliopoulos, S. G. Hubscher, L. S. Young and D. H. Adams, *FASEB J.*, 2001, **15**, 2345-2354.
16 L. P. Cheng, X. Y. Huang, Z. Wang, Z. P. Kai and F. H. Wu, *MonatsheftefürChemie-Chemical Monthly*, 2014, **145**, 1213-1225.
17 M. Dong and Y. Ren, *RSC Adv.*, 2015, **5**, 13754-13761.
18 N. Strynadka, M. Eisenstein, E. Katchalski-Katzir, B. Shoichet, I. Kuntz, R. Abagyan, M. Totrov, J. Janin, J. Cherfils and F. Zimmerman, *Nat.Struct. Mol. Biol.*, 1996, **3**, 233-239.
19 S. K. Singh, V. Saibaba, K. S. Rao, P. G. Reddy, P. R. Daga, S. A. Rajjak, P. Misra and Y. K. Rao, *Eur. J. Med. Chem.*, 2005, **40**, 977-990.
20 X. Y. Lu, Y. D. Chen and Q. D. You, *Chem. Biol. Drug Des.*, 2010, **75**, 195-203.
21 C. Selvam, S. M. Jachak, R. Thilagavathi and A. K. Chakraborti, *Bioorg. Med. Chem.Lett.*, 2005, **15**, 1793-1797.
22 H.L. Qin, Z.P. Shang, I. Jantan, O. U. Tan, M. A. Hussain, M. Sher and S. N. A. Bukhari, *RSC Adv.*, 2015, **5**, 46330-46338.
23 M. Zhou, H. Luo, R. Li and Z. Ding, *RSC Adv.*, 2013, **3**, 22532-22543.

24 R. Zhou, B. J. Berne and R. Germain, *Proc. Natl. Acad. Sci. U S A.*, 2001, **98**, 14931-14936.
25 Y. Tan, Y. Chen, Q. You, H. Sun and M. Li, *J. Mol. Model.*, 2012, **18**, 1023-1036.
26 S. Ma, G. Zeng, D. Fang, J. Wang, W. Wu, W. Xie, S. Tan and K. Zheng, *Mol. BioSyst.*, 2015, 11, 394-406.
27 Y. Shamsudin Khan, H. Gutiérrez de Terán, L. Boukharta and J. Aqvist, *J. Chem. Inf. Model.*, 2014, **54**, 1488-1499.
28 R. D. Cramer, D. E. Patterson and J. D. Bunce, *J. Am. Chem. Soc.*, 1988, **110**, 5959-5967.
29 P. Gaillard, P.-A.Carrupt, B. Testa and A. Boudon, *J.Comput.Aided Mol. Des.*, 1994, **8**, 83-96.
30 S. S. Kulkarni, L. K. Gediya and V. M. Kulkarni, *Bioorg. Med. Chem.*, 1999, **7**, 1475-1485.
31 V. Srivastava, S. Gupta, M. Siddiqi and B. Mishra, *Eur. J. Med. Chem.*, 2010, **45**, 1560-1571.
32 M. Rybka, A. G. Mercader and E. A. Castro, *Chemomer.Intell. Lab.*, 2014, **132**, 18-29.
33 M. Baroni, G. Costantino, G. Cruciani, D. Riganelli, R. Valigi and S. Clementi, *Quantitative Structure - Activity Relationships*, 1993, **12**, 9-20.
34 D. A. Case, T. Darden, T. E. Cheatham III, C. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, D. A. Pearlman and M. Crowley, *University of California, San Francisco*, 2006.
35 P. Dobeš, J. Fanfrlík, J. Řezáč, M. Otyepka and P. Hobza, *J.Comput.Aided Mol. Des.*, 2011, **25**, 223-235.
36 R. Salomon - Ferrer, D. A. Case and R. C. Walker, *Wiley Interdiscip Rev.Comput. Mol. Sci.*, 2013, **3**, 198-210.
37 O. Bertran, B. Zhang, A. D. Schlüter, A. Halperin, M. Kröger and C. Alemán, *RSC Adv.*, 2013, **3**, 126-140.
38 T. I. Cheatham, J. Miller, T. Fox, T. Darden and P. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 4193-4194.
39 P. F. Batcho, D. A. Case and T. Schlick,*J. Chem. Phys.,* 2001, **115**, 4003-4018.
40 P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan and W. Wang, *Accounts of chemical research*, 2000, **33**, 889-897.
41 N. N. Mhlongo and M. E. Soliman, *RSC Adv.*, 2015, **5**, 10849-10861.
42 J. M. Hayes, V. T. Skamnaki, G. Archontis, C. Lamprakis, J. Sarrou, N. Bischler, A. L. Skaltsounis, S. E. Zographos and N. G. Oikonomakos, *Proteins*, 2011, **79**, 703-719.
43 D. R. Roe, A. Okur, L. Wickstrom, V. Hornak and C. Simmerling, *J. Phys. Chem. B*, 2007, **111**, 1846-1857.

