Supporting Information:

A salt bridge turns off the foot-pocket in class-II HDACs

Jingwei Zhou[a], *Zuolong Yang*[a], *Fan Zhang*[a], *Hai-Bin Luo*[a], *Min Li*[*, a] *and Ruibo Wu*[*,a]

[a] School of Pharmaceutical Sciences, Sun Yat-Sen University, Guangzhou 510006, P.R. China

*To whom correspondence should be addressed.
E-mail: wurb3@mail.sysu.edu.cn; limin65@mail.sysu.edu.cn

Computational Method; Experimental Method; References; Structure synthesis and determination of the β-substituted chalcone; Figure S1-S14.

Computational Method

Nine models (5 wild types and 4 mutant types) including the HDAC2-Biaryl Benzamide model (PDB code: 3MAX[1]), HDAC2-SAHA model (PDB code: 4LXZ[1]), HDAC2-Act model (modified from the HDAC2-SAHA model), HDAC7-SAHA model (PDB code: 3C0Z[2]), HDAC7-Act model (modified from HDAC7-SAHA model), E543M HDAC7-SAHA model, R665A HDAC7-SAHA model, E543M HDAC7-Biaryl Benzamide and R665A HDAC7-Biaryl Benzamide (all the four mutation models were also modified from HDAC7-SAHA model, by using the *Molecule Operating Environment (MOE)* package[3]) were built in our simulations. The detailed preparation protocols were similar to our previous studies.[4-6] The protonation states of charged residues were determined by H++ program[7] and by carefully examining their individual local hydrogen bond networks, and then each prepared model was neutralized by adding Na^+ or Cl^- ion at the protein surface with the Amber tool, and solvated into a rectangular box with a 10 Å buffer distance between the solvent box wall and the nearest solute atoms. The TIP3P model[8] and Amber99SB force field[9-11] were employed for water molecules and the proteins, respectively. The charge on Zn ion was set to +2 as used in our previous studies[4-6]. And the force field parameters of the inhibitors (biaryl benzamide, SAHA and Act) were generated from AMBER GAFF force field[12]. The partial atomic charges of these inhibitors were obtained from the restrained electrostatic potential (RESP) charge at the HF/6-31G* level with the Gaussian09 package.[13]

The classical molecular dynamics (MD) simulations were performed in the Amber12 molecular simulation package.[14] Every model was firstly minimized, and then gradually heated from 0 to 300 K over a period of 50 ps, followed by another 100 ps of NPT MD simulations to relax the system density to about 1.0 g/cm3, with the target temperature of 300 K and the target pressure of 1.0 atm. Afterward, 100 ns of NVT MD simulation with a target temperature of 300 K via employing the periodic boundary condition were performed for each model to produce trajectories. All the simulations were accomplished

with applying the GPU-accelerated pmemd program[14] in Amber12, and a time step of 1.0 fs was used for all the simulations. The SHAKE algorithm[15] was applied to constrain all hydrogen-containing bonds with a tolerance of 10^{-5}. The Berendsen thermostat method[16] was used to control the system temperature, and a cutoff of 12 Å was set for both van de Waals and electrostatic interactions. The reliability of our MD simulations was confirmed by investigating RMSD change as well as structures superposition between crystal structures and MD results. The skeleton structures and the heavy atoms of the crystal proteins (PDB codes: 3MAX, 4LXZ and 3C0Z) were set as the reference structures for calculating the RMSD value. The RMSD of Loop3 were also calculated to cluster the open-close states of foot-pocket as shown in Figure 2, with the heavy atoms of Loop3 of the crystal proteins as reference structures. And the structures from the 100 ns MD trajectories were collected by sampling at 10 ps intervals to process the above RMSD and following cluster analysis.

Cluster analysis is a general unsupervised technique for finding structure patterns within simulation data.[17-19] Root mean-square deviation (RMSD)-based clustering was performed with the program *ptraj*, a simulation analysis tools implemented in Amber12. As one of the most popular clustering algorithms, the average linkage cluster algorithm integrated into Amber12 was applied in our simulation results herein.[20] The structures from the 100 ns MD trajectories were collected by sampling at 10 ps intervals to process the cluster analysis based on the RMSD of the heavy atoms of Loop3. For HDAC2-Act, HDAC2-SAHA, E543M HDAC7-SAHA and R665A HDAC7-SAHA models, five clusters were generated. After that, the small percentage clusters were discard and then only the top two clusters were taken into account. While for the HDAC2-Biaryl Benzamide, HDAC7-Act, HDAC7-SAHA, E543M HDAC7-Biaryl Benzamide and R665A HDAC7-Biaryl Benzamide models, only one cluster were obtained.

Experimental Methods

Protein Expression and Purification.

The gene sequence of wild type HDAC7 was firstly obtained from the data bank of NCBI. Then, the truncated HDAC7 (residues 480 to 903) without their plastid targeting sequences were cloned with the primers "CGGGATCCCGGGCTCAGTCTTCCCCAGC" and "CCGCTCGAGGGAGGCCAGGCGCTGCAT" synthesized by *Generay biotech co. Ltd (Shanghai, China)*. The cloned HDAC7 genes were further amplified by PCR as shown in Figure S8 and verified by DNA sequencing. The size of cloned HDAC7 gene was about 1300 *bp* in line with expectations. The cloned HDAC7 fragments were then subcloned into a *pET-28a(+)* vector to obtain the HDAC7-*pET-28a(+)* construct. Meanwhile, for E543M and R665A HDAC7 mutants in which the E543 and R665 residues of HDAC7 were mutated to MET and ALA, respectively, two constructs E543M HDAC7-*pET-28a(+)* and R665A HDAC7-*pET-28a(+)* were prepared by QuickChange (Clonetech, Japan). As shown in Figure S9, the three constructs are about 7000 *bp* in line with expectations.

