

***In silico* design of novel benzohydroxamate-based compounds as inhibitors for histone deacetylase 6 based on 3D-QSAR, molecular docking, and molecular dynamics simulations**

Chu Han^{a,b}, He Qing-xiu^{a,b}, Wang Juan^{a,b}, Hu Yong^{a,b}, Wang Yuan-qiang^{*a,b} and Lin Zhi-hua^{*a,b}

Received 00th

January 20xx,

Accepted 00th

January 20xx

DOI:

10.1039/x0xx00000x

Histone deacetylase 6 (HDAC6) can deacetylate many substrates, including α -tubulin, cortactin, and HSP90 α . This makes HDAC6 an interesting, emerging target. In this study, we selected 52 benzohydroxamate-based compounds and conducted 3D-QSAR studies using CoMFA and CoMSIA models. The results showed that CoMFA ($n = 9$; $q^2 = 0.556$; $r^2 = 0.998$) and CoMSIA ($n = 9$; $q^2 = 0.656$; $r^2 = 0.996$) have robust stability and predictability. The counter maps of the steric field, hydrophobic field, and hydrogen-bond donor field revealed the modified regions of these compounds. Subsequently, molecular docking was used to explore the docking mode of compounds and receptors. Based on these results, we designed 11 new benzohydroxamate-based compounds and predicted their activity and ADME/T properties. The results indicated that these compounds have strong pharmaceutical properties. MD simulation analysis confirmed that Asp742 and His610 in the active site played a key role in this. Overall, these results serve as an important reference for the discovery and design of new HDAC6 inhibitors.

1 Introduction

Metal-dependent histone deacetylases (HDACs) catalyze the hydrolysis of acetyllysine side chains to regulate protein function and participate in a variety of biological processes¹⁻⁵, including metabolism^{6, 7}, cell signal transduction⁸, and epigenetics^{9, 10}. To date, 18 HDACs have been identified in mammals and can be divided into two major categories: zinc-dependent enzymes and NAD⁺-dependent enzymes^{11, 12}. Human histone deacetylase 6 (HDAC6) is composed of a single polypeptide chain of 1215 amino acids and is the largest member of the zinc-dependent HDAC family^{13, 14}. At the same time, HDAC6 is the only isoform with two catalytic domains and a c-terminal zinc finger domain that can bind to ubiquitinated proteins¹⁵. HDAC6 protein is mainly localized in the cytoplasm and can deacetylate many substrates; thus, it involves many cellular processes, such as the following¹⁶: a. HDAC6 affects the formation of immune synapses, cell migration, and chemotaxis by acetylating α -tubulin^{17, 18}; b. HDAC6 influences the regulation of cell migration and actin filament binding through acetylating cortactin¹⁹; and c. HDAC6 impacts the degradation and elimination of misfolded proteins as well as the regulation of glucocorticoid receptors via acetylating HSP90 α ¹⁹. As such, HDAC6 is an interesting and emerging target. Its expression and functional changes are associated with various ailments, such as autoimmune disorders and neurodegenerative diseases²⁰⁻²². In addition, clinical data indicate that HDAC6 is overexpressed in many types of cancer²³⁻²⁵. Inhibition of HDAC6 has significant anticancer properties relevant to prostate cancer²⁶, breast cancer²⁷, melanoma²⁸, and ovarian cancer²⁹. Therefore,

selective inhibition of HDAC6 has become a very attractive cancer treatment strategy.

Currently, four broad-spectrum HDAC inhibitors (romidespin, vorinostat, belinostat, and panobinostat) have been approved for clinical use by the FDA³⁰, though selective HDAC6 inhibitors are not yet mature. HPOB has specific selectivity for HDAC6 and can enhance the anticancer activity of doxorubicin and etoposide³¹. ACY-241 (citarinostat) as well as ACY-1215 (ricolinostat) are hydroxamic acid derivatives. They exhibit specific inhibitory activity against HDAC6 and are currently the only HDAC6 inhibitors in clinical trials^{32, 33}. However, these selective HDAC6 inhibitors are fraught with stability problems or non-optimal pharmaceutical properties. Therefore, it is still necessary to find novel HDAC6 inhibitors. Here, we collected 52 benzohydroxamate-based HDAC6 inhibitors and obtained stable 3D-QSAR models through comparative molecular field analysis (CoMFA) and comparative molecular similarity index analysis (CoMSIA). Then, molecular docking was used to analyze the binding mode of HDAC6 to these compounds. Finally, we designed 11 new benzohydroxamate-based compounds and verified their interactions with HDAC6 through molecular dynamics simulations. Ultimately, this study aims to provide a theoretical basis for the discovery of new benzohydroxamate-based HDAC6 inhibitors.

2 Materials and methods

2.1 Preparation of the data sets

First, 52 benzohydroxamate-based compounds were gathered from the literature³⁴ and divided into training (41 compounds) and test data sets (11 compounds, represented by the symbol "*" superscript) randomly. The bioactivity of these compounds *in vitro* were reported as IC₅₀ values, which were converted to the corresponding pIC₅₀ ($pIC_{50} = -\log IC_{50}$). The relevant properties of the compounds are listed in supplementary material, Table S1.

^a Department of Pharmacy and Bioengineering, Chongqing University of Technology, Chongqing, 400054, P. R. China;

^b Key Laboratory of Screening and Activity Evaluation of Targeted Drugs, Chongqing, 400054, P. R. China;

* Corresponding Author

Email: wangyqnn@cqut.edu.cn; zhlin@cqut.edu.cn

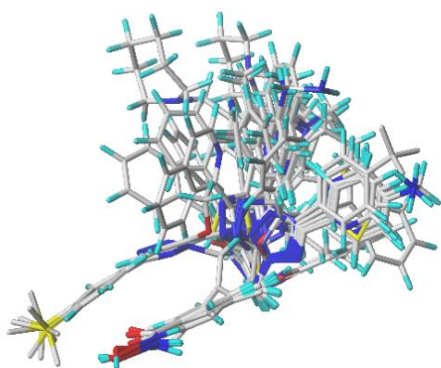


Figure 1. Alignment of all benzohydroxamate-based compounds

2.2 Structural Optimization and alignment

We employed SYBYL2.0 to construct the structure of all 52 compounds, and then added a Gasteiger-Huckel charge as well as a Tripos force field to each. Powell's energy gradient method was used to optimize the structure of these compounds. The maximum optimization limit was 10,000 times, the convergence standard was 0.005 kJ/mol, and the remaining parameters were set as default³⁵. Confirmation of all the treated compounds ensured the active conformation for the next step. At the same time, compound 26 (as shown in Figure S1, the red part represents the common skeleton) had the highest inhibitory activity and could be utilized as an alignment template. The alignment of all compounds is seen in Figure 1.

2.3 3D-QSAR modelling and validation

CoMFA analyses the differences in the molecular fields around a group of similar compounds to determine the non-interacting properties of them with receptors, thereby predicting their biological activity³⁶. Electrostatic and steric fields of CoMFA makes use of Coulomb and Lennard-Jones potential functions. In addition, CoMSIA also introduced a Gaussian distance function that calculates the similarity index between the probe and each atom of the molecule through the hydrophobic as well as hydrogen bond acceptor and donor fields³⁷. Through the leave-one-out (LOO) cross-validation procedure, the optimum number of components (n) with the highest cross-validation correlation coefficient (q^2) of the relevant model are calculated³⁸. The conventional multiple correlation coefficient (r^2), standard error of estimate (SEE), and Fisher test (F) values were obtained by non-cross-validation analysis³⁹. The counter maps of the 3D-QSAR model were employed to evaluate the relationship between molecular structure and biological activity, thereby guiding the design and modification of compound structures.