Probing the binding mechanism of novel dual NF-κB/ AP-1 inhibitors by 3D-QSAR, docking and molecular dynamics simulations

Shaojie Ma,[a] Shepei Tan,[a] Danqing Fang,[b*] Rong Zhang,[a] Shengfu Zhou,[a] Wenjuan Wu,[a*] and Kangcheng Zheng[c]



Potent dual NF-κB/AP-1 inhibitors could effectively treat immunoinflammatory diseases so theoretical studies of dual NF-κB/AP-1 inhibitors are very significant. An integrated computational study was carried out to identify the most favourable binding sites, the structural features and the interaction mechanisms.

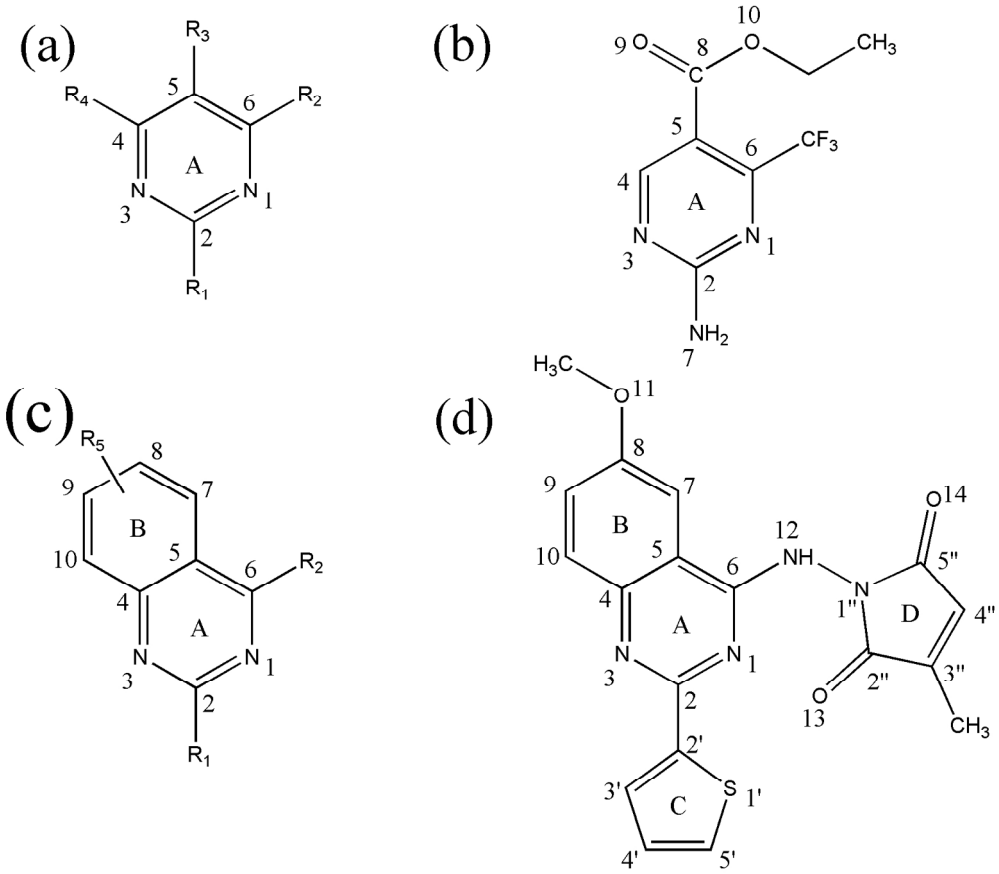


Fig.1 General structural formula and numbering of pyrimidine(a, compounds 1-104) and quinzolinederivatives (c, compounds 105-127), and template molecules (b, compound 50 and d, compound 109).
163x143mm (300 x 300 DPI)

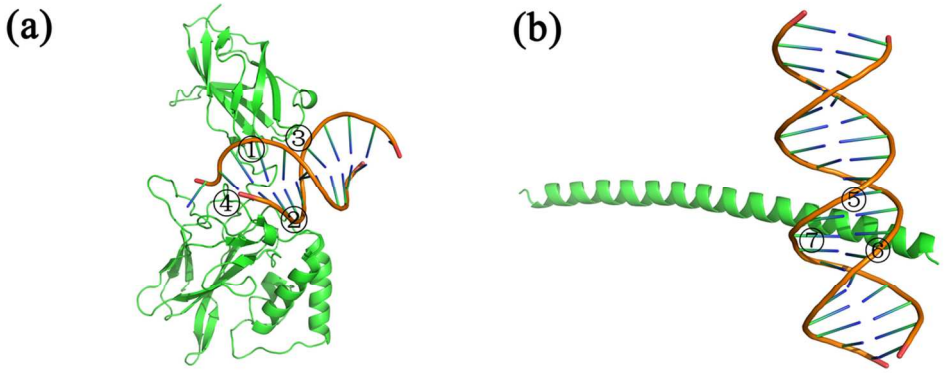


Fig.2 Possible binding sites of the pyrimidine/quinazoline-based derivatives on NF-κB (a) and AP-1 (b) binding to DNA, in which the green represents NF-κB and AP-1 protein, yellow indicates DNA.
54x21mm (600 x 600 DPI)