The three constructs were transformed into BL21 (DE3) competent cells to express full length wild type and mutant HDAC7. The cells were cultured at 37℃ in LB medium with proper antibiotic. When the OD value of cell culture medium reached 0.6, 1 mM isopropyl-1-thio-D-galactopyranoside was added to induce protein expression for 6 h. Expressed proteins were verified by SDS-PAGE and further western blot, the results were shown in Figure S10a and S10b. For protein purification, the well-expressed cells were further transferred to LB fluid medium at 37℃. When the OD value reached to 0.5 ~ 0.6, adjust the cultured temperature to room temperature and 1 mM isopropyl-1-thio-D-galactopyranoside was added to induce protein expression overnight. Expressed proteins were purified by affinity chromatography using Ni-NTA resin (Novoprotein, Shanghai),

and the Purification results were shown in Figure S10c and S10d.

HDAC7 Activity Assay.

The fluorescence activity assay was employed to determine the HDAC7 activity. In the first step, an acetylated lysine substrate was incubated with wild and mutant HDAC7. Deacetylation sensitized the substrate with HDAC developer in the second step released a fluorescent product. The fluorophore was easily analyzed using a fluorescence plate reader with excitation wavelength of 347 nm and emission wavelength of 465 nm. The assay buffer (25 mM Tris-HCl, pH 8.0, 137 mM NaCl, 2.7 mM KCl and 1 mM MgCl2), acetylated lysine substrate (3.4 mM) and HDAC developer were purchased from Cayman (U.S.A). The fluorescence signals represent the activity of the wild type HDAC7 and mutant HDAC7, and it also was background-corrected against the reaction containing HDAC assay buffer, acetylated lysine substrate and HDAC developer but no HDAC7 enzyme. Moreover, the fluorescence signal from each mutant HDAC7 was normalized to the wild type HDAC7 (set as 100%). The mean and standard error form at least three independent assays.

Fluorogenic Assay of HDAC7 Inhibition Activity.

The assay theory is similar to the HDAC7 activity assay. All assays were performed according to the protocol of HDAC inhibitor screening assay kit. First, 140 μL of assay buffer was added into each well on a 96-well plate. HDAC enzymes were diluted to the desired concentrations with assay buffer and then, 10 μL diluted enzyme was dispensed into the well. Second, the biaryl benzamide compound (purchased from Wuhan NCE Biomedical Co.Ltd (China)) was dissolved into assay buffer primarily, and a series of five-fold dilutions of each compound were prepared with assay buffer. 10 μL of each dilution was also added into the correspondent well to mix with enzyme at room temperature. Third, 10 μL acetylated lysine substrate was added to correspondent well to

initiate deacetylation reaction after the mixture above had been mixed at room temperature for 1 hours. The deacetylation reaction was kept up for 30 minutes on a shaker at 37 ℃. Finally, 40 μL of HDAC developer which included TSA (purchased from Cayman (U.S.A), 0.21 mM) used to terminate the deacetylation process was assigned to the well to combine with the deacetylated lysine for 15 minutes on a shaker at room temperature. The fluorescence intensity of the combination product was measured on a microplate reader (FlexStation 3, Molecular Devices) in 30 minutes using an excitation wavelength of 340-360 nm and an emission wavelength of 440-465 nm. In addition, the fluorescence intensity of the 100% initial activity without enzyme (F100) and background activity without inhibitors (F0) were also taken into account to equilibrate the assay result. The fluorescence intensity data were analyzed by Origin software. So in the presence of inhibitors and HDAC enzymes, the fluorescence intensity (F) was laid between F0 and F100. And the accurate relative HDAC activity was calculated based on the following equation: %activity = (F-F0)/ (F100-F0). The TSA inhibitor was also employed as the positive control medicine to demonstrate the reliability of the method above. Each assay was repeated three times and the mean values were used for "S" curve fitting to obtain the IC50 values.

In Vitro Antiproliferative Activity Assay.

In vitro antiproliferative assay was performed by the MTT (3-[4,5-dimethyl-2-thiazolyl]-2,5-diphenyl-2H-tetrazolium bromide) method. Briefly, all cell lines were maintained in RPMI1640 medium containing 10% FBS at 37 ℃ in 5% CO_2 humidified incubator. Cells were passaged the day before dosing into a 96-well plate and allowed to grow for a minimum of 14 h prior to addition of compounds. After compounds addition, the plates were incubated for an additional 48 h, and then 0.5% MTT solution was added to each well. After further incubation for 4 h, formazan formed from MTT was extracted by adding 200 μl of DMSO for 15 min. Absorbance was then determined using an plate reader at 570 nm, and each assay was repeated three times and the mean values were used

for “S” curve fitting to obtain the IC50 values.

Reference

1 J. C. Bressi, A. J. Jennings, R. Skene, Y. Wu, R. Melkus, R. De Jong, S. O'Connell, C. Grimshaw, M. Navre and A. R. Gangloff, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 3142-3145.

2 A. Schuetz, J. Min, A. Allali-Hassani, M. Schapira, M. Shuen, P. Loppnau, R. Mazitschek, N. P. Kwiatkowski, T. A. Lewis, R. L. Maglathin, T. H. McLean, A. Bochkarev, A. N. Plotnikov, M. Vedadi and C. H. Arrowsmith, *J. Biol. Chem.*, 2008, **283**, 11355–11363.

3 *Molecular Operating Environment (MOE), 2013.08, Chemical Computing Group Inc.; 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7*, 2016.

4 J. W. Zhou, M. Li, N. H. Chen, S. L. Wang, H. B. Luo, Y. K. Zhang and R. B. Wu, *ACS Chem. Biol.*, 2015, **10**, 687-692.

5 J. W. Zhou, H. J. Xie, Z. H. Liu, H. B. Luo and R. B. Wu, *J. Chem. Inf. Model.*, 2014, **54**, 3162-3171.

6 J. W. Zhou, R. B. Wu and H. B. Luo, *Phys. Chem. Chem. Phys.*, 2015, **17**, 29483-29488.

7 J. C. Gordon, J. B. Myers, T. Folta, V. Shoja, L. S. Heath and A. Onufriev, *Nucleic Acids Res.*, 2005, **33**, W368–W371.

8 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.

9 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179-5197.

10 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins: Struct., Funct., Bioinf.*, 2006, **65**, 712-725.

11 J. M. Wang, P. Cieplak and P. A. Kollman, *J. Comput. Chem.*, 2000, **21**, 1049-1074.

12 J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.

13 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant,

S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Journal*, 2009.