Furthermore, Y-randomization of the response test⁴⁰ was utilized to ensure the robustness of constructed models. With this test, we repeated the 3D-QSAR calculation procedure several times through randomly shuffling activities of the training set. The relatively low q^2 and r^2 values were yielded with randomized activities of the training sets, indicating that

the constructed models using the current modelling method for the collected data sets were acceptable and reliable⁴¹.

2.4 Molecular docking

Molecular modeling is widely used to illustrate the potential interaction modes of activity inhibitors and protein targets. In this work, the molecular docking simulation is performed by the Surflex-dock module of SYBYL X-2.0 (Tripos, LP). The X-ray co-crystal structure of the HDAC6 complex (PDB code: 5EDU, resolution 2.8) was retrieved from the RCSB protein database and used in the following research⁴². First, extract the original ligand TSN (Trichostatin A, a known class I and II HDAC inhibitor⁴³) from the active site of 5EDU, and then re-dock it into the binding pocket of the protein to evaluate the docking reliability of the generated Surflex-dock solution (<2.0 Å)⁴⁴. Then, we repaired the deletions and terminal residues of the polypeptide chain deleted water, and added hydrogen atoms to prepare proteins as receptors for docking models.

In the Surflex-Dock assay, general parameters including additional starting conformations per molecule, angstroms to expand search grid, max conformations per fragment, and max number of rotatable bonds per molecule were set as 0, 6, 20, and 100, respectively. To obtain the best binding conformation of each molecule, the maximum number of poses for each ligand is set to 20. In addition, the co-crystallized ligand TSN was set as a reference molecule during the docking simulation, and 0.5 nm amino acid residues in its range were used as the interface pocket area. The remaining parameters were default values. The Surflex-Dock scoring function is a weighted sum of non-linear functions involving van der Waals surface distances between the appropriate pairs of exposed protein and ligand atoms. It includes the following items: hydrophobic, polar, entropic, repulsive, solvation, and crash. The ligand-receptor complex with the highest total score was considered as the most stable binding conformation. On the basis of the docking score, based on the understanding of the HDAC6 binding pocket, we also ensured that the oxygen atom in the ligand hydroxamate group of the selected research system directly binds to zinc. The initial conformation selected based on this criterion is used for further research⁴⁵⁻⁴⁷. Finally, apply the above parameters to dock all compounds (Figure 2). The combined visualization is performed using PyMOL software.

2.5 Absorption, distribution, metabolism, and excretion – toxicity(ADME/T) properties prediction

In this study, using the pkCSM online server (<http://biosig.unimelb.edu.au/pkcsml/prediction>), the ADME/T characteristics of the designed HDAC6 inhibitors can be predicted. For a given compound, the web server predicts the percentage that will be absorbed through the human intestinal tract. Molecules with an absorbance less than 30% are considered to be poorly absorbed. The higher the volume of distribution (VD), the more the drug is distributed in tissues rather than plasma. If it is lower than 0.71 L/kg ($\log VD_{ss} < -0.15$), the steady-state distribution (VD_{ss}) is considered low; if it is higher than 2.81 L/kg, the VD_{ss} is considered high ($\log VD_{ss} > 0.45$)⁴⁸. The predictors evaluated the designed molecule to establish whether it may be metabolized by P450 or maybe a

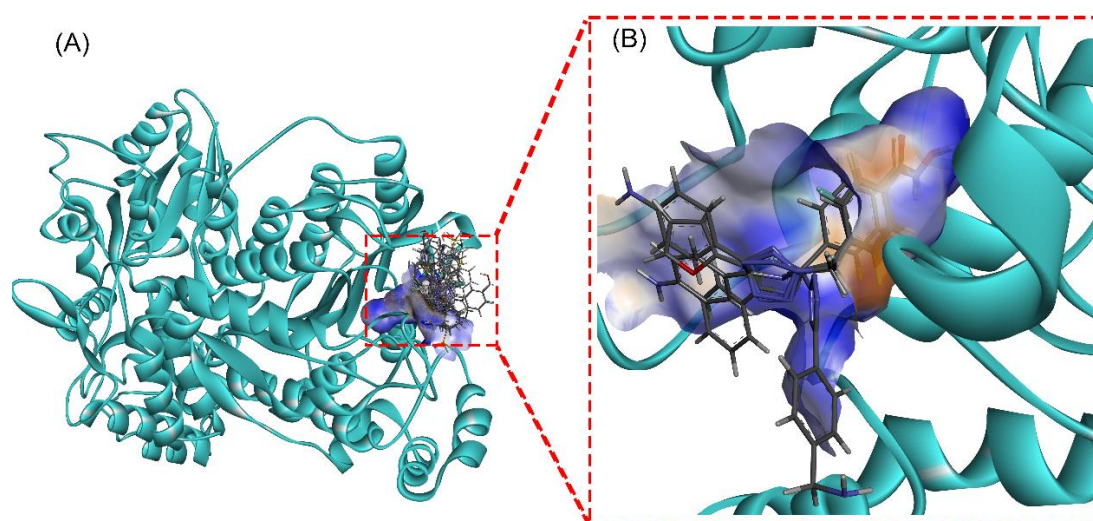


Figure 2. The binding site of HDAC6 with benzohydroxamate-based compounds. (A) All the compounds. (B) The top five chemical structures. (compound 26, 31, 9, 47, and 38). The surfaces of the potential pocket in HDAC6 are presented in orange and blue.

Table 1. Summary of 3D-QSAR results

Method	LOO		PLS			Contribution of force fields				
	<i>n</i>	<i>q</i> ²	<i>r</i> ²	<i>SEE</i>	<i>F</i>	<i>S</i>	<i>E</i>	<i>H</i>	<i>D</i>	<i>A</i>
CoMFA	9	0.556	0.998	0.050	1451	0.619	0.381	-	-	-
CoMSIA	9	0.656	0.996	0.061	961	0.191	0.2466	0.166	0.193	0.204

Note: *n*: The optimum number of components; *S*: Steric; *E*: Electrostatic; *H*: Hydrophobic; *D*: Hydrogen-bond donor; *A*: Hydrogen-bond acceptor; *F*: the Fischer's ratio; *SEE*: the standard deviation of the regression; *q*²: the highest cross-validation correlation coefficient; *r*²: The conventional multiple correlation coefficient.

cytochrome P450 inhibitor. The total clearance rate (log mL/(min • kg)) is used to determine the dose rate to reach steady-state concentration, which is related to bioavailability. Predictive evaluation of AMES toxicity/hepatotoxicity/skin sensitization was intended to predict whether the designed new molecule was toxic.

The online tool, SwissADME (<http://www.swissadme.ch/index.php>), was used to evaluate the synthetic availability of new derivatives⁴⁹. The scale of the tool indicates that a score close to 10 has high structural complexity and is therefore difficult to synthesize. A score of 1 indicates that the synthesis route is relatively easy.

