J Mol Model (2012) 18:295–306
DOI 10.1007/s00894-011-1041-4

ORIGINAL PAPER

# Structural determinants of benzodiazepinedione/peptide-based p53-$HDM_2$ inhibitors using 3D-QSAR, docking and molecular dynamics

**Fangfang Wang · Yan Li · Zhi Ma · Xia Wang · Yonghua Wang**

Received: 17 December 2010 / Accepted: 14 March 2011 / Published online: 27 April 2011


**Abstract** As a tumor suppressor, p53 protein regulates the cell cycle and is involved in preventing tumorgenesis. The protein level of p53 is under the tight control of its negative regulator human double minute 2 ($HDM_2$) via ubiquitination. Therefore, the design of inhibitors of $HDM_2$ has attracted much interest of research on developing novel anticancer drugs. Presently, two classes of molecules, i.e., the 1,4-benzodiazepine-2,5-diones (BDPs) and N-Acylpolyamine (NAPA) derivatives were studied by three-dimensional quantitative structure–activity relationship (3D-QSAR) modeling approaches including the comparative molecular field analysis (CoMFA) and comparative molecular similarity index analysis (CoMSIA) as promising p53-$HDM_2$ inhibitors. Based on both the ligand-based and receptor-guided (docking) alignments, two optimal 3D-QSAR models were obtained with good predictive power of $q^2$=0.41, $r^2_{pred}$=0.60 for BDPs, and $q^2$=0.414, $r^2_{pred}$=0.69 for NAPA analogs, respectively. By analysis of the model and its related contour maps, it is revealed that the electrostatic interactions contributed much larger to the compound binding affinity than the steric effects. And the contour maps intuitively suggested where to modify the molecular structures in order to improve the binding affinity. In addition, molecular dynamics simulation (MD) study was also carried out on the dataset with purpose of exploring the detailed binding modes of ligand in the $HDM_2$ binding pocket. Based on the CoMFA contour maps and MD-based docking analyses, some key structural aspects responsible for inhibitory activity of these two classes of compounds were concluded as follows: For BDPs, the $R_1$ and $R_3$ regions should have small electronegativity groups; substituents $R_2$ and $R_4$ should be larger, and $R_3$ substituent mainly involves in H-bonds forming. For NAPA derivatives, bulky and electropositive groups in ring B and ring A, small substituent at region P is favorable for the inhibitory activity. The models and related information, we hope, may provide important insight into the inhibitor-p53-$HDM_2$ interactions and be helpful for facilitating the design of novel potent inhibitors.

**Keywords** p53 · $HDM_2$ · 3D-QSAR · CoMFA · CoMSIA · Molecular docking · Molecular dynamics

**Electronic supplementary material** The online version of this article (doi:10.1007/s00894-011-1041-4) contains supplementary material, which is available to authorized users.

F. Wang · Z. Ma · X. Wang · Y. Wang
Bioinformatics Center, Northwest A&F University, Yangling, Shaanxi 712100, China

Y. Li
School of Chemical Engineering, Dalian University of Technology, Dalian, Liaoning 116024, China

Y. Wang (✉)
College of Life Sciences, Northwest A&F University, Yangling, Shaanxi 712100, China
e-mail: yh_wang@nwsuaf.edu.cn

## Introduction

The tumor suppressor p53 protein is a transcription factor which involves in controlling the cell cycle and monitoring the integrity of the genome which has made it known as the "guardian of the genome" [1]. Under non-stressed conditions, p53 has a short half-life [2] and maintains in a low level. While under stresses, such as DNA damage and abnormal growth regulation, p53 will be activated and initiates a succession of events which result in growth arrest or apoptosis leading to the elimination of genetically altered cells [3]. Because of its significant roles in keeping the

integrity of genome, it has become a main target of much tumor therapeutics. The loss of p53 functions in many cancers is a result of disabling p53 through mutation or alterations of various components of the pathways that regulate p53. Evidence has shown that p53 mutates in ~50% of human cancers, and in the remaining of human cancers, p53 is in the wild-type but dysfunctional situation [4, 5], due to the precise control of its negative regulator $HDM_2$ in cells [6, 7].

$HDM_2$ is a ring finger protein, which can form autoregulatory feedback loop with p53 protein [8, 9]. On one hand, $HDM_2$ binds to the N-terminal transactivation domain of p53 to block p53 transcriptional activity [10, 11]. On the other hand, $HDM_2$ exports the complex from the nucleus to the cytoplasm to promote its degradation. More importantly, $HDM_2$ serves as an $E_3$ ligase to target p53 for degradation through the proteasome pathway [12]. The detailed interaction mode between $HDM_2$ and p53 has been revealed by the X-ray crystallography, which structure shows that $HDM_2$ has a large hydrophobic cleft to bind to α-helix of p53 protein, involving three hydrophobic residues (Phe19, Trp23, and Leu26) in p53 [11], which enhances the stability of the binary complex.

Recently, several approaches have been developed to activate or restore the activity of p53, for example, the activation of wild-type p53 can be made through drugging its modification proteins, stimulating family member p63/p73, or disturbing interactions between $HDM_2$ and wild-type p53 [13]. And the latter case has allowed for successful identification of many $HDM_2$-p53 complex inhibitors for cancer treatment [14–16], such as the analogs of cis-imidazoline [14], spiro-oxindole [17], benzodiazepinedione [18], quilinol [19] and terphenyl derivatives [20].

By now, the most potent and well-characterized p53-$HDM_2$ inhibitors are nutlins identified by high-throughput screening method, and the second most potent series are benzodiazepinediones (BDPs) [21]. These inhibitors showed high binding affinity to p53-$HDM_2$ complex and desirable pharmacokinetic profiles in cells [22]. N-Acylpolyamine analogs (NAPA), another class of peptide inhibitors, have been proved as potent antagonists of the $HDM_2$-p53 interaction in vitro and in cell-based assays [18, 23]. NAPA shows characteristics of cyclic peptides in β-sheet conformations, and these helical β-peptides can be chemically modified to bind to $HDM_2$ with good affinity.

Clearly, in vitro assessment of the activity of p53-$HDM_2$ inhibitors remains a labor-intensive and time-consuming operation. Therefore, more efficient and economical alternative methods should be employed, such as in silico molecular modeling approaches that are used for the purpose of predicting and prioritizing chemicals for subsequent in vitro and in vivo screening. 3D-QSAR (quantitative structure-activity relationship) method, as effective computational method, has been widely used as a way to find out various interactive fields making impact on the activity and thus to help forecasting and designing of novel inhibitors [24–27, 33, 35], as well as the $HDM_2$-p53 inhibitors, including Isoindolinone [28], Terphenyls [20] and Chalcones Derivatives [29]. To our knowledge, there is still a lack of in silico modeling of these BDPs and NAPA analogs of the complex inhibitors.

BDPs and NAPA analogs have been reported as promising $HDM_2$-p53 interaction inhibitors [18, 30–32]. BDPs bound to $HDM_2$ in the p53-binding cavity in micromolar concentration, and the most potent compound 41 possessed very strong anti-$HDM_2$ activity with an $IC_{50}$ of 0.25 μM. Recent evidence showed that the structures of this class of inhibitors had four available sites, i.e., $R_1$, $R_2$, $R_3$ and $R_4$ shown in Table S1 for modifications, thus improving inhibitory activity and cell permeability for a candidate [18, 30, 31]. For peptide-based NAPA analogs, a synthesized single trimeric inhibitor showed very potent inhibition on $HDM_2$ protein, and the most potent compound 102 inhibits $HDM_2$ binding to p53 with $IC_{50}$ value of 2.3 μM [32]. Traditionally, large molecules would be difficult to produce in sufficient quantities for drug development; however, NAPA derivatives offer the possibility of using a short oligomeric scaffold to develop potent p53-$HDM_2$ inhibitors. Although these two classes of inhibitors are important for treatment of cancer, they have not received much attention from a theoretical perspective so far. Therefore, in this paper, these two classes of inhibitors were analyzed using theoretical computations, with purpose to build related 3D-QSAR models by comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) approaches [33]. In addition, molecular docking and molecular dynamics were also performed to further understand the structure properties and the probable binding modes of these inhibitors at the allosteric site of the receptors. Evidence has proved that this was the first time on 3D-QSAR modeling of these two classes of p53-$HDM_2$ inhibitors. All these methods applied on the inhibitors could not only help understanding ligand–receptor interactions but also provide useful and rational suggestions for further design of new drug candidates of p53-$HDM_2$ for cancer therapy.