14 D. A. Case, T. A. Darden, I. Cheatham, T.E., C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Götz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko and P. A. Kollman, *AMBER 12, University of California, San Francisco.*, 2012.

15 J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.* , 1977, **23**, 327-341.

16 H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.

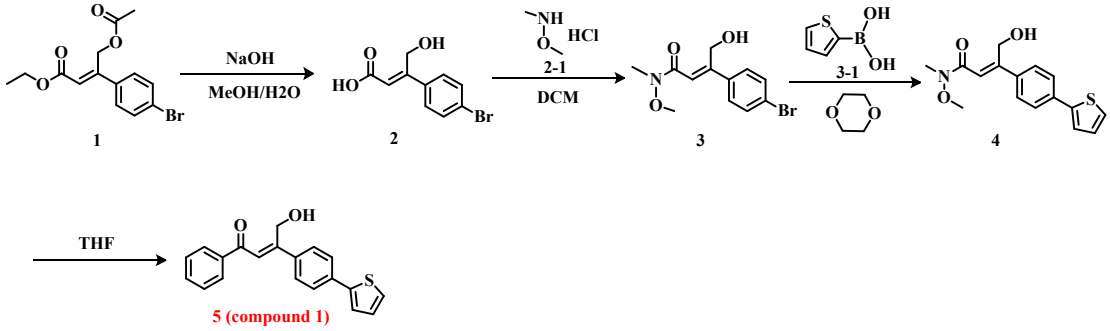
17 S. G. Estacio, R. Moreira and R. C. Guedes, *J Chem Inf Model*, 2011, **51**, 1690-1702.

18 R. E. Amaro, R. V. Swift, L. Votapka, W. W. Li, R. C. Walker and R. M. Bush, *Nat Commun*, 2011, **2**.

19 G. Colombo, G. Morra, M. Meli and G. Verkhivker, *P Natl Acad Sci USA*, 2008, **105**, 7976-7981.

20 J. Y. Shao, W. T. Stephen, T. Nephi and E. C. Thomas, *J.Chem.Theory Comput.*, 2007, **3**, 2312-2334.

Structure synthesis determination of the *β*-substituted chalcone



1. To a stirred solution of **1** (2 g) in a mixture of MeOH (20 ml) and H_2O (2 ml), NaOH (2 eq) was added. Then the mixture was stirred for 2 h at 60 ℃. It was concentrated and adjusted pH to 4~5 by HCl (5 M). The mixture was extracted by DCM : MeOH = 10:1 (80 ml * 3). The DCM layer was dried by Na_2SO_4, concentrated to **2** which used for next step without purification.

2. To a stirred solution of **2** in DCM (60 ml), EDCI (1.3 eq) and HOBT (1.1 eq) were added. The mixture was stirred for 1 h at RT. **2-1** (3 eq) and TEA (3 eq) were added to the solution. Then it was stirred over night at RT. DCM (60 ml) was added and the DCM layer was washed by $NaHCO_3$ aq (100 ml), HCl aq (100 ml), and NaCl aq (100 ml), dried and concentrated, which was purified by Flash with PE : EA = 1:1 ~ 1:4 (1.3 L). The solution was concentrated to **3** (900 mg, 50% for two steps).

3. To a stirred solution of **3** and **3-1** (2.5 eq) in dioxane (30 ml), Pd(pddf)Cl (0.1 eq) and K_2CO_3 (2 eq) were added under N_2 protection. The mixture was stirred over night at 100 ℃. After cooled to RT, it was concentrated and extracted by DCM (100 ml). The DCM layer was washed by NaCl aq (60 ml), dried and concentrated, which was purified by Flash with PE : EA = 1:1 ~ 1:4 (1.0 L). The solution was concentrated to **4** (450 mg, 49.8%).

4. To a stirred solution of **4** in THF (60 ml), phenylmagnesium bromide solution (1 M in THF, 6 eq) was added at -20 ℃ under N2 protection. Then it was stirred for 2 h. NH_4Cl aq (10 ml) was added, which was extracted by DCM (200 ml). The DCM layer was washed by NaCl aq (60 ml), dried and concentrated, which was purified by Prep-TLC

with DCM : MeOH = 30 : 1 (400 ml). The slica gel was washed by DCM : MeOH = 10 : 1 (100 ml), which was concentrated to the **5** (***compound 1***, 39 mg, 8.21%). The synthesis method of compound 2 is similar to that of compound 1.

Compound 1. A white solid, 100% pure. ^{1}H NMR (400 MHz, DMSO-d6) δ 4.564 (s, 2H), 7.045-7.065 (m,1H), 7.209-7.277 (m,5H), 7.396-7.415 (m, 2H), 7.508-7.528 (m, 3H), 7.909-7.930 (m, 2H); LC-MS (ESI) m/s 320.9; Retention Time: 1.93 min, Mobile phase: A 0.01% TFA in water B 0.01% TFA in CAN.

Compound 2. A white solid, 99.19% pure. ^{1}H NMR (400 MHz, DMSO-d6) δ 4.366 (d, J = 3.2 Hz, 2H), 7.034 (s, 1H), 7.259-7.275 (m, 3H), 7.453-7.638 (m, 9H), 7.866-7.883 (m, 2H); LC-MS (ESI) m/s 315.1; Retention Time: 1.62 min, Mobile phase: A 0.01% TFA in water B 0.01% TFA in CAN.