2.6 Molecular dynamics simulation and free-energy analysis

Like most HDAC inhibitors, the compounds in this study contain a hydroxamate group that binds to zinc in the active site of HDAC6. Therefore, we used the MCPB.py module of the AMBER16 software package for modelling to establish the topology and coordinate files necessary for MD simulation⁵⁰. Before simulation, molecular mechanics methods are usually applied to optimize biological macromolecule systems. The GAFF force field is used for small molecules and the ff99 force field for proteins. Then, for example, the complexes were placed into 0.15 M NaCl solution within a cubic water box, which included 62160 water molecules, 172 Na⁺ ions, and 168 Cl⁻

ions. The initial configurations of the receptors and ligands were taken from docking studies. Waters (TIP3PBOX model) were added at a radius within 12 Å from the molecular surface of the complex. The other systems were set up with the same protocol. Next, the MD simulations were carried out using the PMEMD.mpi and PMEMD.cuda modules in AMBER16^{51, 52}. First, five minimization steps were performed on the system obtained by MCPB.py to avoid possible steric crashes. Then, gradually, heating each system from 0 K to 300 K took place during the heating stage and it was maintained at 300 K during the subsequent equilibrium and production stages. A time step of 2 fs was employed during the heating stage, the equilibrium stage, and the entire production stage. Periodic boundary conditions are employed to maintain constant temperatures and pressure ensembles. The pressure was set to 1 atm and controlled by an anisotropic (x-, y-, z-) pressure-scaling protocol, and the pressure relaxation time is 1 ps. The temperature was adjusted with Langevin dynamics with a collision frequency of 2 ps⁻¹^{53, 54}. The particle mesh Ewald (PME) method^{55, 56} was used to deal with long-range electrostatics, and a 10 Å cut-off value was set to deal with real-space interactions. All covalent bonds involving hydrogen atoms were constrained by the SHAKE algorithm⁵⁷. Each system underwent a 100-ns MD simulation,

and the trajectory of the simulated system was saved every 100 ps.

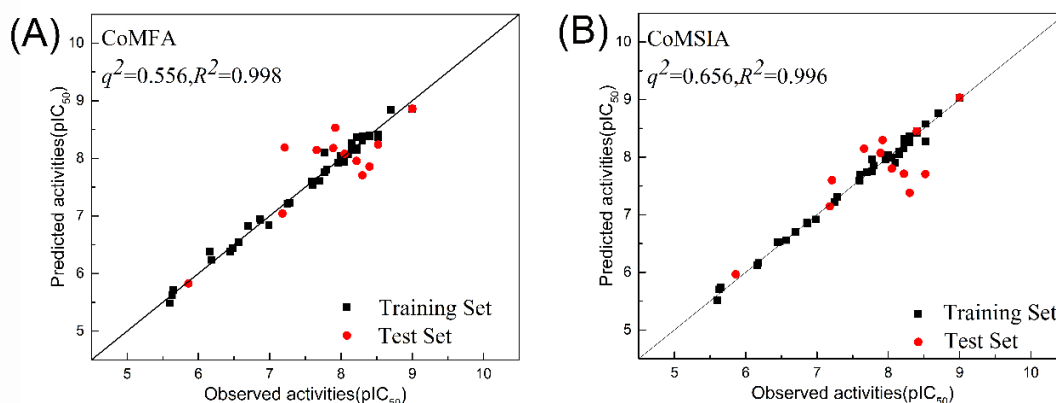


Figure 3. Correlation between training and test data sets' observed value data and predicted value data. (A) CoMFA; and (B) CoMSIA.

For the saved MD trajectories, the MM/GBSA method^{58, 59} was utilized to calculate the binding energy between different ligands and receptor residues. The energy of each residue was decomposed into main chain and side chain atoms. A total of 200 snapshots were taken from 80 to 100 ns. The energy decomposition could be analyzed to determine the contribution of key residues to binding⁶⁰.

3 Results and analysis

3.1 3D-QSAR modelling and validation

CoMFA and CoMSIA modelling results are shown in Table 1 (31 possible combinations of CoMSIA fields are shown in supplementary material, Table S2). The stability of the CoMFA and CoMSIA models can be reflected by the q^2 obtained by cross-validation. In general, when q^2 is greater than 0.3, the constructed model is only statistically significant at the 5% level, and when q^2 is greater than 0.5, the model has significant statistical significance. The CoMFA model constructed by the steric field and electrostatic field is the best, and the q^2 of LOO is 0.556, indicating that the model has robust stability and predictability. In addition, the n of the CoMFA model is 9, r^2 is 0.998, F value is 1451 ($F > 100$ is considered reliable), and SEE is 0.050, indicating that the model has strong statistical significance. In the CoMSIA model, among 31 possible field combinations, the combination of the steric field, electrostatic field, hydrophobic field, and hydrogen-bond donor and acceptor fields performed well ($q^2 = 0.656$, $r^2 = 0.996$). At the same time, the remaining parameters also demonstrate this ($F = 961$, $SEE = 0.061$). The linear correlation analysis of the training and test data sets of the 3D-QSAR model is found in Figure 3. Most of the compounds in the training and test data sets in the CoMFA and CoMSIA models are located at or near the trend line, proving that the actual activity values of the compounds fit well with the predicted activity values (indicated

by pIC_{50}). Moreover, the low q^2 and r^2 values obtained using Y-randomization test indicated that the robust results of our original models were not because of the accidental correlation or structural dependence of the training set (Supplementary Material, Table S5).

3.2 Counter-map analysis of CoMFA and CoMSIA

Compound 26 with the highest actual activity was used as a template for counter-map analysis of CoMFA and CoMSIA. The CoMFA steric field, CoMSIA hydrogen-bond donor field, and hydrophobic field play an important role in guiding the structural modification process of compounds. Selected counter maps are displayed in Figure 4.

Figure 4A features a steric-field contour map of CoMFA. The combination of green (80% contribution) and yellow (20% contribution) regions indicates that substitution with a bulky group at the five-membered ring in the middle of the compound can enhance activity. Figure 4B was a hydrophobic-field contour map of CoMSIA. The white area indicates that adding a hydrophilic group here is beneficial to increase the activity of the compound, and the yellow area reflects that it is better to replace it with a hydrophobic group. Therefore, the $-CH_2$ group linking the five-membered ring and benzene ring can be used as an important region for modification. The hydrogen-bond donor-field contour map of CoMSIA is shown in Figure 4C. The blue-green region indicates that increasing the hydrogen-bond donor here is beneficial to activity. Therefore, more hydrogen-bond donors are needed for the $-NH_2$ group on the benzene ring. This has been confirmed by certain compounds within the system. In the case of compounds 26 and 30, compound 26 (pIC_{50} value of 9.00) was substituted with $-NH_2$, which provided more hydrogen-bond donors, and its activity was significantly higher than $-CF_3$ -substituted compound 30 (pIC_{50} value of 7.96).