## Methods and materials

### Data sets and biological activity

A total of 108 molecules with available $IC_{50}$ were collected from the references [18, 30–32], these molecules comprised two different classes of diverse 1,4-benzodiazepine-2,5-diones (BDPs) and N-Acylpolyamine (NAPA) derivatives, whose $IC_{50}$ values were converted into $pIC_{50}$ ($-\log IC_{50}$)

values (shown in Tables S1-S2) and which were used as dependent variables in 3D-QSAR analysis. Division of the training and test sets (Tables S1-S2 marked with [a]) was carried out by considering that the test compounds should represent the structural diversity and a range of biological activities similar to that of the training set. A11 molecular modeling calculations were performed using SYBYL package (Tripos Associates, St.Louis, MO), where Gasteiger–Marsili charges for BDPs and Gasteiger-Hückel charges for NAPA respectively were added and energy minimization for each molecule was performed by Tripos force field [34] with the Powell conjugate gradient minimization algorithm. The minimization wouldn't be finished until the energy gradient convergence criterion of 0.05 kcal $mol^{-1}$ $Å^{-1}$ was reached.

## Conformational sampling and alignment

For both the CoMFA and CoMSIA studies, the most important step of modeling is that all molecular structures are aligned to a suitable conformational template, which is assumed to be the most potent conformation. In this work, we applied two alignment methods, the first one of which denoted by ligand-based alignment selected the most potent inhibitor for each class (compounds 41, 102, respectively) as the template. All inhibitors for each class in the data set were then aligned to a common substructure (shown in Fig. 1) of the template using the "align database" command in SYBYL software to produce valid and reliable CoMFA and CoMSIA models. The second alignment method we employed was the receptor-based alignment, where all molecules were docked into the receptors and then the top scored conformations were further used directly for CoMFA and CoMSIA analysis. During this process, we not only explored the traditional docking alignment but also a novel receptor-guided consensus dynamics alignment for comparison of the docking difference between them. Finally, the conformations derived from the dynamics alignment, which is the optimal one we found, was applied for construction of the 3D-QSAR models. Figures 2 and 3 show the aligned conformations derived from the two methods.

## 3D-QSAR analysis

CoMFA and CoMSIA studies were carried out with the 3D cubic lattice with grid spacing of 2 Å in x, y and z directions, using an sp3 carbon probe atom with a van der Waals radius of 1.52 Å and a charge of +1.0. Cutoff values for both steric and electrostatic fields were set to 30 kcal $mol^{-1}$. The CoMSIA models were based on the molecular similarity indices with the same lattice box used for the CoMFA calculations. However, it had several advantages over CoMFA such as greater robustness regarding both region shifts and small shifts within the alignments [35]. Other than steric, electrostatic fields described in the CoMFA method, five physicochemical properties: steric (S), electrostatic (E), hydrophobic (H), hydrogen bond donor (D), and hydrogen bond acceptor (A) were evaluated using the standard settings: probe with charge +1, radius 1 Å and hydrophobicity +1, hydrogen-bond donating +1, hydrogen-bond accepting +1, attenuation factor of 0.3 and grid spacing 2 Å. CoMSIA similarity indices ($A_F$) for a molecule j with atoms $i$ at a grid point $q$ are calculated by Eq. 1:

$$A_{\mathrm{F,K}}{}^{q}(j) = -\sum \omega_{\mathrm{probe},k}\omega_{ik}e_{iq^2}{}^{-ar}. \quad (1)$$

In this equation, $k$ represents five physicochemical properties (S, E, H, D, and A), $W_{ik}$ is the actual value of physicochemical property $k$ of atom $i$, and $W_{probe,k}$ is the value of the probe atom. $\alpha$ is the attenuation factor and the default value of 0.3 was used. A Gaussian type distance dependence was used between the grid point $q$ and each atom $i$ of the molecule.

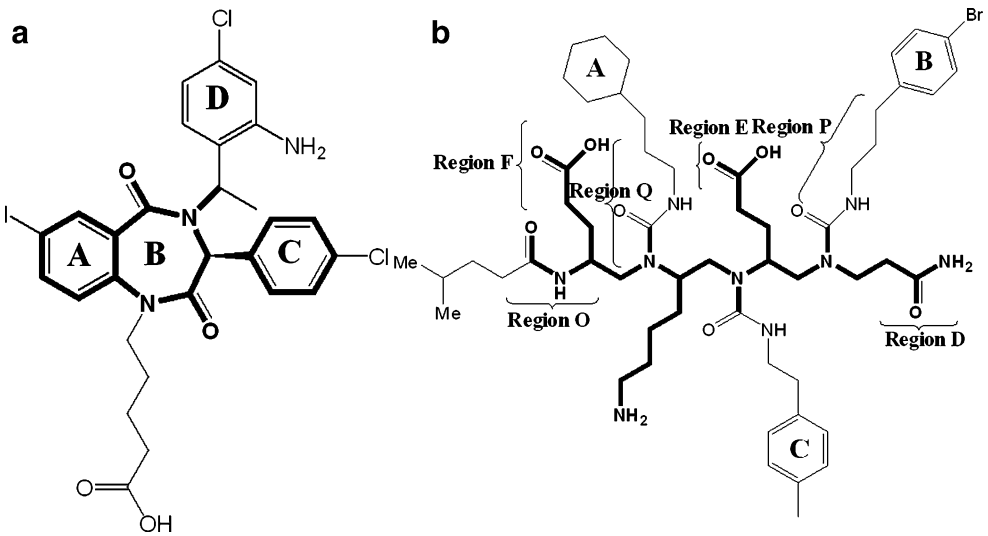


**Fig. 1** Representative skeletons and the most active molecules showing different regions which are used in contour analysis. (**a**) Compound 41 in the group of 1,4-benzodiazepine-2,5-diones analogs. (**b**) Compound 102 in the class of N-acylpolyamine analogs

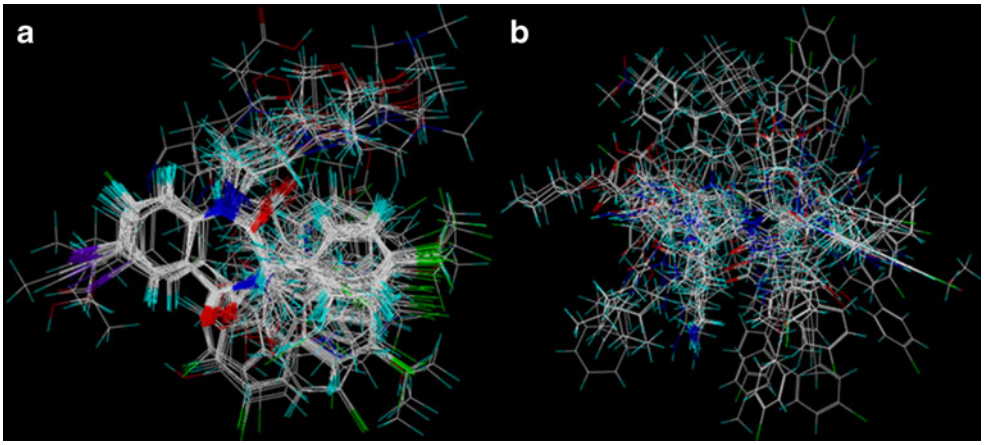


**Fig. 2** Ligand-based alignment. (**a**) Alignment of 1, 4-benzodiazepine-2, 5-diones derivatives with compound 41 selected as the template. (**b**) Alignment of N-acylpolyamine derivatives with compound 102 selected as a template

Partial least-square (PLS) method was often used for all 3D-QSAR analyses [36]. It was used to linearly correlate the CoMFA and CoMSIA descriptors to the inhibitory activity values. Cross-validation based on the training set was performed by the leave-one-out (LOO) procedure to determine the optimum number of components (ONC) and the coefficient $R^2_{cv}$. Then the ONC was further employed to do non-validation analysis to produce non-cross-validated correlation coefficient $R^2_{ncv}$. The molecules which were not included in the training set were used to evaluate the predictive ability of the models derived using the training set. The predictive correlation coefficient $r_{pred}^2$ was calculated using Eq. 2:

$$r_{\text{pred}}^2 = (SD - PRESS) / SD, \quad (2)$$

where SD is the sum of the squared deviation between the actual $pIC_{50}$ values of the compounds in the test set and the mean $pIC_{50}$ value of the training set ones. PRESS is the sum of squared deviation between predicted and actual activities of the test set compounds.