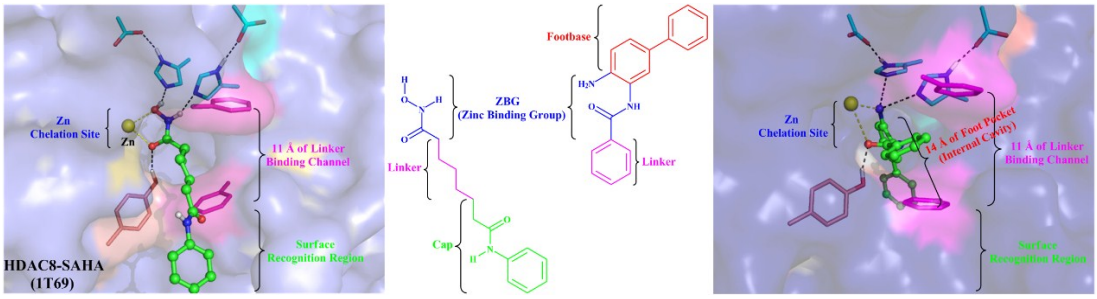


Figure S1. The partition of the HDAC active site and the HDAC inhibitors. The active site could be described by surface recognition region (green)-11 Å linker binding channel (pink)-zinc chelation site (blue)-14 Å foot-pocket (red). And corresponding to the four regions of the active site, the HDAC inhibitors could be characterized by cap (green)-linker (pink)-ZBG (blue)-foot (red).

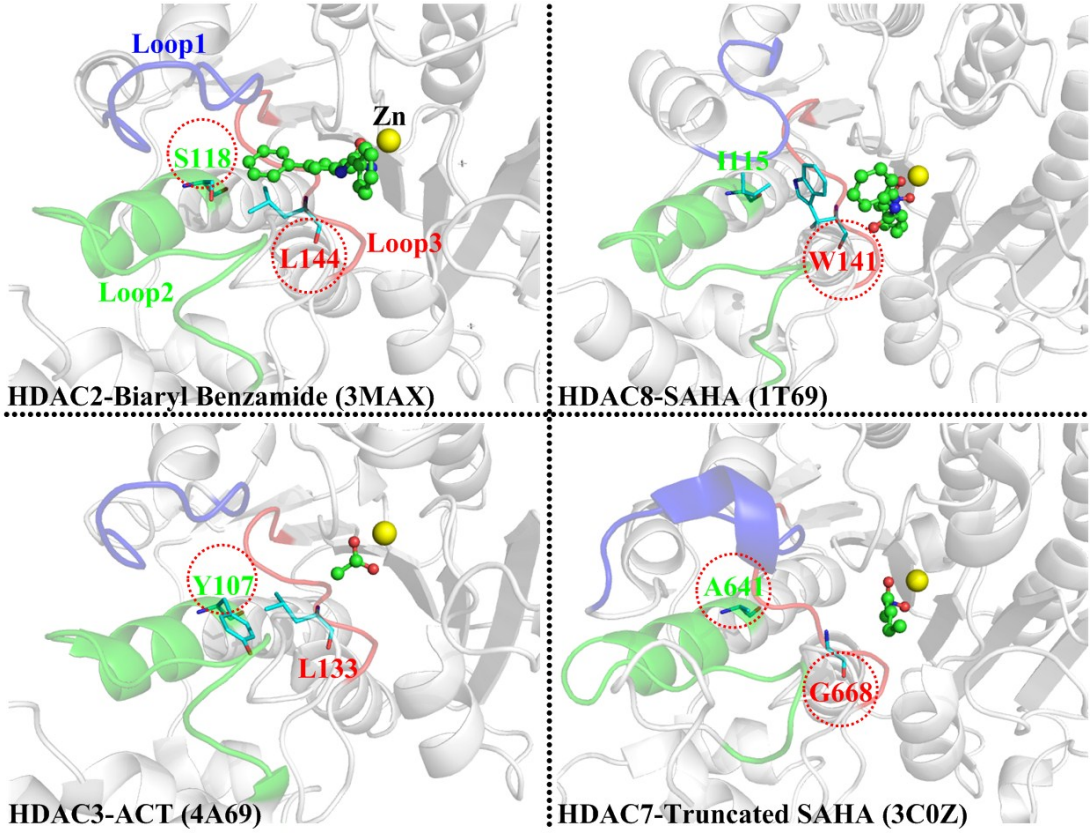


Figure S2. The "steric hindrance" hypothesis on the selective inhibition of Benzamide towards class I HDACs. For HDAC8, the big residue W141 located in Loop3 blocks the entrance of the foot pocket, which excludes the Biaryl Benzamide and Simple Benzamide from the foot pocket. For HDAC3, the aromatic residue Y107 located in Loop2 blocks

the exit of the foot pocket to form a very small foot-pocket, which excludes the Biaryl Benzamide from the foot pocket but it could accommodate Simple Benzamide. However, this hypothesis is inconsequential to explain why benzamide couldn't inhibit the class II HDACs, since the above mentioned residue is replaced by a small residue G668 in class II HDAC (take HDAC7 for numbering, see residue conservation in Figure 1), that is, no steric hindrance effect to block the benzamide enter into the foot-pocket in view of "steric hindrance" hypothesis.

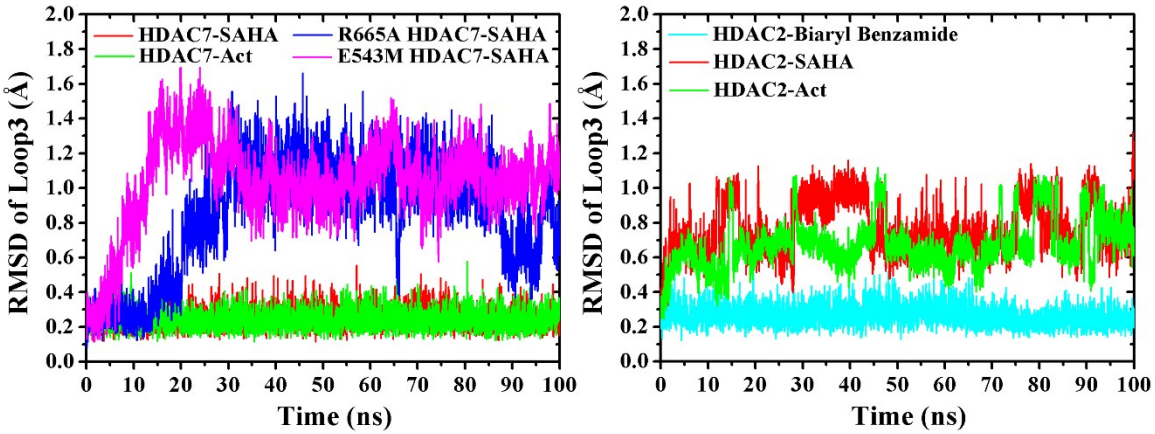


Figure S3. Flexibility of the Loop 3 in the different HDACs-Inhibitors models. The Loop3 is very stable in the wild type HDAC7-Act, HDAC7-SAHA and HDAC2-Biaryl Benzamide models. However, it becomes a little flexible in wild type HDAC2-Act, HDAC2-SAHA models, and especially flexible in mutant type R665A HDAC7-SAHA and E543M HDAC7-SAHA models. The heavy atoms of Loop3 in the crystal proteins (PDB codes: 3MAX, 4LXZ and 3C0Z) were set to the reference structures for calculating the RMSD.