3.3 Analysis of molecular docking

In order to explore the binding mode of HDAC6 and ligands, the Surflex-Dock module of SYBYL software was employed for molecular docking. Before docking all the compounds, the original ligand TSN of HDAC6 was re-docked into the binding

pocket to verify whether the method and parameters we used were reliable and feasible. As shown in Figure S1

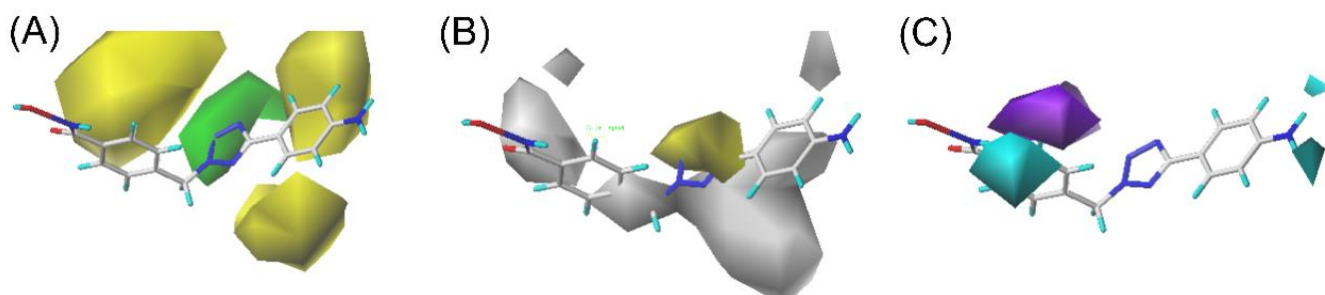


Figure 4. Three-dimensional counter maps of CoMFA and CoMSIA (compound 26). (A) CoMFA steric field; (B) CoMSIA hydrophobic field; (C) CoMSIA hydrogen-bond donor field.

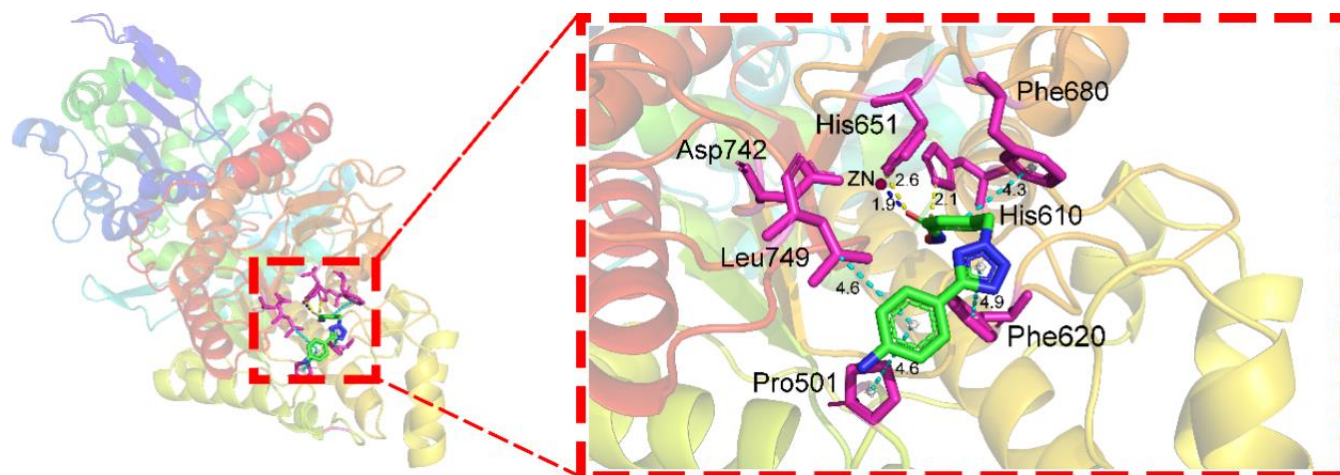


Figure 5. Docking model of HDAC6 (SEDU⁴²) with compound 26. The hydrogen bonds are shown as yellow dashed lines, hydrophobic bonds as light-blue dashed lines, and ZN-O binding as dark-blue dashed lines. Key residues are shown in purple and compound 26 is shown in green.

(Supplementary Material), the re-docked configuration of the TSN molecule is basically the same as its initial configuration, and the RMSD of the two configurations is 1.2 Å (< 2.0 Å). The hydroxyl oxygen atom in the compound directly binds to zinc, and this characteristic interaction is consistent with previous studies³⁴. This also shows that the generated docking protocol is reliable and can be used for subsequent research. All compounds were then docked into the binding site using the same docking parameters as in Figure S3 (Supplementary Material) while the total scores of all compounds are located in supplementary material, Table S1. We selected compound 26 as a template for analyzing binding mode and displayed it with a two-dimensional diagram (Supplementary Material, Figure S2) and three-dimensional diagram (Figure 5). First, the oxygen atom on the hydroxamate group of compound 26 forms a ZN-O bond (1.9 Å) with zinc in the active site. Interestingly, this oxygen atom also forms a hydrogen bond (2.6 Å) with His651. This combination is very stable. It is additionally worth noting that the hydroxamate group also forms two hydrogen bonds with His610 and Gly619. In addition, there is a rich hydrophobic

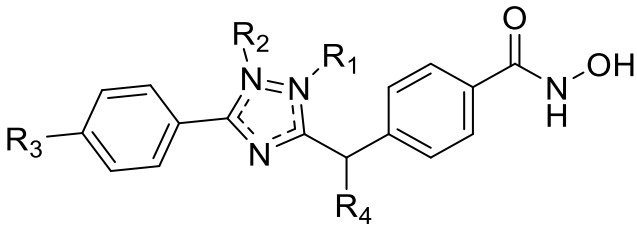



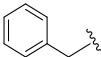
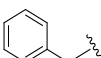
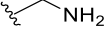
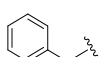

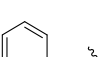

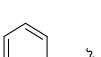
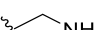
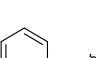
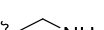
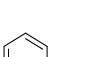
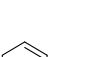

interaction between compound 26 and HDAC6. A Pi-Pi interaction exists between Phe680 and the benzene ring containing the hydroxamate group, Phe620, and a five-membered ring. There are also hydrophobic interactions between Pro749, Pro501, and benzene rings containing -NH₂ in the compounds. Rich hydrophobic interactions and hydrogen bonds play a key role in the binding of ligands to receptors. They can make the binding of compounds to receptors more stable. In Figure S2, the light-green amino acid residue without attachment indicates van der Waals interaction between the residue and ligand. Among these residues that interact with the ligand to generate van der Waals interactions, His500, Asp742, and Tyr782 contributed higher binding free energy, especially Asp742. This is consistent with the results of free-energy decomposition.