Molecular docking

To investigate the ligand-protein interactions, all the molecules were docked into the binding site of the receptor using the Surflex module in SYBYL package. During the docking procedure, standard parameters were employed. Crystal structures of $HDM_2$ (1T4E, 1RV1) were retrieved from RCSB Protein Data Bank (http://www.pdb.org). Before docking, all water molecules were removed from the X-ray structure, which is based on two reasons. 1) Normally, it is a standard procedure to remove waters from X-ray structures before docking, because in most cases, the studied compounds are different from the ligands in the crystal structure. Therefore, the interactions between the original ligands and waters would not be true for other different compounds. This phenomenon is also found in this work. For 1T4E, in the crystal structure, only one water molecule forms H-bond with the –NH of ring B for compound 19. However, this H-bonding interaction will disappear for compound 41 as the hydrogen atom of -NH was substituted by $–(CH_2)_4COOH$. 2) The MD simulations have been also performed to investigate if there are possible

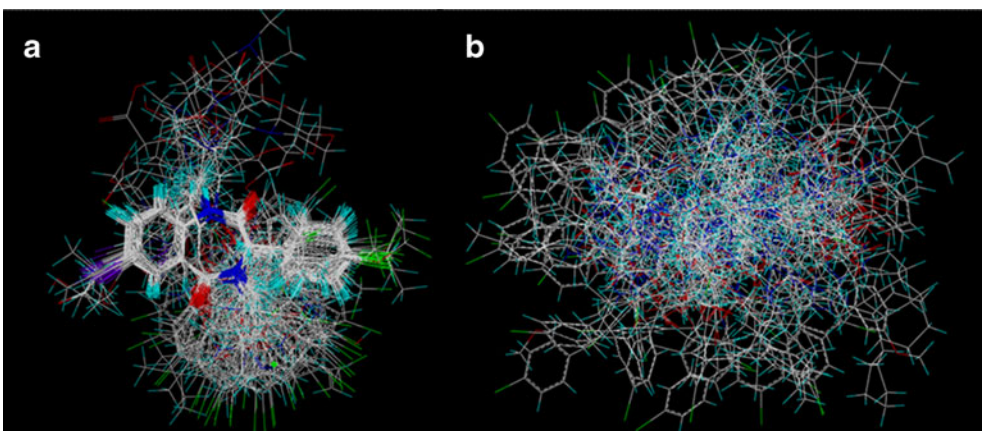


**Fig. 3** Receptor-guided consensus dynamics alignment. (**a**) Alignment of 1,4-benzodiazepine-2,5-diones derivatives with compound 41 selected as a template. (**b**) Alignment of N-acylpolyamine derivatives with compound 102 selected as a template

effects provided by water molecules. The MD results revealed that no waters were found H-bonded with the ligands as those in both 1RV1 and 1T4E structures, which further demonstrated that the deletion of waters before docking is reasonable. Subsequently, the relative ligands were extracted and polar hydrogen atoms were added. Protomol, served as an idealized representation of a ligand that made every potential interaction with the binding site, was used to guide molecular docking. During docking process, two parameters, i.e., protomol_bloat and protomol_threshold, were applied to determine how far from a potential ligand the site should extend and how deep into the protein, the atomic probes were used to define the protomol could penetrate. Ligand docking method was employed presently to generate the protomol, and finally each conformer of all inhibitors in two classes was docked into the binding site 10 times with the D_score [37], G_score [38], Chemscore [39] and PMF_score [40] values further used to evaluate the docking analysis. Finally, the top ranked conformations for each molecule were extracted, then aligned together and subsequently utilized in CoMFA and CoMSIA modeling.

### MD simulations

The MD simulations were performed with GROMACS software package [41] using the GROMOS96 force field [42]. The molecular topology files for compounds BDPs and NAPA were generated by the program PRODRG 2.5 [43–46], which are shown in the supporting information. For MD simulation, a cubic periodic box of side length of 67.44 Å and 64.70 Å was applied for each class of the inhibitors. The minimum distance between the protein and box walls was set to larger than 8 Å. Three $Cl^-$s were employed to neutralize the net charge of each system, i.e., 1T4E.pdb and 1RV1.pdb, respectively. The total number of the atoms was 28057 for 1T4E and 24574 for 1RV1 including the protein complexes and waters, respectively. The remaining box volume was filled using the extended simple point charge (SPCE) water [47].

Prior to the simulation, an energy minimization was applied to the full system without constraints using the steepest descent integrator for 5000 steps, then the system was equilibrated via a 200 ps MD simulations at 300 K. Finally, a 5 ns simulation was performed with a time step of 2 fs. During MD simulation, the standard parameters and main calculation methods were set as follows:

Both models (BDPs, NAPA) used NPT ensemble at 300 K with periodic boundary conditions, the temperature was kept constant by the Berendsen thermostat, the values of the isothermal compressibility were set to $4.5\times10^{-5}$ $bar^{-1}$ while the pressure was maintained at 1 bar using the Parrinello-Rahman scheme [48], electrostatic interactions were calculated using the particle mesh Ewald method [49], cut-off distances for the calculation of Coulomb and van der Waals interactions were 1.0 and 1.4 nm, respectively. All the MD simulations lasted 5 ns to ensure that the whole systems were stable.

## Results and discussion

### 3D-QSAR model

Table 1 summarizes the optimum results derived from the CoMFA models where the predictive ability of the models was assessed by predicting the inhibitory ability of the test set molecules. The developed 3D-QSAR models were analyzed by a number of parameters including the cross-validated correlation coefficient ($r^2_{cv}$), standard error estimate (SEE) and F-statistic values (F), etc. For CoMFA and CoMSIA analysis, two alignment methods were applied to produce the models, with results that the models derived from receptor-guided consensus dynamics alignment were much worse than those derived from ligand-based ones. Consequently, we mainly focused on the models derived by ligand-based method. S and E fields were generated for CoMFA models, while H, D and A, three other fields in addition to the S and E were generated for CoMSIA ones. Since it is argued that the five different descriptor fields may not be totally independent of each other and such dependencies of individual fields usually decrease the statistical significance of the models [50, 51], in the present work all 31 possible combinations of the descriptors for each group were attempted to build the optimum models with highest $R^2_{cv}$ values and other proper statistical results for each class. Figure 4 illustrates the observed versus predicted activities of the optimal CoMFA models for the two classes of dataset. As to the results of other combinations of CoMFA or CoMSIA descriptors, Tables S3 to S6 in the supporting information give a full summation.