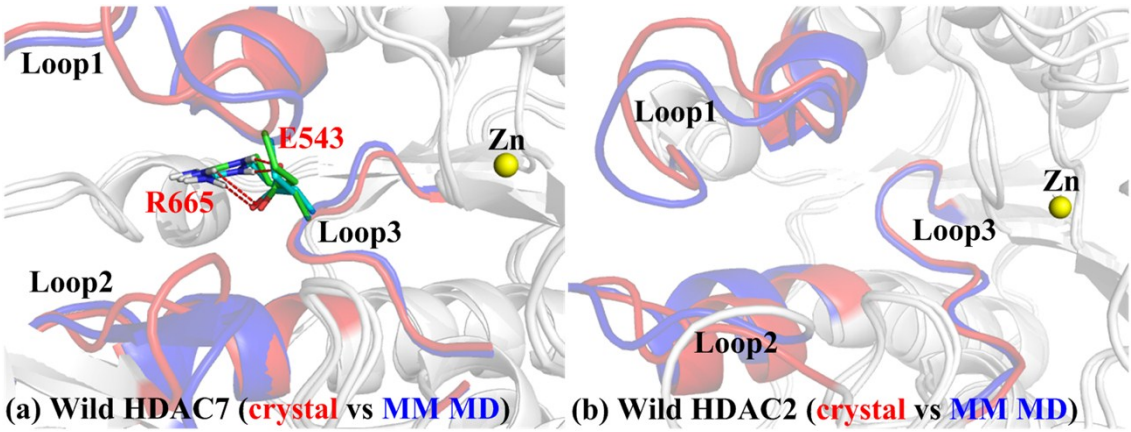


Figure S4. Comparison of the wild type HDAC2/7 crystal structures and MM MD structures. It indicates that our MD simulations could reproduced the crystal structures well.

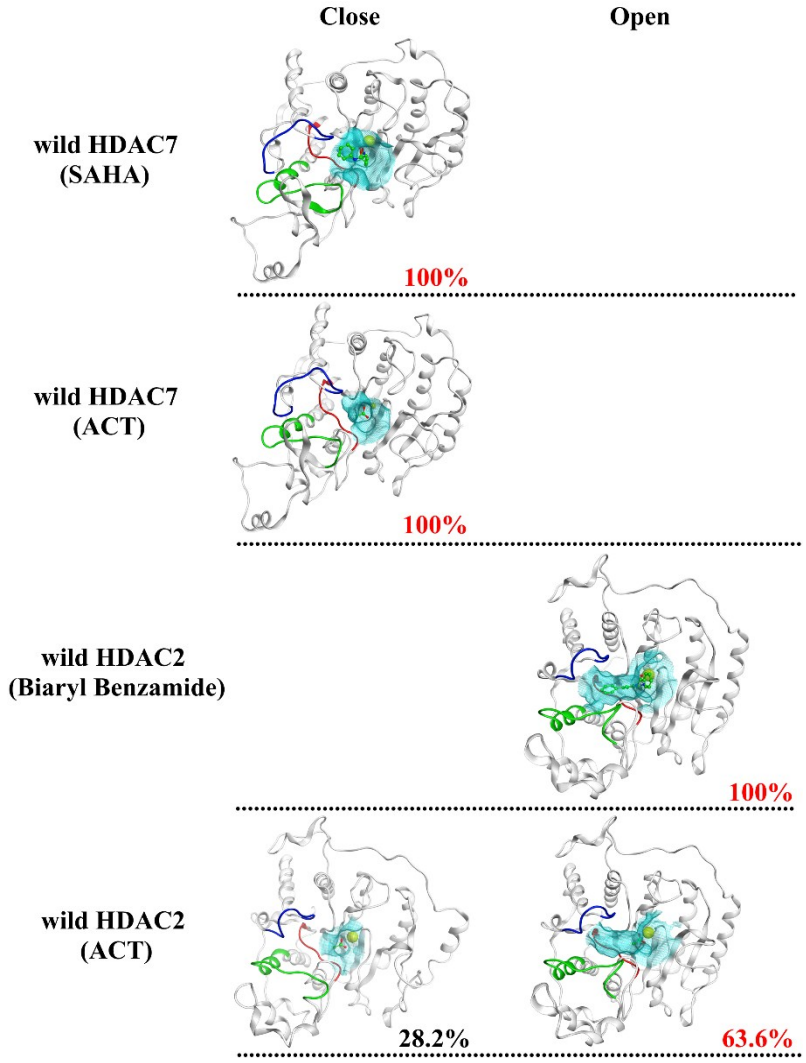


Figure S5. The comparison of the active pocket in wild HDAC7-SAHA/Act and HDAC2-Act/Biaryl Benzamide models from cluster analysis on Loop3 of the MM MD trajectory. In HDAC7-SAHA, HDAC7-Act models and HDAC2-Biaryl Benzamide model, the foot pocket only presents close conformation and open conformation,

respectively. While in HDAC2-Act, HDAC2-SAHA (see Figure 2), E543M HDAC7-SAHA and R665A HDAC7-SAHA models (two mutants also shown in Figure 2), both the close conformation and open conformation of foot pocket are observed.