3.4 Molecular design and activity prediction

By analyzing the 3D-QSAR counter map and the results of molecular docking, we identified four modified regions of the compounds (Table 2). Compound 26 is the most active

compound and was selected as the lead compound for the new compound design (compound 26, CoMFA=8.86, CoMSIA=9.03,

Table 2. Structures and predicted activities of the newly designed compounds

								
No.	R ₁	R ₂	R ₃	R ₄	Predicted activity(pIC ₅₀)		Score	Synthetic accessibility
					CoMFA	CoMSIA		
26a			-NH ₂	-H	7.89	8.02	8.52	3.24
26b			-NH ₂	-H	7.52	8.20	8.14	3.27
26c			-NH ₂	-H	7.24	7.18	7.98	3.29
26d			-NH ₂	-H	8.75	8.95	8.91	3.32
26e				-H	8.85	9.13	9.49	3.48
26f				-H	8.69	8.63	9.62	3.51
26g				-H	8.85	9.01	7.94	3.44
26h				-Cl	9.34	9.23	8.03	4.01
26i				-F	9.32	9.36	8.51	4.05
26j			-NH ₂	-F	9.41	9.29	8.15	3.89
26k				-F	9.47	9.32	7.70	4.01

Docking Score=7.97, Synthetic accessibility=2.79). The analysis results show that one side of the five-membered ring (R₁ or R₂)

of the lead compound needs to introduce a larger group. Based on this, we tried to construct compounds 26a-26d. The results

showed that the introduction of a benzyl group at the R1 position significantly improved the docking score of the compound. Next, a hydrogen bond donor group was introduced into the R3 position of the lead compound (compound 26e-26g). We found that the docking score and CoMSIA of compound 26e were significantly improved. Next, we modified the R4 position of the lead compound (compound 26h and compound 26i), and the results showed that the introduction of -F greatly increased the values of compound CoMFA and CoMSIA. From the perspective of the docking score, compound 26i was significantly better than compound 26h. Finally, we introduced a benzyl group at the R1 position and -F at the R4 position and modified the R3 region (compound 26j-26k) again. The CoMFA and CoMSIA of the two increased slightly, but the docking score dropped significantly. In the end, we determined that 26i was the optimal compound for subsequent analysis. The docking scores of the above 11 compounds and the CoMFA and CoMSIA values predicted by the 3D-QSAR model are listed in Table 2 (the slimes of these compounds are listed in Table S3 of supplementary materials).

3.5 Molecular dynamics simulations and binding free-energy analysis

We selected compound 26i with a higher docking score and greater predicted activity, and compound 26 with the highest

actual activity value. Using the complex formed by docking them with HDAC6 as the initial conformation, a 100 ns MD simulation was performed to determine the stability of the complex and further understand the detailed dynamic combination mode. The overall system convergence and stability of the MD simulation were monitored by the root mean-square deviation (RMSD) of the skeleton atoms (C, Ca, N, and O) relative to the initial docking structure (Figures 7A, 7B). From Figure 6A, it can be seen that the complex is stable after 30 ns, the RMSD of compound 26 fluctuates around 2.2 Å, and the RMSD of the acceptor main-chain atoms fluctuates around 2.4 Å. At roughly 80 ns, the RMSD of the acceptor backbone atoms drop to 1.5 Å. Yet, at 90 ns, the RMSD returned to around 2.4 Å and remained stable. This indicates that compound 26 has a flip in 80-90ns. But then stabilized. In Figure 6B, the RMSD of HDAC6/26i fluctuates around 2.0 Å and 2.2 Å, respectively, after 50 ns. At 80 ns, the RMSD of compound 26i fluctuated, but then recovered. At this stage, HDAC6 has a partial flip, but it does not affect the stability of the entire system. These results indicate that compounds 26/26i and the acceptor backbone atoms undergo minor conformational changes during the simulation process, thus achieving a stable conformation. At the same time, we showed the initial and final complex structures of two (26/26i). (Supplementary Material, Figure S5) to verify the reliability of our MD simulation.

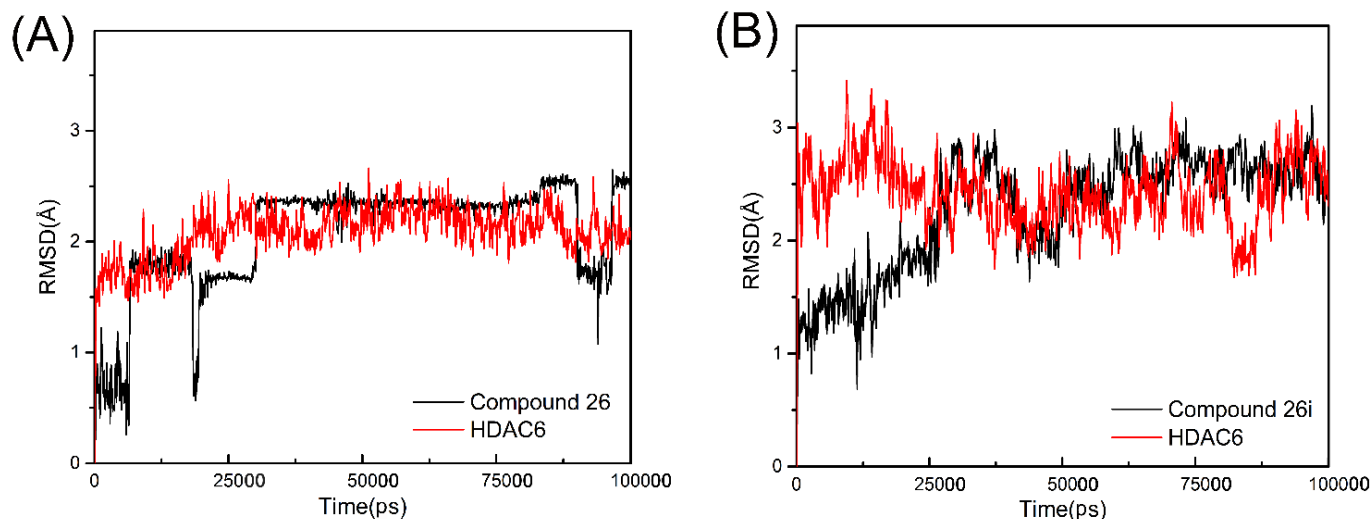


Figure 6. RMSDs of complex HDAC6/compounds (26, 26i) for 100-ns MD simulations.