#### *BDPs*

For this class of inhibitors, the data set was divided into a training set of 59 and a test set of 17 molecules with the statistical parameters obtained from the CoMFA model summarized in Table 1. The best predictions were obtained by the CoMFA standard model using five components, with statistical results of $R^2_{cv}$=0.41, $R^2_{ncv}$=0.855, F value =62.39 and a low SEE=0.26 for the training set obtained, which suggested that this model should be considerably reliable to predict the $IC_{50}$ values. The steric and electrostatic contributions were 38.7% and 61.3%, respectively, which indicated that the electrostatic field made more contribution to ligand binding affinity. Overall, this model showed reasonable statistical features. However, the CoMSIA model derived

**Table 1** The optimum 3D-QSAR results for the two classes of $HDM_2$ inhibitors

| Parameters | 1,4-benzodiazepine-2,5-diones | | | | N-acylpolyamine | | | |
|---|---|---|---|---|---|---|---|---|
| | Ligand-based | | Structure-based | | Ligand-based | | Structure-based | |
| | CoMFA | CoMSIA | CoMFA | CoMSIA | CoMFA | CoMSIA | CoMFA | CoMSIA |
| $R^2_{cv}$ | 0.405 | 0.506 | 0.101 | 0.06 | 0.414 | 0.474 | 0.07 | 0.12 |
| SEE | 0.259 | 0.244 | 0.38 | 0.096 | 0.184 | 0.097 | 0.087 | 0.059 |
| F | 62.388 | 51.394 | 56.289 | 257.36 | 34.862 | 75.878 | 413.34 | 919.321 |
| $R^2_{pred}$ | 0.6017 | 0.314 | 0.033 | 0.424 | 0.6913 | 0.5941 | 0.116 | 0.17 |
| SEP | 0.523 | 0.486 | 0.626 | 0.692 | 0.473 | 0.495 | 0.549 | 0.535 |
| Nc | 5 | 7 | 2 | 10 | 5 | 9 | 2 | 2 |
| Field contribution | | | | | | | | |
| S | 0.387 | - | 0.345 | 0.143 | 0.461 | - | 0.386 | 0.389 |
| E | 0.613 | - | 0.655 | 0.556 | 0.539 | 0.395 | 0.614 | - |
| H | - | 0.529 | - | - | - | 0.605 | - | - |
| D | - | 0.471 | - | - | - | - | - | 0.611 |
| A | - | - | - | 0.301 | - | - | - | - |

$R^2_{cv}$=Cross-validated correlation coefficient using leave-one-out method; SEE=Standard error of estimate; F=Ratio of $R^2_{ncv}$ explained to unexplained= $R_{ncv}{}^2/(1-R^2_{ncv})$; $R^2_{pred}$=Predicted correlation coefficient for the test set of compounds; SEP=Standard error of prediction; Nc=Optimal number of principal components; S=steric, E=electrostatic, H=hydrophobic, D=H-bond donor, A=H-bond acceptor.

from the same training set showed poor internal predictions ($R^2_{cv}$ =0.13) using the same S and E fields. Meanwhile, incorporation of H, D and A fields also could not improve the model performance. All these results suggested that the CoMFA model was superior to the CoMSIA one for this set of compounds.

These models were further validated using the external test compounds. The optimum CoMFA model gave good predictions exhibiting the large $R^2_{pred}$ value of 0.60. However, the CoMSIA models possessed poor predictions with low $R^2_{pred}$ value. The correlation between the predicted activities and the actual activities are displayed in Fig. 4a. From this figure, we can see that all points are uniformly distributed along the regression line which suggests that no systematic error exists in the method. Additionally, in the optimal CoMFA model, compounds 10 and 61 were regarded as outliers, since their prediction errors were larger than 1.2 standard deviation (SD). We also found that for compound 10, which possessed a low inhibitory activity and was the only molecule had a cyclohexyl substituent at the $R_2$ position. As for compound 61, its outlier status might be due to the low inhibitory activity, which was beyond the forecast range.

*NAPA*

Using both S and E fields with five components, a best CoMFA model for NAPA inhibitors was obtained, giving a good internal predictive ability illustrated by an $R^2_{cv}$ of 0.41, $R^2_{ncv}$ of 0.91, F value of 34.86 and a low SEE value of 0.18. The steric and electrostatic contributions were found to be 46.1% and 53.9%, respectively. Therefore, the electrostatic field had a greater effect than the steric field on the inhibitory activity, indicating that the electrostatic interactions of the molecules with the receptor could be an important factor for $HDM_2$ antagonistic activity. However, the CoMSIA model

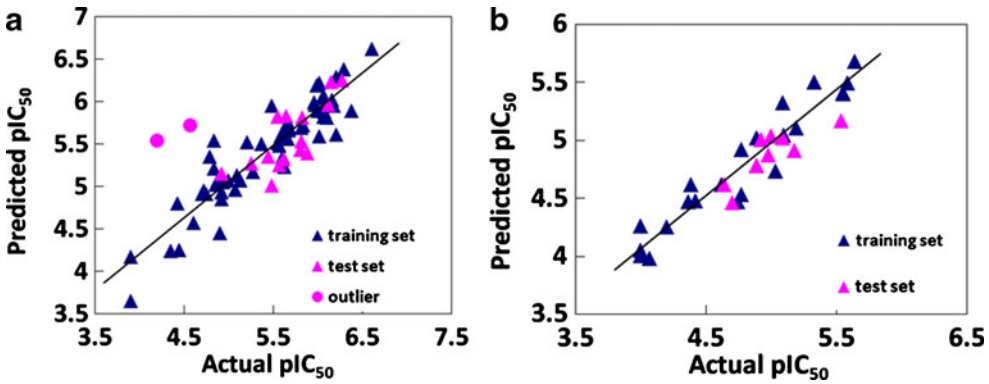


**Fig. 4** Graphs of the predicted $pIC_{50}$ versus the experimental $pIC_{50}$ values of the optimal CoMFA models. (**a**) CoMFA model of 1,4-benzodiazepine-2,5-diones. (**b**) CoMFA model of N-acylpolyamine derivatives

using the same fields (SE) exhibited a low $R^2_{cv}$ of 0.34, $R^2_{ncv}$ of 0.84, F value of 17.32 and SEE value of 0.25. Other combinations like SHE, SEHD, SEHA, SEDA, SEHDA were also performed to build models, ending in an average $R^2_{cv}$>0.4 and $R^2_{pred}$ >0.6 (shown in Table 1 and S4). Overall, the performance of the CoMFA model was superior to that of the CoMSIA one.

To test the predictive ability of the model, a test set of nine molecules excluded from the model derivation was employed. The predictive correlation coefficient $R^2_{pred}$ was 0.69. The graph of actual activity versus predicted $pIC_{50}$ of the training set and test set was illustrated in Fig. 4B, the plots represent a uniform distribution around the regression line with respective slope and intercept very close to one and zero, indicating the satisfactory predictive capability and accuracy of the model.

## 3D-QSAR contour maps

To visualize the results of the CoMFA and CoMSIA models, contour maps which denote the areas where the molecules would favorably or unfavorably interact with the receptor were generated. For steric field, the green contours represent areas where bulky group would enhance activity while the yellow contours represent areas where bulky group would decrease the activity. In the electrostatic field, the red contour represents the electronegative charge favorable area while the blue color represents the electropositive charge favorable area.

### *BDPs*

CoMFA steric and electrostatic contributions for this class of inhibitors with the most potent compound 41 as a reference were shown in Fig. 5. A small green contour at position 7 of ring A indicates that a sterically bulky group was favored in this region, which is coincide with compounds 29 and 30. Interestingly, compared with compound 33, compound 34 has a larger substituent at this position, leading to an increased inhibitory activity. Sterically unfavored yellow regions were found around ring C suggesting the need for small substituent at this region to enhance the inhibitory activity, for example, the activity of compound 19 (−Cl) was shown to be higher than that of compound 20 (−Br). There were two large green contours observed outside the ring C and around the yellow contours respectively, which indicates that if the substituents at the para position of ring C could touch these green contours, large groups would enhance the activity, which may exactly be the reason why compounds 10, 11, 12, 13 with small groups were less potent than molecules 14, 15, 16 with bulky substituents at the same location. Another yellow contour located between ring B and C, suggesting the desire of the position for groups with small steric barrier to increase the activity, such as compounds 57, 58, 59 and 60. Two yellow contours were also observed near $R_3$ substituent, indicating the requirement of less bulky substituents in this region for more potent p53-$HDM_2$ inhibitors.