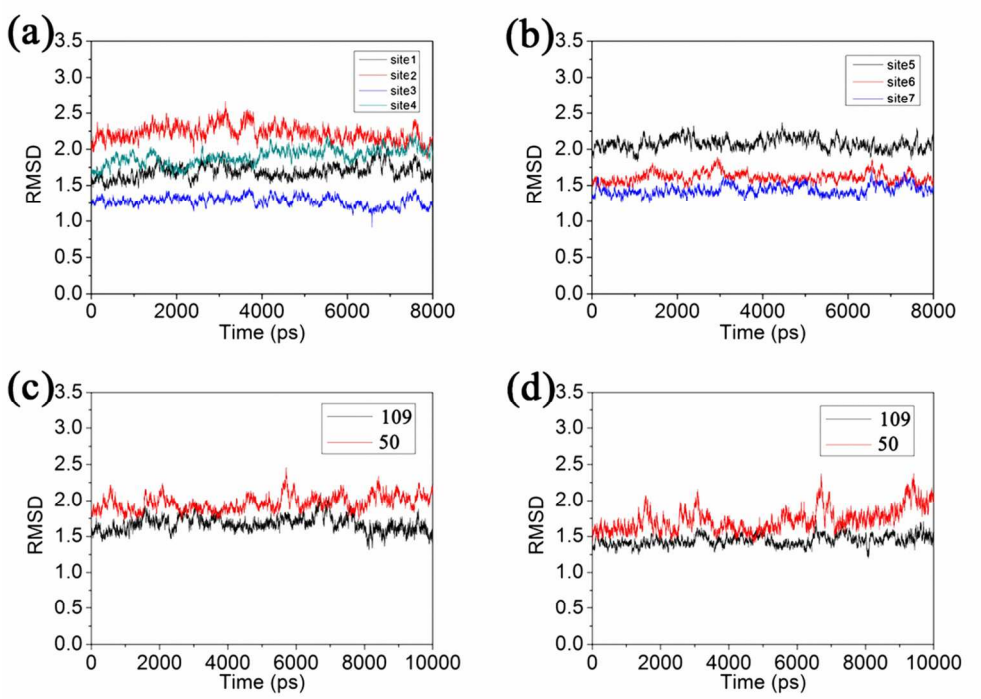


Fig. 3 The rmsd of backbone atoms of the studied complexes during MD simulations.(a) complexes109-1NKF in sites 1-4 of NF-κB. (b) complexes109-2H7H in sites 5-7 of AP-1.(c) complexes109-1NFK and 50-1NFK. (d) complexes109-2H7H and 50-2H7H.
98x69mm (300 x 300 DPI)