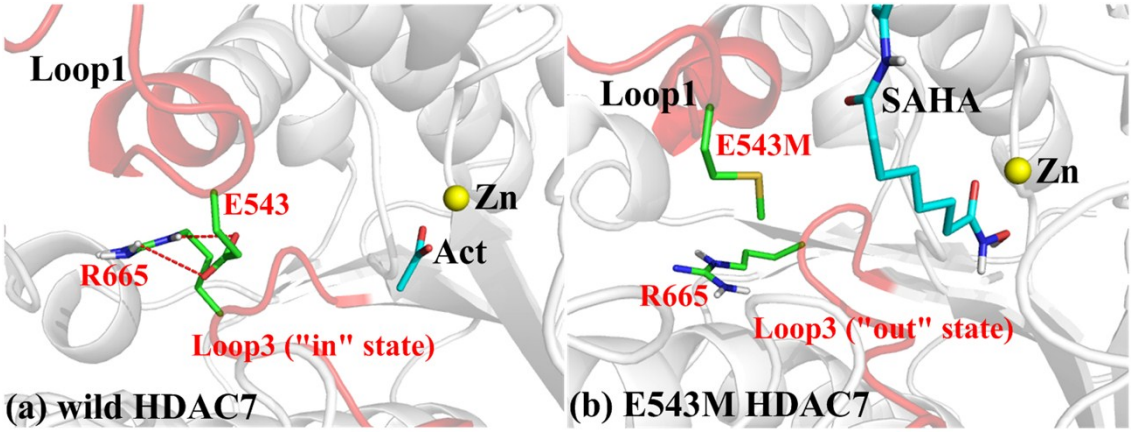


Figure S6. (a) The stable E-R salt bridge and the representative "in" conformation of loop3 in HDAC7-Act complex. (b) The "out" conformation of loop3 as E-R salt bridge broken by E543M mutant in HDAC7.

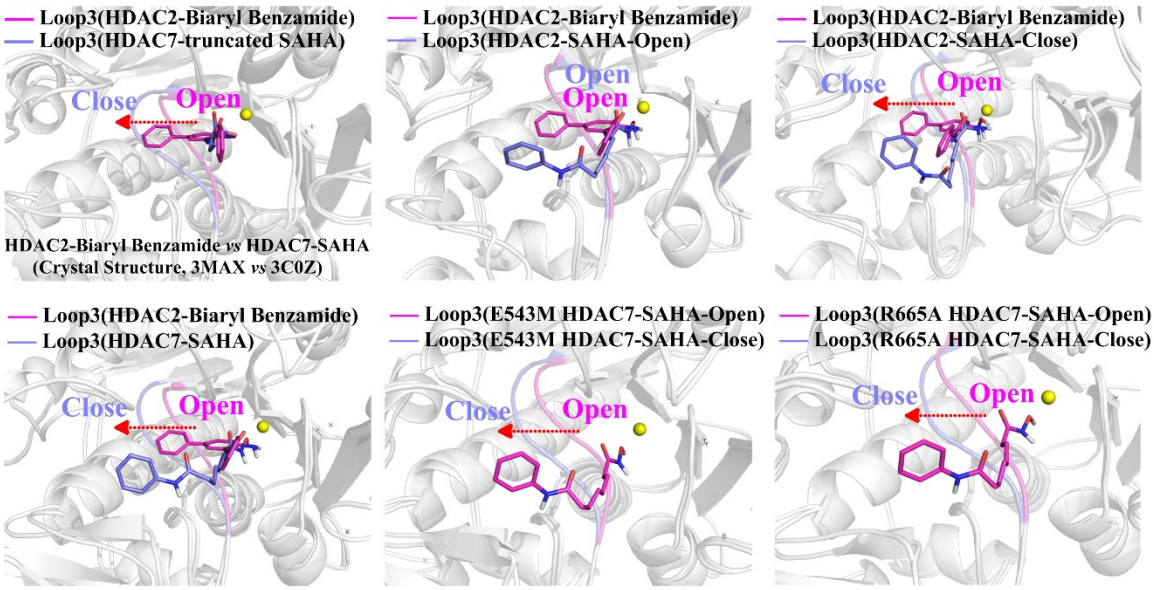


Figure S7. The Loop3 conformational dynamics corresponding to the open/close states of foot-pocket in the selected HDAC2 and HDAC7 models.

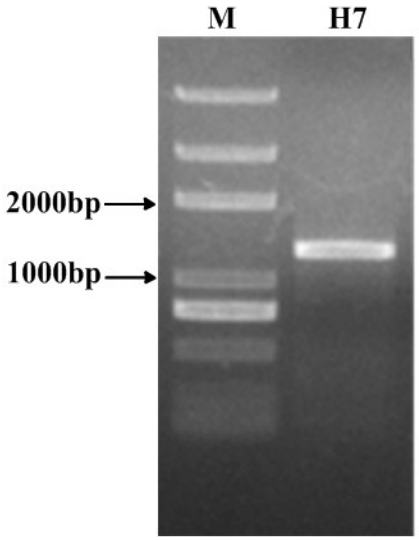


Figure S8. The opening reading frames encoding human HDAC7 were amplified by PCR. The size of obtained HDAC7 fragments is about 1300bp. M: marker; H7: the fragments of HDAC7 (wild type).

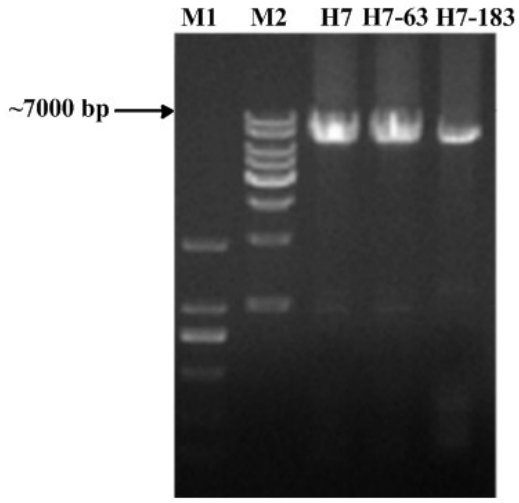


Figure S9. The constructs of full length (wild type) and mutant HDAC7 were confirmed by agarose electrophoreis. M1, M2: DNA marker; H7: wild HDAC7-*pET-28a(+)*; H7-63：E543M HDAC7-*pET-28a(+)*; H7-183: R665A HDAC7-*pET-28a(+)*.