Figure 5b shows the CoMFA electrostatic filed. A medium sized red isopleth at the para position of ring C represents an area where negative charge is favorable, for example, though all having a common structure, compounds 1 to 7 are slightly different at the para position of ring C, where molecules 1, 6, 7 possess negatively charged groups (−$CF_3$, -Cl, -$OCF_3$ respectively) and thus increased activity, while chemicals 2, 3, 4, 5 having positive charged substituents (−H, -$CH_3$, -$CH_2CH_3$, -$CH(CH_3)_2$ respectively) at the position and decreased activity. A large red contour around the negatively charged carboxyl group at the $R_3$ substituent indicates that the negatively charged carboxyl or other negative groups at this position were necessary to increase the activity. Another one big and two medium sized red contours at the -ortho, -meta position of ring D suggests that electronegative substituents would increase the activity, bringing us a conclusion that modifications such as incorporation of electronegative substituents in these areas may

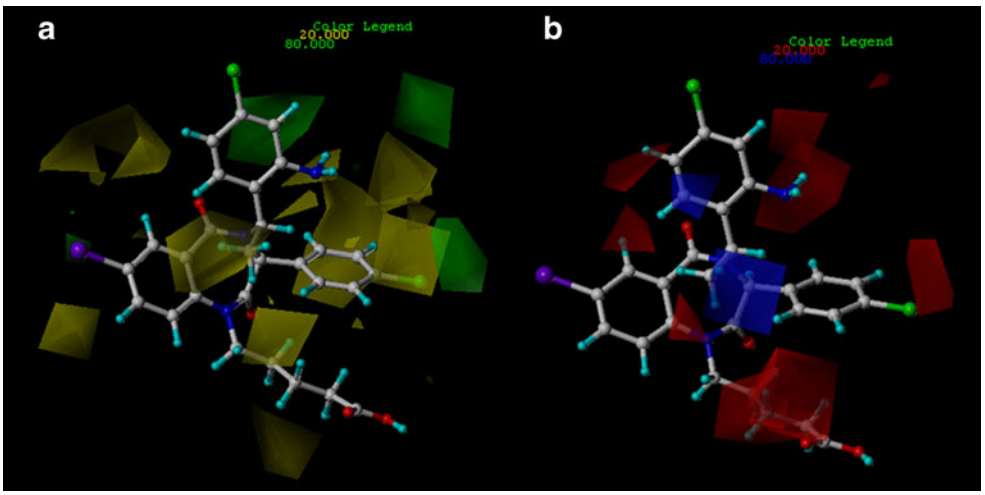


**Fig. 5** CoMFA StDev*Coeff contour plots for 1, 4-benzodiazepine-2,5-diones analogs in combination of compound 41. (**a**) The steric contour map, where the green and yellow contours represent 80% and 20% level contributions, respectively. (**b**) The electrostatic contour map, where the blue and red contours represent 80% and 20% level contributions, respectively

greatly improve the inhibitory activity. Meanwhile, a red and a blue contour simultaneously appeared at the linker area between rings B and D, i.e., the electropositive and electronegative contours emerged at the same region showing that a balance of these properties among the groups presented at this region is required for optimum binding.

*NAPA*

The CoMFA steric map (Fig. 6a) displays a medium sized green contour at the para position of ring B suggesting that bulky groups at this position would significantly improve the affinity. Thus, compound 102, which has –Br located at the green region, showed more potent inhibitory activity than those compounds of 101, 103 and 104 which possess rather smaller substitutent (F) at the same location. Another large green contour around ring A suggests that steric bulky substituents are favored there. A large region of yellow contour around Region P represents the disfavor of bulky group at the region, which is consistent with compounds 103 and 104.

The electrostatic contour map is displayed in Fig. 6b. Clearly, an electropositive favored contour was observed outside the para position of ring B, which is well illustrated by the decreased activity of compound 87 with an oxygen atom touching the blue contour. It is also validated by compound 85 which also exhibited lower activity than compound 102 by extending a substituent of –Cl to the blue contour. One big red contour in the vicinity of region O suggests that electronegative substituent could favor the activity, the fact was that all of the derivatives involved in the present study possessed electronegative substituents (−NH) at this site.

## Molecular dynamics simulations and comparison with 3D contour maps

Molecular dynamic simulations were performed to obtain the 'real' bioactive conformation which could not be acquired only from the crystal structure of protein–ligand complex [48]. During molecular dynamics simulation, the protein complex was placed in a bulk system which can mimic the real physiological conditions, which can fully consider the flexibility of the receptor, thus making the complex structure more reasonable. However, in the general molecular docking procedure, including the Surflex program, the receptor is normally treated rigidly to model the ligand-receptor binding mode. Therefore, in this work, in order to obtain the optimal conformation for complexes, the docked complex structures of 1T4E with compound 41, 1RV1 with 102 were treated as the initial structures for molecular dynamics simulations. The root-mean-square deviation (RMSDs) of the trajectory with respect to their initial structures ranged from 1.3 to 2.5 Å, 1.3 to 1.8 Å for 1T4E and 1RV1, respectively (Fig. 7a and Fig. 8a). The small RMSDs variations indicate the system of the complexes reached about 2.2 Å, 1.5 Å after 2 ns, and retained this value throughout the simulation.

The superposition of the average structures of the whole trajectory and the docked structures was shown in Fig. 7b and 8b. The comparisons of binding models from MD simulation and the QSAR maps are shown as follows:

*BDPs*

Figure 9b indicates that the main amino acid residues in the active site cavity responsible for significant interactions were Gln18, Gln24, Lys51, Gly16, Ser17, Leu54, Leu57, Val93, Phe55 and His73. The most potent compound 41 formed three H-bonds: 1) the keto oxygen of ring B (H-bond acceptor) hydrogen-bonded with the –NH group of Ser17 (−O···HN, 1.73 Å, 121.1°); 2) the carboxyl oxygen atom at $R_3$ substituent formed two H-bonds with the $–NH_2$ of Gln24 (−O···HN, 1.91 Å, 152°) and the $–NH_2$ of Gly16 (−O···HN, 2.58 Å, 140°). In addition, ring D of compound 41 formed π-π interaction with His73. As illustrated in Fig. 9b, the $R_3$ substituent was surrounded by electropositive Lys51 and

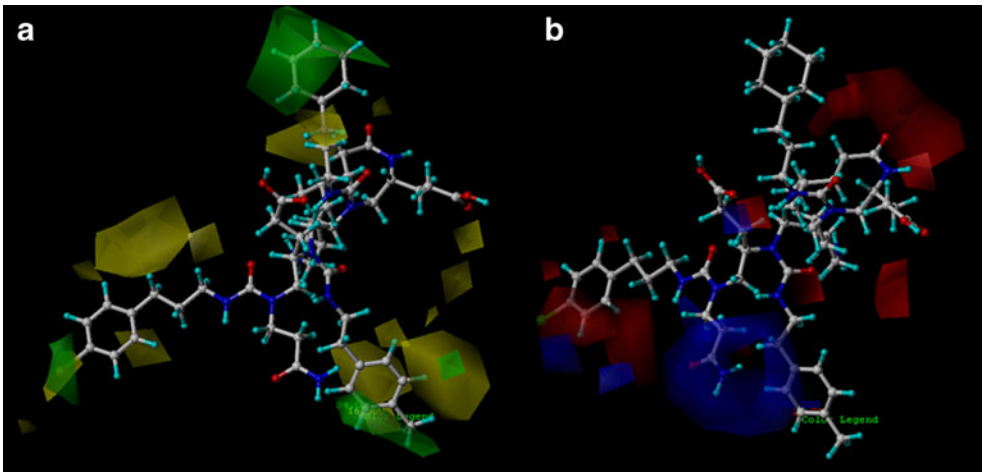


**Fig. 6** CoMFA StDev*Coeff contour plots for N-acylpolyamine analogs in combination of compound 102. (**a**) The steric contour map, where the green and yellow contours represent 84% and 16% level contributions, respectively. (**b**) The electrostatic contour map, where the blue and red contours represent 82% and 18% level contributions, respectively