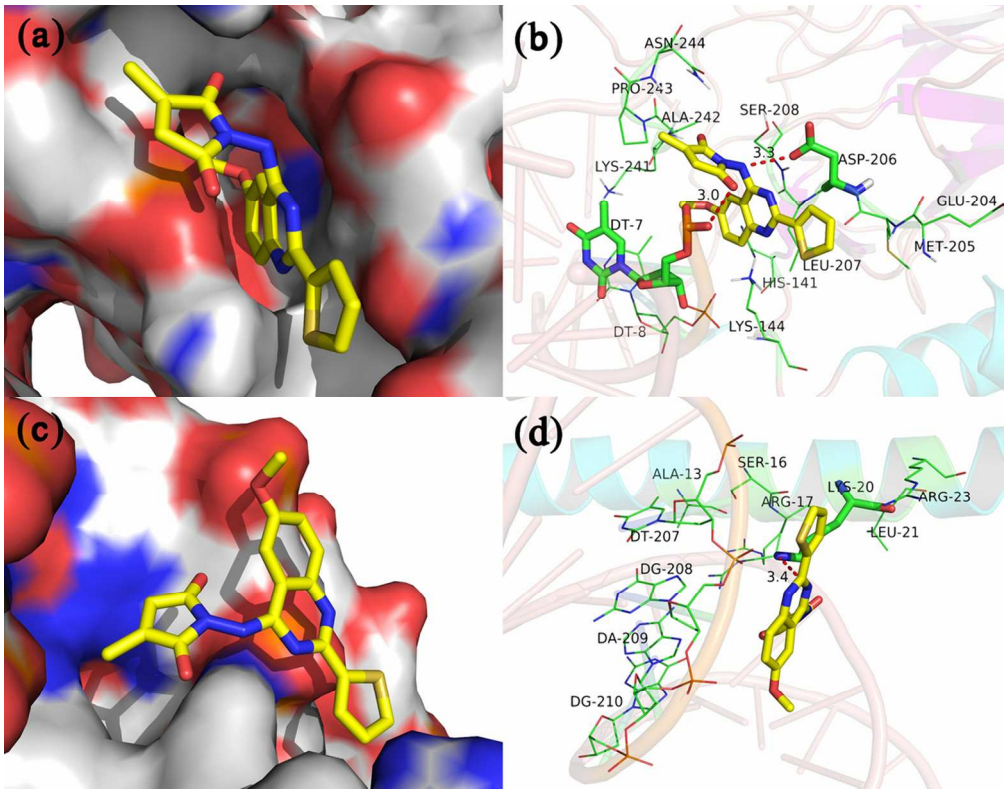


Fig.4 Docking structure of the highly potent compound 109andcorrespondingsurfaceofNF-κB (a) and AP-1 (c). Interactions between the NF-κB (b) and AP-1 (d) active binding site and compound 109. Hydrogen bonds are depicted as dotted lines.
109x85mm (300 x 300 DPI)