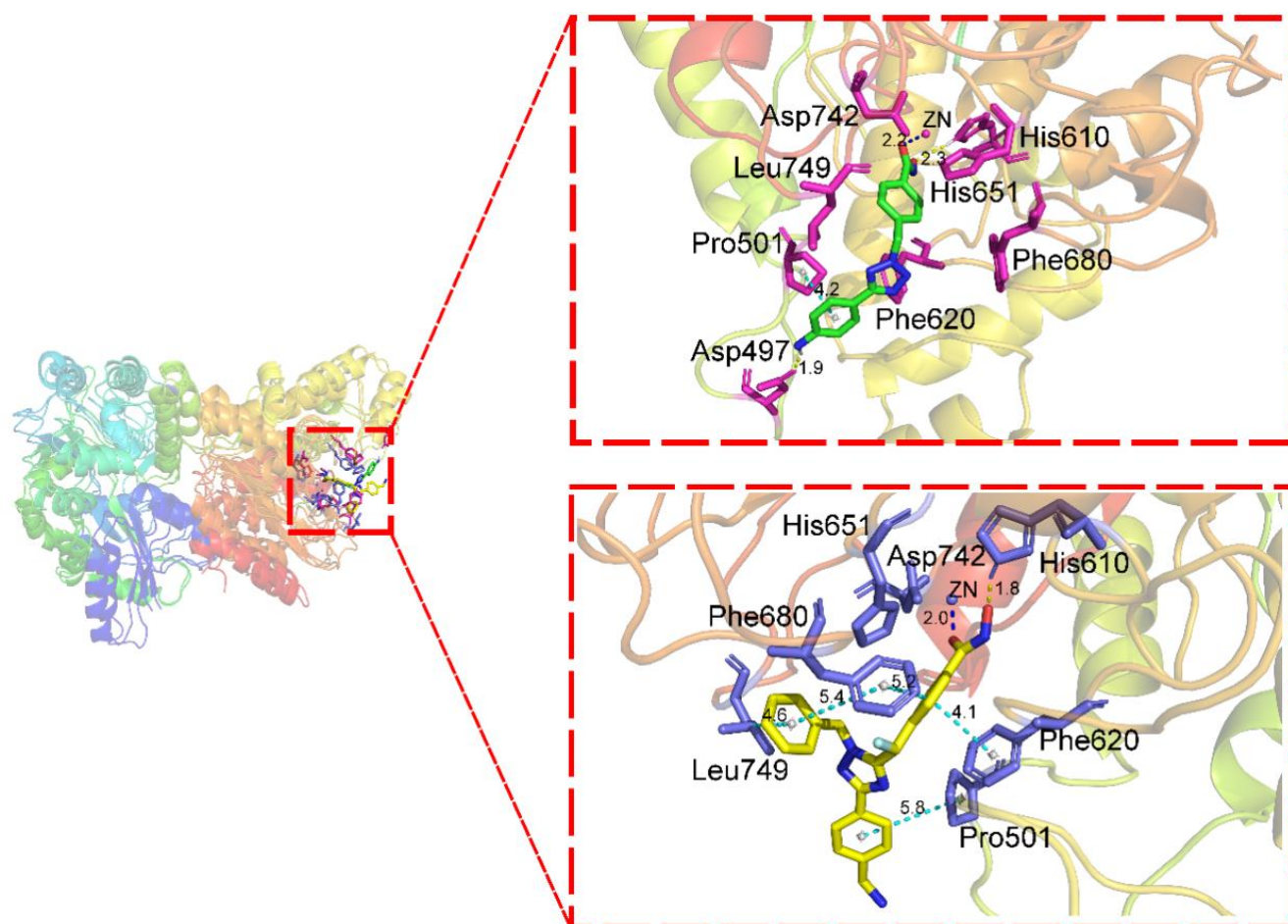


Figure 7. Analysis of the binding mode of ligand molecules to HDAC6. (A) Interaction between compound 26 and HDAC6 after 100-ns MD. (B) Interaction between compound 26i and HDAC6 after 100-ns MD. Hydrogen bonds are depicted as yellow dashed lines; hydrophobic bonds are depicted as light-blue dashed lines; Zn-O bonds depicted as dark-blue dashed lines.

Next, we assessed in depth the interactions between compound 26/26i and key residues of HDAC6. The stable conformation of the two complexes after MD simulation is portrayed in Figure 7. As seen in Figure 7A, the oxygen atom on the hydroxamate group of compound 26 forms a Zn-O bond (2.2 Å) with the zinc of the active site, and at the same time, another oxygen atom on the hydroxamate group forms a hydrogen bond with His610 with a bond length of 2.3 Å. In addition, Asp497 and the -NH₂ group on the benzene ring also form a hydrogen bond with a bond length of 1.9 Å. The benzene ring and Pro501 form a weak hydrophobic bond with a distance of 4.2 Å. In Figure 7B, like compound 26, the hydroxamate group forms a hydrogen bond and a Zn-O bond with His610 and Zn, respectively (bond lengths are 1.8 Å and 2.0 Å, respectively). In addition, the benzene ring containing the hydroxamate group forms two weak hydrophobic interactions with Phe680 and Phe620 (bond lengths are 5.2 Å and 4.1 Å, respectively). The benzene ring containing -NH₂ and Pro501 form a weak hydrophobic bond with a bond length of 5.8 Å. In addition to the aforementioned

two benzene rings, another benzene ring also forms weak hydrophobic interactions with Phe680 (5.4 Å) and Leu749 (4.6 Å) through van der Waals forces.

To evaluate the interaction between key residues of HDAC6 and compound 26/26i, we calculated the interaction energy between key residues of HDAC6 and compound 26/26i using the MM/GBSA method. The results are shown in Figure 8 (data is listed in Supplementary Table S4). From the overall results, the interaction energy between Asp742 and compound 26/26i is lower than -8.0 kcal/mol. In addition to van der Waals forces contributions and electrostatic interactions, the contribution of polar solvents is even greater. The interaction energy of His610 and compound 26/26i reduced the energy of 6.0 kcal/mol for the whole system, among which the electrostatic interaction contributed the most. In addition, His651, Gly780, Pro501, Leu749, Phe620, Asp497, Tyr782, His500, and Phe680 also have a significant effect on binding affinity, rendering the entire system more stable.

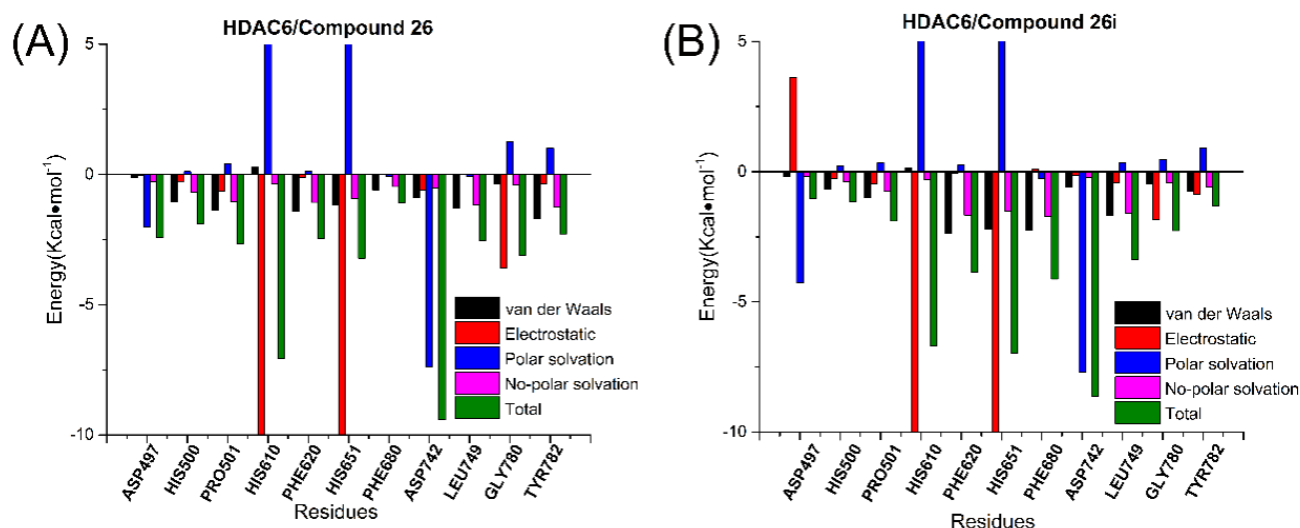


Figure 8. Contributions of free energy calculated by the MM/GBSA method for key residues of HDAC6.