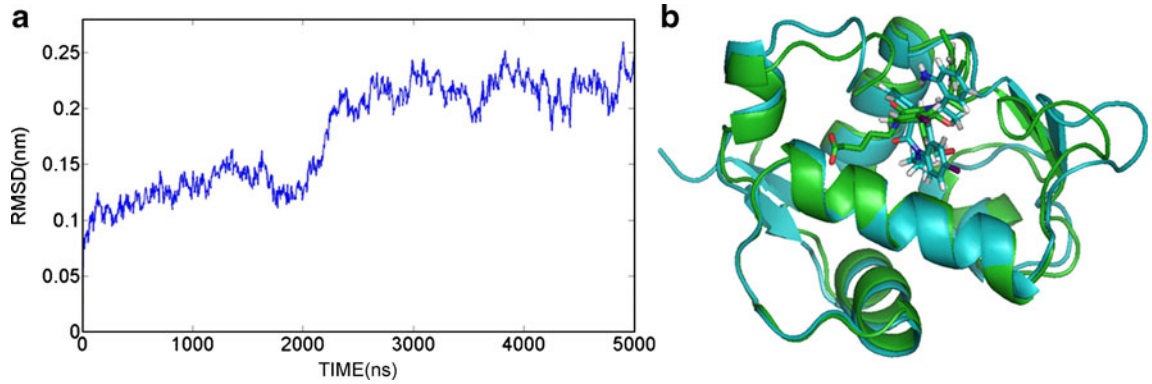


**Fig. 7** (**a**) Plot of the root-mean-square deviation (RMSD) of docked complex versus the MD simulation time in the MD-simulated structures. (**b**) View of superimposed backbone atoms of the lowest energy structure of the MD simulation (green) and the initial structure (cyan) for compound 41-$HDM_2$ complex. Compound 41 is represented as carbon-chain in cyan for the initial complex and carbon-chain in green for the lowest energy complex

Gln24 indicating that electronegative groups at this position are favorable for the inhibitory activity, which result matches well with the electrostatic contour map of the CoMFA model (depicted in Fig. 5b). In addition, the MD-simulated structures of the binding site shows that residues His73, Leu54, Val93 appeared near $R_1$ substituent, where electronegative substituents would favor electrostatic interaction between ligands and $HDM_2$. This result is well consistent the electrostatic contour map of the CoMFA model (Fig. 5b), which owns a red contour at the same location. It is worthy to note that near ring C there is sufficient room to accommodate a large hydrophobic substituent into the hydrophobic pocket formed by Leu54, Leu57, Gly58, Gln72 and Val93. This observation is in full agreement with the steric contour map with a green contour (Fig. 5a) around this region. In addition, several yellow contours were observed at this position suggesting that too large substituents would make clash effects on the residues.

Additionally, $R_1$ substituent was surrounded by some non-polar amino acids Leu54, Leu57, Ile61 and Ile99, suggesting that non-polar groups at this position is favorable for the binding affinity, this is consistent with all compounds employed in this work possessing a non-polar benzene ring in this region. In addition, the $R_3$ substituent is surrounded by polar residues Gln24 and Lys36, therefore, a relatively polar substituent at this position is needed. This is in agreement with the fact that the binding affinities of compound 41 with polar substituent –COOH is more active than compounds 43, 44 and 45 which possess non-polar $–CH_3$ groups in this position.

*NAPA*

Figure 8b shows the conformation derived from compound 102 with the allosteric binding site of 1RV1, where compound 102 was suitably localized at the binding site. The MD-based docking simulation showed that Glu25, Thr26, Tyr100, Tyr104, Met50, Lys94, His96 and Arg97 were the main residues presented at the binding site and main contributors to the ligand-receptor interaction (shown

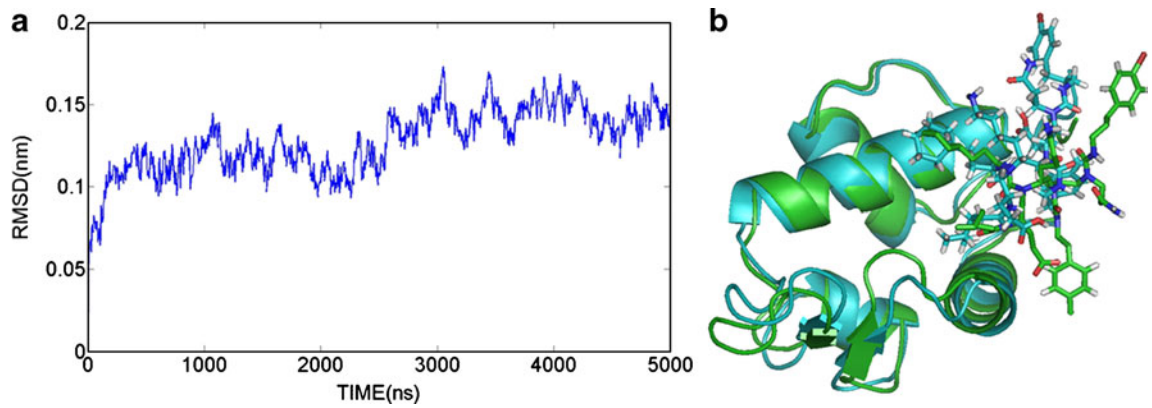


**Fig. 8** (**a**) Plot of the root-mean-square deviation (RMSD) of docked complex versus the MD simulation time in the MD-simulated structures. (**b**) View of superimposed backbone atoms of the lowest energy structure of the MD simulation (green) and the initial structure (cyan) for compound 102-$HDM_2$ complex. Compound 102 is represented as carbon-chain in cyan for the initial complex and carbon-chain in green for the lowest energy complex

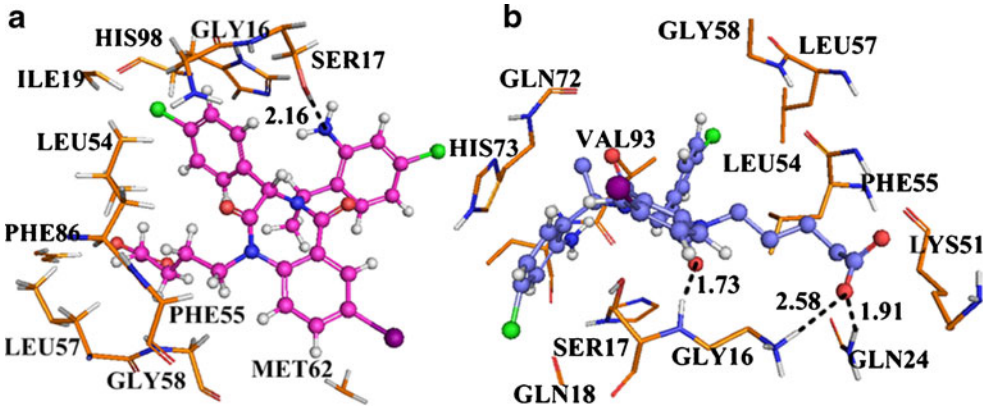


**Fig. 9** (**a**) Docked conformation derived for compound 41 in complex to the active site of $HDM_2$ protein. (**b**) Plot of the MD-simulated structures of the binding site with ligand 41. H-bonds are shown as dotted black lines, active site amino acid residues are represented as sticks, the inhibitors are shown as stick and ball model

in Fig. 10b). And a total of nine hydrogen bonds were formed in this system. The carboxyl oxygen at region Q acted as an acceptor to form H-bond with the side chain hydroxyl group of Tyr100 (−O···HO, 1.42 Å, 175.6°), the carboxyl oxygen at region F formed two H-bonds with the side chain NH of Arg97 (−O···HN, 1.95 Å, 139.6°) and (−O···HN, 1.76 Å, 153.8°). In addition, the carboxyl oxygen at region E interacted through H-bonding with the backbone NH of Glu25 (−O···HN, 1.81 Å, 171.8°) and the OH of Thr26 (−O···HO, 1.52 Å, 161.9°), also the hydroxyl oxygen atom formed two H-bonds with Glu25 (−O···HN, 2.36 Å, 114.9°), (−O···HN, 2.74 Å, 90.7°). The carboxyl oxygen at region P was involved in hydrogen bonding with the backbone $NH_2$ of Glu25 (−O···HN, 2.57 Å, 89.4°), (−O···HN, 2.64 Å, 85.6°). As displayed in Fig. 10b, several basic electropositive residues Arg97, His96 were close to region O which suggests that electronegative substituents at these positions would enhance the inhibitory activities. This could be confirmed by the electrostatic contour map (Fig. 6b) which holds a red contour at this position. A large binding cavity formed by residues Lys94, His96, Arg97 and Thr100 over ring A could contain a bulky substituent, which is corroborated by the contour analysis results of the green contour obtained from the CoMFA model (Fig. 6a). Meanwhile, ring C established arene-cation interaction with Arg97 which further enhanced the stability of the ligand-receptor interaction.