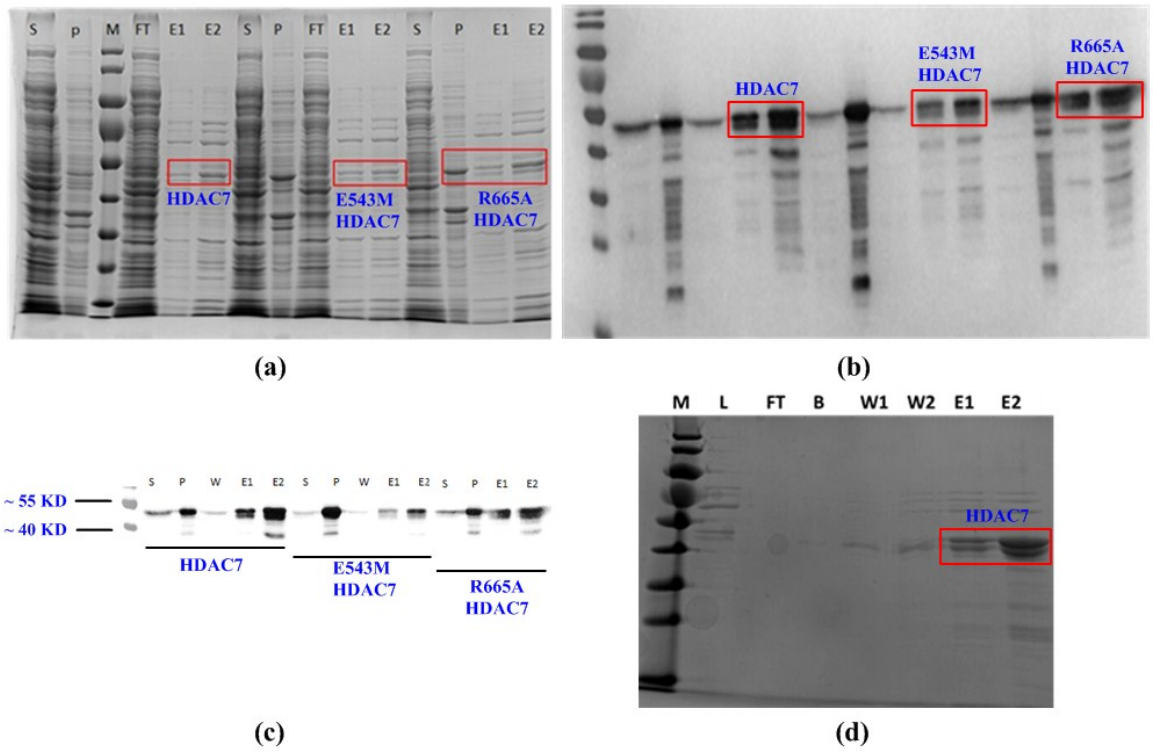


Figure S10. The expression (a and b) and purification (c and d) results of wild type and mutant HDAC7. S: Supernatant; P: Precipitate; M: Maker; FT: Flow Liquid; E1, E2: Eluent buffer. Most target proteins appeared in the eluent buffer (the size of wild and mutant HDAC7 is about 46 *kd*) and the purification of the proteins (wild and mutant) are up to 80%.

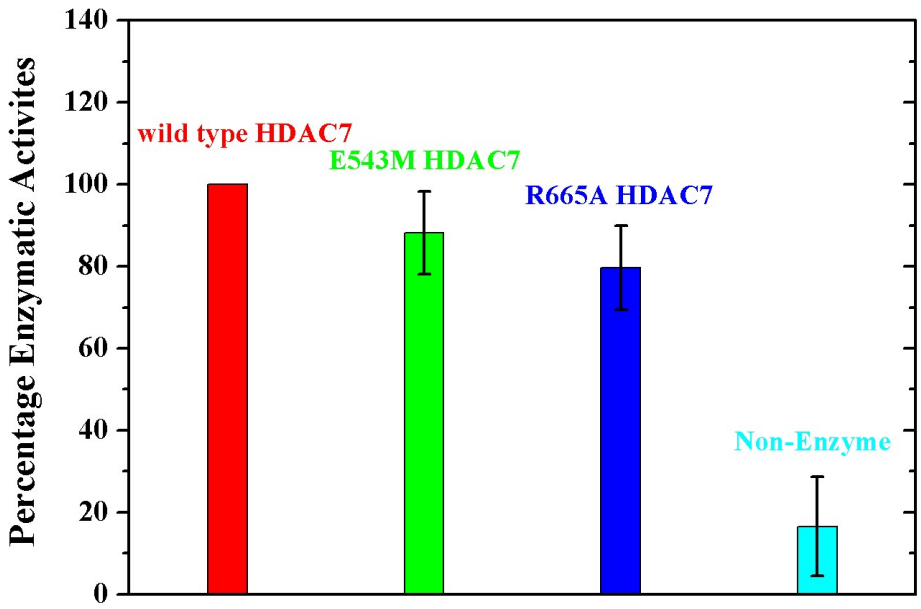


Figure S11. The effect of the E543M and R665A mutations on the HDAC7 deacetylation activity.