Table 3. ADME/T properties of the novel designed molecules

No.	Absorption	Distribution	Metabolism CYP							Excretion		Toxicity	
	Intestinal	VDss (human) (logL/Kg)	2D6	3A4	1A2	2C19	2C9	2D6	3A4	Total	AMES	Hepatotox icity	Skin Sensitization
	Absorption (human) (%absorbed)		Substrate		Inhibitor			Clearance (log mL/min/kg)					
26a	87.753	-1.08	No	Yes	Yes	Yes	Yes	No	No	0.515	Yes	Yes	No
26b	23.724	-1.031	No	No	No	No	No	No	No	0.332	Yes	Yes	No
26c	86.908	-0.911	No	Yes	Yes	Yes	Yes	No	Yes	0.516	Yes	Yes	No
26d	87.292	-1.086	No	Yes	Yes	Yes	Yes	No	Yes	0.553	Yes	Yes	No
26e	85.541	-0.962	No	Yes	Yes	Yes	Yes	No	Yes	0.546	No	Yes	No
26f	93.676	-1.128	No	Yes	No	Yes	Yes	No	Yes	0.675	No	Yes	No
26g	83.840	-1.334	No	Yes	No	Yes	Yes	No	Yes	0.373	No	Yes	No
26h	86.952	-0.872	No	Yes	No	Yes	Yes	No	Yes	0.504	No	Yes	No
26i	89.473	-0.853	No	Yes	No	No	Yes	No	Yes	0.417	No	Yes	No
26j	91.231	-0.999	No	Yes	No	Yes	Yes	No	Yes	0.281	Yes	Yes	No
26k	87.716	-1.138	No	Yes	No	Yes	Yes	No	Yes	0.148	No	Yes	No
26	73.445	-0.673	No	Yes	Yes	No	No	No	No	0.176	No	Yes	No

3.6 Absorption, distribution, metabolism, and excretion - toxicity profiles

To further study the pharmacological properties of the virtually designed compounds, the pkCSM online server was used to predict the absorption, distribution, metabolism, and excretion - toxicity (ADME/T) characteristics (Table 3). The intestinal absorption rate of most compounds is in the range of 73% to 91% (except for compound 26b, which has an intestinal absorption rate of 23%), exhibiting very high absorption characteristics, especially for compound 26j. The logVDss of all compounds was

higher than -0.15, which means that they are distributed more so in tissues than plasma. For metabolism, all compounds except compound 26b were predicted to be substrates of the CYP450 3A4 subtype, which means they may be metabolized by CYP 3A4. In addition, all compounds may not inhibit the CYP450 2D6 subtype, but may inhibit other CYP450 subtypes. Based on the predicted total clearance, liver and kidney tissues can clear all compounds in a combined manner. The expected toxicity indicates that all compounds may be harmful to the liver, and that they all compounds may not be responsible for skin

allergies. In addition, compounds 26e, 26f, 26g, 26h, 26i, and 26k were not mutagenic. The synthesis accessibility scores of all compounds ranged from 2.79 to 4.05, showing a low structural complexity, thus demonstrating the feasibility of synthesis (Table 2). Computational pharmacokinetic and toxicological studies along with available synthetic methods indicate that the virtually designed compounds can be used as lead compounds for further development.

4 Conclusion

HDAC6 has attracted much attention as an emerging and interesting target. In this study, we combined 3D-QSAR, molecular docking, and MD simulations to study benzohydroxamate-based HDAC6 inhibitors. We established a 3D-QSAR model, revealed the key structural factors of HDAC6 inhibitors, and determined the modified regions of these compounds. Based on the results of 3D-QSAR and molecular docking, we designed and predicted 11 new compounds. The results show that these compounds have robust predicted activity and ADME/T curves. Molecular docking and MD simulation results indicate that the compound binds to HDAC6 through various interactions, including the bond between the compound and zinc of the receptor. Over the course of this research, we also found some key residues in the binding pocket (such as Asp742 and His610). Overall, these results provide a useful reference for the rational design of such compounds.

Conflicts of interest

There are no conflicts to declare

Acknowledgments

This work was supported by the [National Natural Science Foundation of China] under Grant [81171508, 31400667]; [Key project of Chongqing Natural Science Foundation] under Grant [cstc2018jcyjAX0683, cstc2015jcyjBX0080]; and [Scientific and Technological Research Project of Chongqing Municipal Education Commission] under Grant [KJZD-K201801102, KJ1600908, KJQN201801132].

Notes and references

1. J. D. Osko and D. W. Christianson, *Biochemistry*, 2019, **58**, 4912-4924.
2. A. Norvell and S. B. McMahon, *Science*, 2010, **327**, 964-965.
3. J. E. López, E. D. Sullivan and C. A. Fierke, *ACS chemical biology*, 2016, **11**, 706-716.
4. T. Kouzarides, *The EMBO journal*, 2000, **19**, 1176-1179.
5. C. Choudhary, C. Kumar, F. Gnad, M. L. Nielsen, M. Rehman, T. C. Walther, J. V. Olsen and M. Mann, *Science*, 2009, **325**, 834-840.
6. S. Zhao, W. Xu, W. Jiang, W. Yu, Y. Lin, T. Zhang, J. Yao, L. Zhou, Y. Zeng and H. Li, *Science*, 2010, **327**, 1000-1004.
7. Q. Wang, Y. Zhang, C. Yang, H. Xiong, Y. Lin, J. Yao, H. Li, L. Xie, W. Zhao and Y. Yao, *Science*, 2010, **327**, 1004-1007.
8. C. Choudhary, B. T. Weinert, Y. Nishida, E. Verdin and M. Mann, *Nature reviews Molecular cell biology*, 2014, **15**, 536-550.
9. G. P. Delcuve, D. H. Khan and J. R. Davie, *Clinical epigenetics*, 2012, **4**, 5.
10. S. L. Berger, T. Kouzarides, R. Shiekhattar and A. Shilatifard, *Genes & development*, 2009, **23**, 781-783.
11. G. Yan, D. Li, X. Zhong, G. Liu, X. Wang, Y. Lu, F. Qin, Y. Guo, S. Duan and D. Li, *Journal of Biomolecular Structure and Dynamics*, 2020, 1-22.
12. C. M. Grozinger and S. L. Schreiber, *Chemistry & biology*, 2002, **9**, 3-16.
13. A. Verdel and S. Khochbin, *Journal of Biological Chemistry*, 1999, **274**, 2440-2445.
14. C. M. Grozinger, C. A. Hassig and S. L. Schreiber, *Proceedings of the National Academy of Sciences*, 1999, **96**, 4868-4873.
15. R. J. Mobley, D. Raghu, L. D. Duke, K. Abell-Hart, J. S. Zawistowski, K. Lutz, S. M. Gomez, S. Roy, R. Homayouni and G. L. Johnson, *Cell reports*, 2017, **18**, 2387-2400.
16. H. Losson, M. Schnekenburger, M. Dicato and M. Diederich, 2020.
17. J. B. Mortenson, L. N. Heppler, C. J. Banks, V. K. Weerasekara, M. D. Whited, S. R. Piccolo, W. E. Johnson, J. W. Thompson and J. L. Andersen, *Journal of Biological Chemistry*, 2015, **290**, 12487-12496.
18. A. Matsuyama, T. Shimazu, Y. Sumida, A. Saito, Y. Yoshimatsu, D. Seigneurin - Berny, H. Osada, Y. Komatsu, N. Nishino and S. Khochbin, *The EMBO journal*, 2002, **21**, 6820-6831.
19. Y. Li, D. Shin and S. H. Kwon, *The FEBS journal*, 2013, **280**, 775-793.
20. R. Prior, L. Van Helleputte, Y. E. Klingl and L. Van Den Bosch, *Expert opinion on therapeutic targets*, 2018, **22**, 993-1007.
21. O. Moreno-Gonzalo, F. Mayor Jr and F. Sánchez-Madrid, *Trends in immunology*, 2018, **39**, 591-595.
22. J. H. Kalin and J. A. Bergman, *Journal of medicinal chemistry*, 2013, **56**, 6297-6313.
23. K. Maharaj, J. J. Powers, A. Achille, S. Deng, R. Fonseca, M. Pabon-Saldana, S. N. Quayle, S. S. Jones, A. Villagra and E. M. Sotomayor, *Blood advances*, 2018, **2**, 3012-3024.
24. T. Li, C. Zhang, S. Hassan, X. Liu, F. Song, K. Chen, W. Zhang and J. Yang, *Journal of hematology & oncology*, 2018, **11**, 1-10.
25. A. Li, P. Chen, Y. Leng and J. Kang, *Oncogene*, 2018, **37**, 5952-5966.
26. C. Seidel, M. Schnekenburger, A. Mazumder, M.-H. Teiten, G. Kirsch, M. Dicato and M. Diederich, *Biochemical pharmacology*, 2016, **99**, 31-52.
27. Y.-L. Hsieh, H.-J. Tu, S.-L. Pan, J.-P. Liou and C.-R. Yang, *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 2019, **1866**, 992-1003.
28. X.-X. Wang, R.-Z. Wan and Z.-P. Liu, *European journal of medicinal chemistry*, 2018, **143**, 1406-1418.
29. J. Haakenson and X. Zhang, *International journal of molecular sciences*, 2013, **14**, 9514-9535.
30. Y. Li and E. Seto, *Cold Spring Harbor perspectives in medicine*, 2016, **6**, a026831.
31. J.-H. Lee, A. Mahendran, Y. Yao, L. Ngo, G. Venta-Perez, M. L. Choy, N. Kim, W.-S. Ham, R. Breslow and P. A. Marks, *Proceedings of the National Academy of Sciences*, 2013, **110**, 15704-15709.
32. L. Santo, T. Hideshima, A. L. Kung, J.-C. Tseng, D. Tamang, M. Yang, M. Jarpe, J. H. van Duzer, R. Mazitschek and W. C. Ogier, *Blood, The Journal of the American Society of Hematology*, 2012, **119**, 2579-2589.
33. P. Huang, I. Almeciga-Pinto, M. Jarpe, J. H. van Duzer, R. Mazitschek, M. Yang, S. S. Jones and S. N. Quayle, *Oncotarget*, 2017, **8**, 2694.
34. B. Vergani, G. Sandrone, M. Marchini, C. Ripamonti, E. Cellupica, E. Galbiati, G. Caprini, G. Pavich, G. Porro and I. Rocchio, *Journal of medicinal chemistry*, 2019, **62**, 10711-10739.
35. A. P. Tiwari, V. B. Giliyar, G. G. Shenoy and V. K. Eshwara, *Letters in Drug Design & Discovery*, 2020, **17**, 31-47.