Moreover, the polar substituent –COOH of region E extends to polar amino acids residues Glu25 and Thr26, indicating that polar groups are favorable for the inhibitory activity. This is the same as the –COOH of region F which is surrounded by Ary97 and His96. It is remarkable to note that the non-polar amino acid residues Phe55 and Leu54 are near ring A, which indicates that the non-polarity of ring A is needed for the inhibitory activity.

## Comparison of the results of docking and MD simulation

When predicting the binding mode of the ligand to receptor, molecular docking provides a good starting to evaluate the stability of the predicted interactions involved in binding [51]. While molecular dynamics was employed for further calculations considering the flexibility of the receptor and the effect of water solvation. In this study, firstly the most potent compounds (41 and 102) were docked into the binding site of the receptors, which complexes were further used to undertake MD simulations. As depicted in Fig. 7b and Fig. 8b, the initial conformations changed a lot after the MD simulations. Furthermore, we further examined the stability of the hydrogen bond formed during the docking process by calculating the presence rates of hydrogen bonds

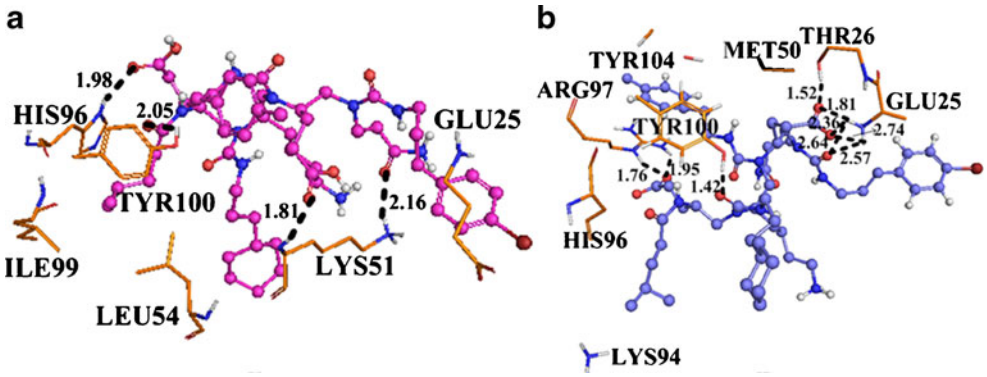


**Fig. 10** (**a**) Docked conformation derived for compound 102 in complex to the active site of $HDM_2$ protein. (**b**) Plot of the MD-simulated structures of the binding site with ligand 102. H-bonds are shown as dotted black lines, active site amino acid residues are represented as sticks, the inhibitors are shown as stick and ball model

during MD simulation. The hydrogen bonds produced between all possible donors D and acceptors A. A geometrical criterion shown in Eq. 3 is used to determine if the hydrogen-bond exist.

$$\begin{aligned} r &\leq r_{HB} = 0.35\,nm \\ \alpha &\leq \alpha_{HB} = 30^\circ \end{aligned} \quad (3)$$

As seen from Fig. 9 and Fig. 10, it is concluded that the presence of corresponding H-bonds between the ligands and receptors exhibited higher frequencies than the docking analysis.

For BDPs analogs, one hydrogen bond was formed based on docking simulation between the amino nitrogen atom in ring D and the hydroxyl group of Ser17 (−N···HO, 2.16 Å, 177.1°), which was yet not preserved during MD simulation. It was interesting to note that three new hydrogen bonds were constructed to enhance the stability of the complex.

As for the lost hydrogen bond, it could be compensated by the electrostatic interactions. As illustrated in Fig. 9b, the electronegative nitrogen atom was surrounded by electropositive residues His73 and Gln18 which could still be involved in electrostatic interactions.

During docking simulation, four hydrogen bonds were formed in compound 102 and 1RV1 (Fig. 10a), i.e., region O oxygen atom…OH of Tyr100 (−O···HO, 2.05 Å, 108.8°), region F oxygen atom…NH of His96 (−O···HN, 1.98 Å, 158.2°), region E oxygen atom…NH of Lys51 (−O···HN, 1.81 Å, 143.8°), region D oxygen atom…NH of Lys51 (−O···HN, 2.16 Å, 147.7°). However, after MD simulations, only two hydrogen bonds were preserved, and the amino acids involved in forming the H-bonds were also changed as explained above. In a word, comparatively the molecular dynamics simulation provides a more stable conformation, so we chose the MD-simulated structures for further 3D-QSAR studies.

### Comparison of binding modes for each class

In order to explore the differences and similarities between the two classes of p53-$HDM_2$ inhibitors, we compared the binding mode of each class. The CoMFA steric field in the BDPs model makes more contributions to the binding activity, as well as the NAPA derivatives. In addition, the docked models reveal that the hydrogen bonding is also important for interactions between the ligand and the target. For BDPs, three H-bonds were formed between the most potent compound 41 and Ser17 (−O···HN, 1.73 Å, 121.1°), Gln24 (−O···HN, 1.91 Å, 152°), Gly16 (−O···HN, 2.58 Å, 140°), at the same time, compound 41 also formed π-π interaction with His73. The NAPA derivative 102 formed a total of nine hydrogen bonds, mainly through interacting with Tyr100, Arg97, Glu25 and Thr26. Additionally, arene-cation interaction has been established between Arg97 and this compound. The NAPA derivatives formed more hydrogen bonds than the BDPs analogs which also demonstrate why they were more active than BDPs. In addition, from the comparison of interacting patterns for the two series of inhibitors, we also found that there were no common amino acids which were involved in forming H-bonds among these two classes. Thus a conclusion can be made that these inhibitors might have different binding modes with p53-$HDM_2$.

## Conclusions

In this study, 3D-QSAR models for these two different classes of p53-$HDM_2$ inhibitors, i.e., BDPs and NAPA analogs were built using both the ligand-based and receptor-guided alignments to find out the structural relationship with the activity. We also compared two alignment schemes, with respect to the internal predictive ability and the robustness of their respective models for structurally diverse data sets. As a result, using the ligand-based alignment was superior to that based on the receptor alignment and the CoMFA models were superior to the CoMSIA ones. In addition to the structural features of stetic and electrostatic fields, this article still elucidated the structural features of hydrogen bond related to the $HDM_2$ activity. The external test validation results indicate that the CoMFA models could be successfully used for predicting the inhibitory efficiency of these two classes of compounds. The optimal models from the BDPs and NAPA analogs implied the significant roles of electrostatic field effects on their binding affinity. For the BDP model, the bulky substituent at ring C played a main contribution to inhibitory activity. The crucial role of $R_3$ group was revealed by the fact that it formed strong hydrogen bond with its receptor and the more potent compounds had bulky groups at the para position of ring A, small and electronegative groups around ring D. In addition, the non-polarity of $R_1$ and the polarity of $R_3$ had important effect on the binding affinity. For the NAPA model, the electropositive and bulky substituents at ring B had a major impact on the inhibitory activity, and hydrogen-bond interactions also improve the compound potency. For the more active compounds, they possessed bulky and electropositive groups at ring B, large and non-polar groups at ring A and electronegative substituents in region O.