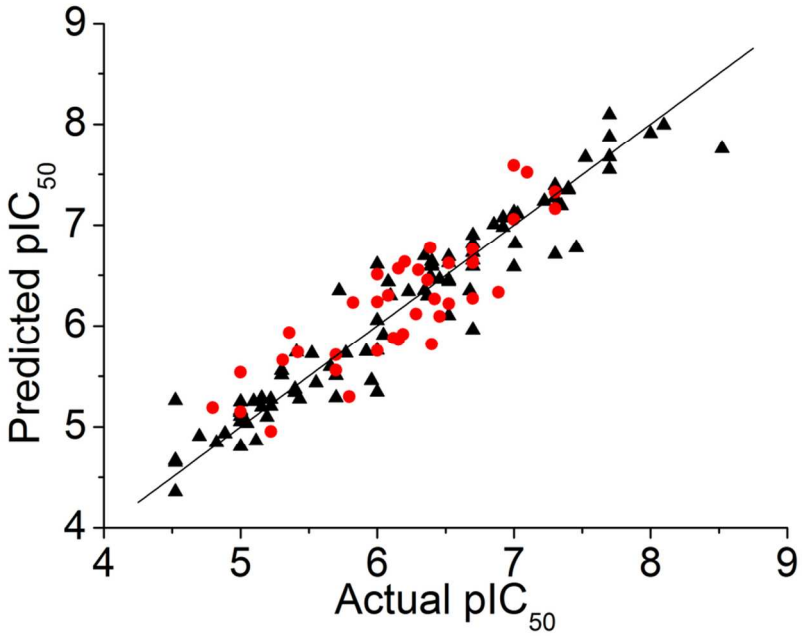


Fig.5 Plots of the predicted versus actual values using the training set (triangle) and test set (dot) based on the CoMFA models.
49x34mm (600 x 600 DPI)

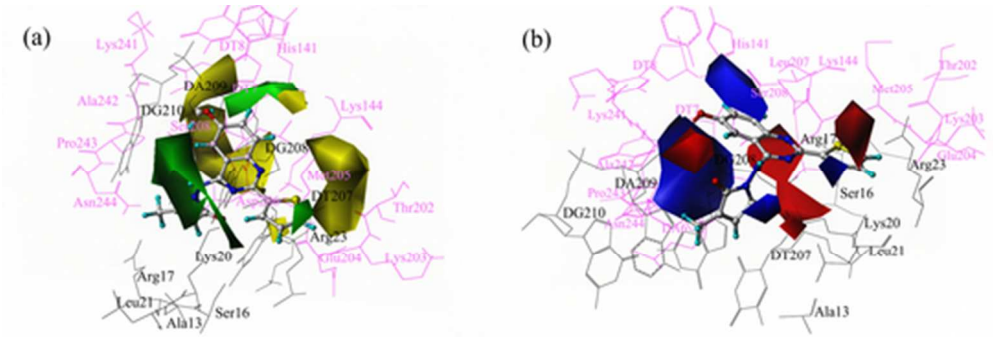


Fig.6 CoMFA contour maps of the highly active compound 109, in which the purple and black represent NF-κB and AP-1 protein, respectively. (a) Steric contour map, (b) Electrostatic contour map.
46x15mm (300 x 300 DPI)