36. R. D. Cramer, D. E. Patterson and J. D. Bunce, *Journal of the American Chemical Society*, 1988, **110**, 5959-5967.
37. G. Klebe, U. Abraham and T. Mietzner, *Journal of medicinal chemistry*, 1994, **37**, 4130-4146.
38. H. Chu, Q. X. He, J. W. Wang, Y. T. Deng, J. Wang, Y. Hu, Y. Q. Wang and Z. H. Lin, *Journal of biomolecular structure & dynamics*, 2019, DOI: 10.1080/07391102.2019.1697366, 1-12.
39. Q. He, H. Chu, Y. Wang, H. Guo, Y. Wang, S. Wang, Z. Feng, X.-Q. Xie, Y. Hu and H. Liu, *Journal of Biomolecular Structure and Dynamics*, 2019, 1-12.
40. E. Vrontaki, G. Melagraki, T. Mavromoustakos and A. Afantitis, *Methods*, 2015, **71**, 4-13.
41. H. Guo, Y. Wang, Q. He, Y. Zhang, Y. Hu, Y. Wang and Z. Lin, *Journal of Molecular Structure*, 2019, **1193**, 223-230.
42. Y. Hai and D. W. Christianson, *Nature Chemical Biology*, 2016, **12**, 741-747.
43. J. Ma, X. Guo, S. Zhang, H. Liu, J. Lu, Z. Dong, K. Liu and L. Ming, *Molecular medicine reports*, 2015, **11**, 4525-4531.
44. Y. Gao, Y. Chen, Y. Tian, Y. Zhao, F. Wu, X. Luo, X. Ju and G. Liu, *New Journal of Chemistry*, 2019, **43**, 17004-17017.
45. Y. Zhang, G. Zheng, T. Fu, J. Hong, F. Li, X. Yao, W. Xue and F. Zhu, *Physical Chemistry Chemical Physics*, 2020, **22**, 5132-5144.
46. Y. Zhang, J. B. Ying, J. J. Hong, F. C. Li, T. T. Fu, F. Y. Yang, G. X. Zheng, X. J. Yao, Y. Lou and Y. Qiu, *ACS chemical neuroscience*, 2019, **10**, 2467-2480.
47. G. Zheng, W. Xue, F. Yang, Y. Zhang, Y. Chen, X. Yao and F. Zhu, *Physical Chemistry Chemical Physics*, 2017, **19**, 28885-28896.
48. D. E. Pires, T. L. Blundell and D. B. Ascher, *Journal of medicinal chemistry*, 2015, **58**, 4066-4072.
49. A. Daina, O. Michielin and V. Zoete, *Scientific Reports*, 2017, **7**, 42717.
50. P. Li and K. M. Merz Jr, *Journal*, 2016.
51. R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand and R. C. Walker, *Journal of chemical theory and computation*, 2013, **9**, 3878-3888.
52. A. W. Götz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand and R. C. Walker, *Journal of chemical theory and computation*, 2012, **8**, 1542-1555.
53. J. A. Izaguirre, D. P. Catarella, J. M. Wozniak and R. D. Skeel, *The Journal of chemical physics*, 2001, **114**, 2090-2098.
54. R. J. Loncharich, B. R. Brooks and R. W. Pastor, *Biopolymers: Original Research on Biomolecules*, 1992, **32**, 523-535.
55. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *The Journal of chemical physics*, 1995, **103**, 8577-8593.
56. T. Darden, D. York and L. Pedersen, *The Journal of chemical physics*, 1993, **98**, 10089-10092.
57. J.-P. Ryckaert, G. Ciccotti and H. J. Berendsen, *Journal of computational physics*, 1977, **23**, 327-341.
58. H. Sun, L. Duan, F. Chen, H. Liu, Z. Wang, P. Pan, F. Zhu, J. Z. H. Zhang and T. Hou, *Physical Chemistry Chemical Physics*, 2018, **20**, 14450-14460.
59. F. Chen, H. Sun, J. Wang, F. Zhu, H. Liu, Z. Wang, T. Lei, Y. Li and T. Hou, *RNA*, 2018, **24**, 1183-1194.
60. J. Hu, Z. Feng, S. Ma, Y. Zhang, Q. Tong, M. H. Alqarni, X. Gou and X.-Q. Xie, *Journal of chemical information and modeling*, 2016, **56**, 1152-1163.