In addition, molecular dynamics simulations were also performed for the docking complexes. The analysis revealed that the steric, electrostatic, and H-bonding interactions between ligands and key amino-acid residues in binding pocket of $HDM_2$ correlated well with CoMFA analysis and data of the inhibitory activity. The binding modes showed that for BDPs, the key residues were Gln24,

Lys51, Gly16, Ser17, Leu54, Gln72, Gly58, Leu57, Val93, Phe55 and His73. Yet in the NAPA-protein complex, those key residues changed to Glu25, Thr26, Tyr100, Tyr104, His96, Leu54, Met50 and Phe55. Further analysis showed that the CoMFA contour maps had a good correlation with the docking analysis after MD simulations, further proving the reliability of the models and promoting the understanding of the ligand-receptor interactions.

In this study, the CoMFA analyses gained some insights into the key structural factors affecting the bioactivity of these inhibitors. The excellent predictive ability of the developed CoMFA models indicated that they could be used for predicting the $IC_{50}$ values of this two class of inhibitors. Furthermore, the CoMFA contour maps along with the docking results after MD simulation offered enough information to understand the structure–activity relationship and identify structural features influencing the inhibitory activity. Overall, the correlation of the results obtained from 3D-QSAR, molecular dynamics simulation and molecular docking studies could, we hope, be served as a useful guideline for further modification and designing of new compounds as p53-$HDM_2$ inhibitors.

**Acknowledgments** The research is supported by high-performance computing platform of Northwest A & F University. The research is financially supported by the Fund of Northwest A & F University. The authors are grateful to Prof. L. Yang for access of Sybyl software.

## References

1. Lane DP (1992) Nature 358:15–16
2. Gudkov AV, Komarova EA (2007) Hum Mol Genet 16:R67–72
3. Komarova EA, Gudkov AV (2001) Biochem Pharmacol 62:657–667
4. Vogelstein B, Lane D, Levine AJ (2000) Nature 408:307–310
5. Feki A, Irminger-Finger I (2004) Rev Oncol Hematol 52:103–116
6. Kubbutat MH, Jones SN, Vousden KH (1997) Nature 387:299–303
7. Kussie PH et al (1996) Science 274:948–953
8. Wu X, Bayle JH, Olson D, Levine AJ (1993) Genes Dev 7:1126–1132
9. Barak Y, Juven T, Haffner R (1993) OrenM. EMBO J 12:461–468
10. Momand J, Zambetti GP, Olson DC, George D, Levine AJ (1992) Cell 69:1237–1245
11. Kussie PH, Gorina S, Marechal V, Elenbaas B, Moreau J, Levine AJ et al (1996) Science 274:948–953
12. Michael D, Oren M (2003) Cancer Biol 13:49–58
13. Chen F, Wang W, Wafik S (2010) El-Deiry. Biochem Pharmacol 80:724–730
14. Vassilev LT (2004) Cell Cycle 3:419–421
15. Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, Filipovic Z et al. (2004) Science 303:844–848
16. Brown CJ, Lain S, Verma CS, Fersht AR, Lane DP (2009) Nat Rev Cancer 9:862–873
17. Ding K, Lu Y, Nikolovska-Coleska Z, Qiu S, Ding Y, Gao W, Stuckey J, Krajewski K, Roller PP, Tomita Y, Parrish DA, Deschamps JR, Wang S (2005) J Am Chem Soc 127:10130–10131
18. Parks DJ, Lafrance LV, Calvo RR, Milkiewicz KL, Gupta V, Lattanze J, Ramachandren K, Carver TE, Petrella EC, Cummings MD, Maguire D, Grasberger BL, Lu T (2005) Bioorg Med Chem Lett 15:765–770
19. Lu Y, Nikolovska-Coleska Z, Fang X, Gao W, Shangary S, Qiu S, Qin D, Wang S (2006) J Med Chem 49:3759–3762
20. Chen L, Yin H, Farooqi B, Sebti S, Hamilton AD, Chen J (2005) Mol Cancer Ther 4:1019–1025
21. Lu F, Chi SW, Kim DH, Han KH, Kuntz ID, Guy RK (2006) J Comb Chem 8:315–325
22. Shangary S, Wang S (2008) Clin Cancer Res 14:5318–5324
23. Koblish HK, Zhao S, Franks CF, Donatelli RR, LaFrance LV, Leonard KA, Gushue JM, Parks DJ, Calvo RR, Milkiewicz KL, Marugan JJ, Cummings Raboisson P, MD Grasberger BL, Lu T, Molloy CJ, Maroney AC (2006) Mol Cancer Ther 5:160–169
24. Li Y, Wang YH, Yang L, Zhang SW, Liu CH (2006) J Mol Des 5:1–12
25. Li Y, Wang YH, Yang L, Zhang SW, Liu CH, Yang SL (2005) J Mol Struct 733:111–118
26. Wang X, Yang W, Xu X, Zhang H, Li Y, Wang Y (2010) Curr Med Chem 17:2788–2803
27. Wei SP, Ji ZQ, Zhang HX, Zhang JW, Wang YH, Wu WJ (2010) J Mol Model. doi:10.1007/s00894-010-0765-x
28. Hardcastle IR, Ahmed SU, Atkins H, Farnie G, Golding BT, Griffin RJ, Guyenne S, Hutton C, Källblad P, Kemp SJ, Kitching MS, Newell DR, Norbedo S, Northen JS, Reid RJ, Saravanan K, Willems HM, Lunec J (2006) J Med Chem 49:6209–6221
29. Stoll R, Renner C, Hansen S, Palme S, Klein C, Belling A, Zeslawski W, Kamionka M, Rehm T, Mühlhahn P, Schumacher R, Hesse F, Kaluza B, Voelter W, Engh RA, Holak TA (2001) Biochemistry 40:336
30. Parks DJ, LaFrance LV, Calvo RR, Milkiewicz KL, Gupta V (2006) Bioorg Med Chem Lett 16:3310–3314
31. Leonard K, Marugan JJ, Raboisson P, Calvo R, Gushue JM (2006) Bioorg Med Chem Lett 16:3463–3468
32. Hayashi R, Wang D, Hara T, Iera J, Durell SR, Appella DH (2009) Bioorg Med Chem 17:7884–7893
33. Klebe G, Abraham U, Mietzner T (1994) J Med Chem 37:4130–4146
34. Matthew C, Richard DC III, Van Nicole O (1989) J Comput Chem 10:982–1012
35. Pirhadi S, Ghasemi JB (2010) Europ J Med Chem 1–7
36. Wold S, Albano C, Dunn WJ, Edlund U, Esbenson K, Geladi P, Hellberg S, Lindburg W, Sjostrom M (1984) Multivariate data analysis in chemistry. In: Kowalski BR (ed) Chemometrics: MaThematics and Statistics in Chemistry, vol 138, NATO ASI Series- Reidel. Dordrecht, Holland, pp 17–96
37. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE (1982) J Mol Biol 161:269–288
38. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) J Mol Biol 267:727–748
39. Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP (1997) J Comput Aided Mol Des 11:425–445
40. Muegge I, Martin YC (1999) J Med Chem 42:791–804
41. van der Spoel D, van Buuren AR, Tieleman DP, Berendsen HJC (1996) J Biomol NMR 8:229–238
42. Lindahl E, Hess B, van der Spoel D (2001) J Mol Model 7:306–317
43. van Aalten DMF, Bywater R, Findlay JBC, Hendlich M, Hooft RWW, Vriend G (1996) J Comput Aided Mol Des 10:255–262
44. Barreca ML, Ortuso F, Iraci N (2007) Biochem Biophys Res Commun 363:554–560
45. Liu R, Li X, Li Y, Jin P, Qin W, Qi J (2009) Biosens Bioelectron 25:629–634
46. Niu C, Xu Y, Xu Y, Luo X (2005) J Phys Chem B 109:23730–23738
47. Berendsen HJC, Grigerra JR, Straatsma TPJ (1987) Phys Chem 91:6269–6271
48. Parrinello M, Rahman A (1981) J Appl Phys 52:7182–7190
49. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) J Chem Phys 103:8577–8593
50. Bohm M, Stürzebecher J, Klebe G (1999) J Med Chem 42:458–477
51. Bringmann G, Rummey CJ (2003) J Chem Inf Comput Sci 43:304–316