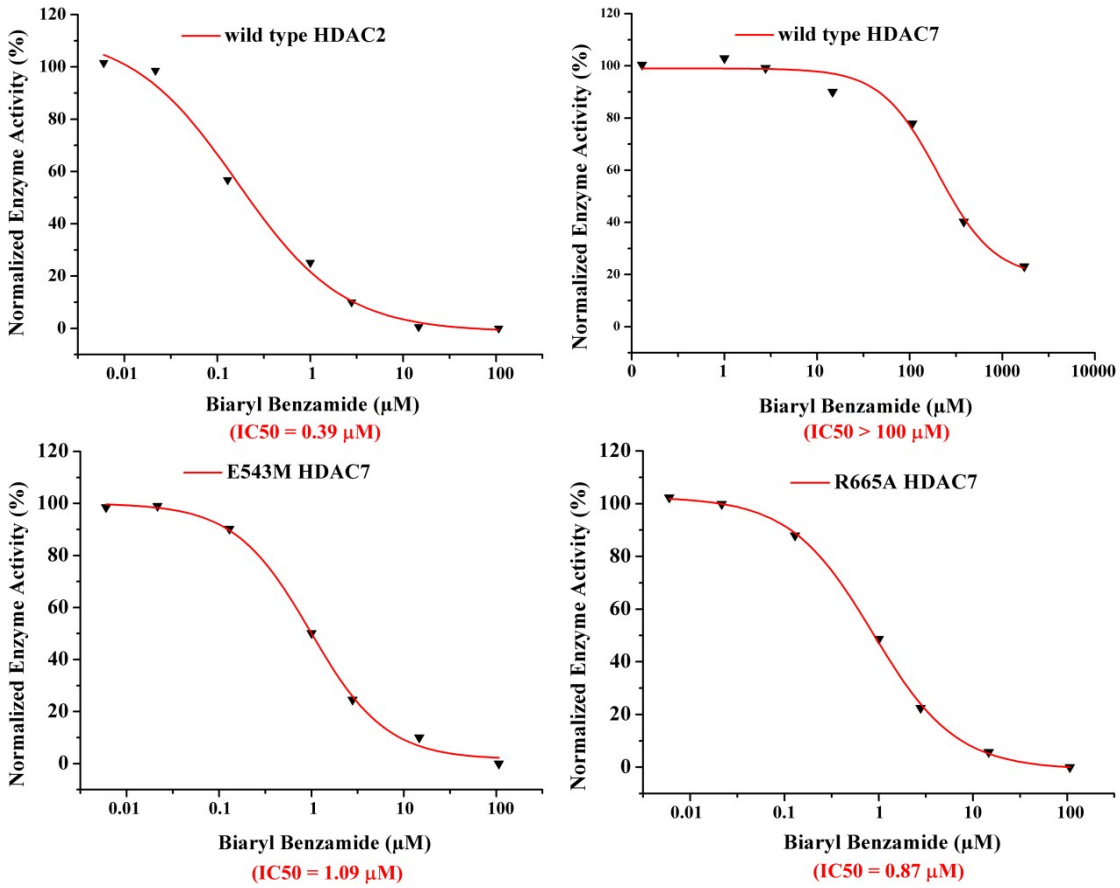


Figure S12. The inhibition of the Biaryl Benzamide against wild type HDAC2/7 and mutant HDAC7. Due to the destruction of the R665-E543 salt bridge, the foot pocket would present the open conformation, and the Biaryl Benzamide could enter into the foot pocket and finally inhibit the mutant HDAC7.

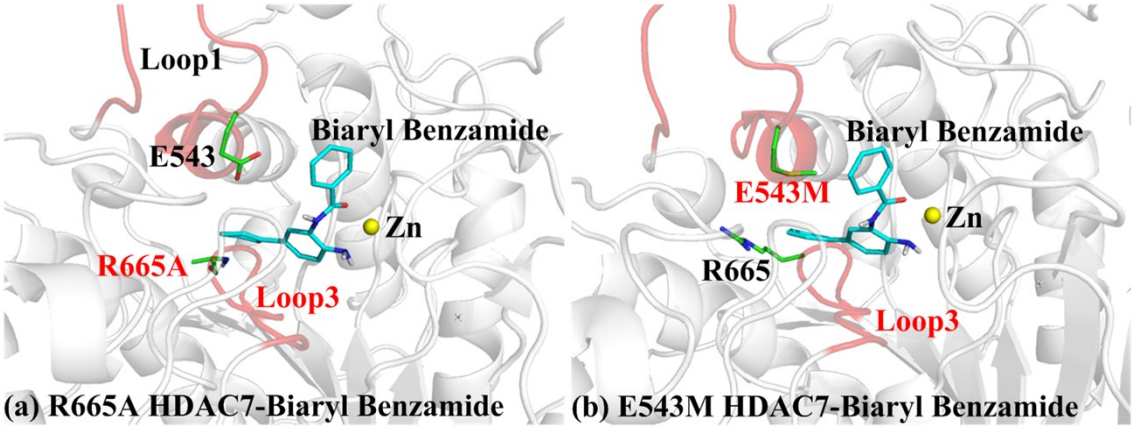


Figure S13. The representative structures of the mutant HDAC7-Biaryl Benzamide models. Due to the destruction of the R665-E543 salt bridge, the foot pocket would be turn on with loop3 in "out" conformation, and the additional C-ring of Biaryl Benzamide could enter into the foot-pocket and thus present inhibitory effect on the mutant HDAC7 as experiment observation which summarized in Figure 2.

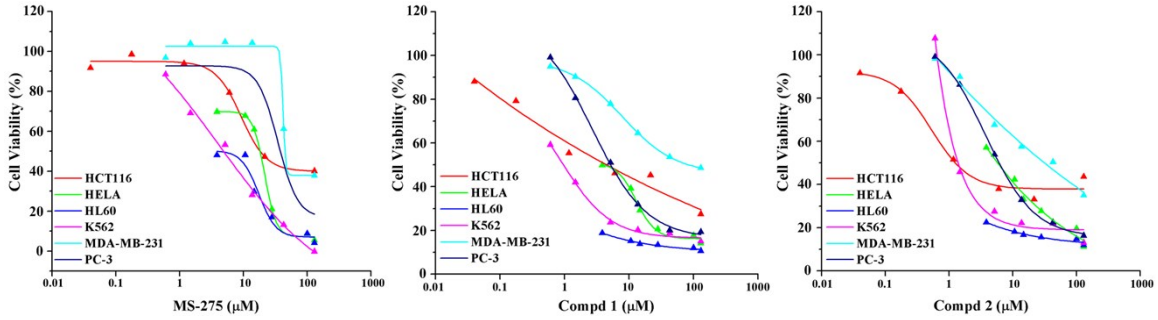


Figure S14. The antiproliferative activities of compounds against several tumor cell lines with MS-275 as positive control.

Table S1. The inhibition of the Biaryl Benzamide against wild type HDAC2/7 and mutant HDAC7.

Compd	IC50 (**μM**)			
	Wild type HDAC2	Wild type HDAC7	E543M HDAC7	R665A HDAC7
Biaryl Benzamide	0.39	>100	1.09	0.87