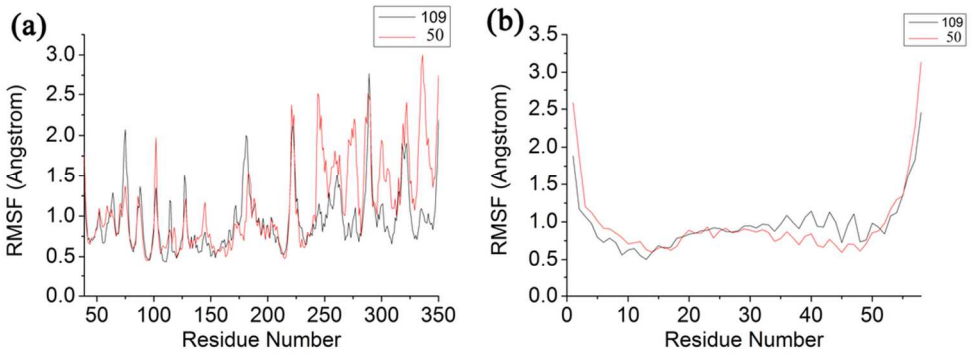


Fig.7 The RMSFs of each residue of the protein for the four systems. (a) complexes109-1NFK and 50-1NFK, (b) complexes 109-2H7H and 50-2H7H.
49x17mm (600 x 600 DPI)

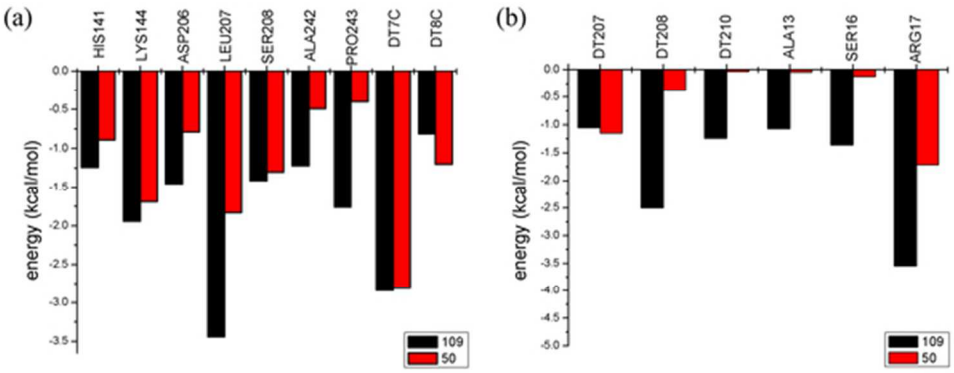
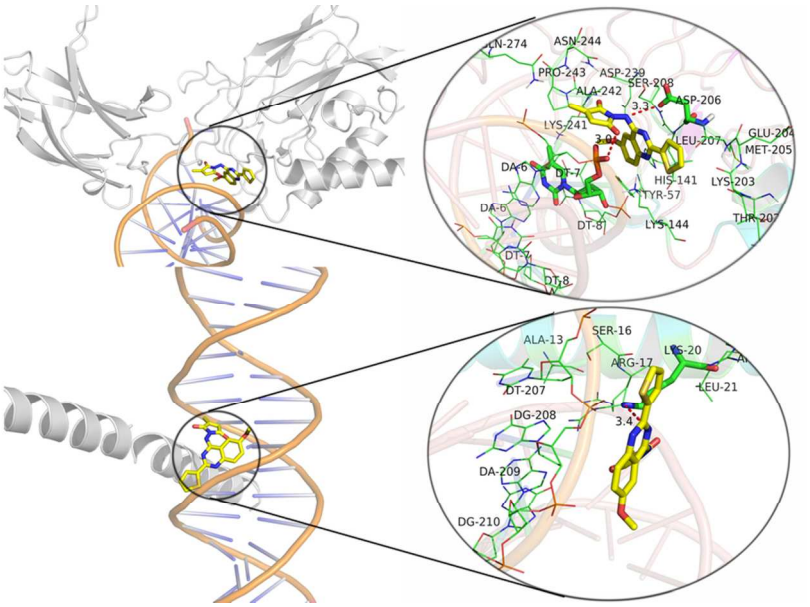


Fig.8 Free energy decomposition plots for the four systems. (a) complexes109-1NFK and 50-1NFK, (b) complexes 109-2H7H and 50-2H7H.
52x20mm (300 x 300 DPI)



47x28mm (600 x 600 DPI)