Structure-based discovery of selective histone deacetylase (HDAC) 3 and 4 inhibitors

Rupesh Agarwal[1,2*], Pawat Pattarawat[3], Michael R. Duff[2], Hwa-Chain Robert Wang[3], Jerome Baudry[4], Jeremy C. Smith[1,2*]

1 UT/ORNL Center for Molecular Biophysics, Oak Ridge National Laboratory, TN

2 Department of Biochemistry & Cellular and Molecular Biology, University of Tennessee, Knoxville, TN

3 Department of Biomedical and Diagnostic Sciences, College of Veterinary Medicine, University of Tennessee, Knoxville, TN

4 Department of Biological Sciences, The University of Alabama in Huntsville, Huntsville, Alabama.

*Co-corresponding authors

Abstract

Histone deacetylases (HDACs) are a family of proteins that are very important cancer drug targets. Several FDA-approved drugs target the conserved catalytic pocket of HDACs and thus lack selectivity for subfamilies (classes) of HDAC. In this study, we aimed to target HDAC4 from Class IIA, which is linked to various common cancers. There are no known inhibitors reported to bind specifically to HDAC4. Using classical and accelerated molecular dynamics (MD) simulations, we identified novel pockets in the interface between HDAC4 and the protein NCOR, which are not observed in the available crystal structures. These pockets were then targeted using an ensemble docking approach combined with consensus scoring. Using this approach, we selected 18 compounds for *in vivo* testing on two cancer cell lines. Of these, 5 compounds decreased cell viability to less than 60%. One inhibited the catalytic activity of HDAC4 but not HDAC3, which belongs to a different family of HDACs (Class I). The most potent compound has an IC_{50} comparable to the FDA-approved compound SAHA (Vorinostat). Targeting pockets other than the active site is thus a viable strategy for developing selective HDAC inhibitors.

Introduction

Histone deacetylases (HDACs) catalyze the deacetylation of histone tail lysines, resulting in compaction of DNA and suppression of transcription[1]. HDACs work in opposition to histone acetyltransferase (HAT), which transfers an acetyl group to lysines in the histone tail[2]. In a healthy cell, the transcription levels regulated by HAT and HDAC are well balanced[1]. However, abnormality in the function of HDACs contributes to the initiation and progression of several cancers[1] and can cause abnormal transcription of critical genes that control vital cell functions, such as proliferation, cell cycle regulation, and apoptosis[3, 4]. HDACs also influence other important genomic functions such as DNA repair, chromatin assembly, and recombination[5].

HDACs have been assigned to four major classes based on their catalytic mechanisms and sequence homology: Class I (HDAC 1, 2, 3, 8), IIa (HDAC 4, 5, 7, 9), IIb (HDAC 6, 10) and IV (HDAC 11)[6]. These HDACs have a zinc (Zn^{+2}) ion in the catalytic pocket as a cofactor in enzymatic activity. In contrast Class III HDACs, or sirtuins,[7] require nicotine adenine dinucleotide as a cofactor. HDACs cannot bind to DNA by themselves and exist as components in various multiprotein complexes. Structures of the catalytic domains of all class I HDACs show only minor differences between them.

Unlike class I HDACs, class IIa proteins lack intrinsic deacetylase activity because a key tyrosine residue, which stabilizes the tetrahedral intermediate during deacetylation of the native substrate (acetylated lysine) is absent[8]. The catalytic domain contains two zincs: structural and catalytic. A structural zinc is one of the features distinguishing Class IIa from Class I HDACs. However, it is believed that class IIa HDACs, especially HDAC4, function by forming multiprotein complexes with other enzymatically active class I HDACs. HDAC4 forms a complex with HDAC3-Nuclear Receptor Corepressor (NCoR) and is believed to operate as a scaffold for the recruitment of multiprotein complexes, thus increasing the range of deacetylation in specific regions of chromatin[9].

HDAC inhibitors (HDACi) have emerged to be important drugs for cancer therapy, as they can inhibit upregulated HDACs, thus reducing histone deacetylation

leading to normal levels of acetylated histones[10]. To date, four HDACi: Vorinostat (SAHA), Belinostat (PXD101), Panobinostat (LBH589), Romidepsin (FK228) have been approved by the US FDA, and one HDACi, namely Chidamide (CS005) by the China FDA, for the treatments of hematologic malignancies, and more than 20 inhibitors are in clinical trials [11, 12]. All these US FDA-approved inhibitors target the HDAC catalytic pocket and share common structural features, including a Zinc-binding group (ZBG), a linker region, and a surface recognition domain[12, 13]. Based on the ZBG chemical structures, HDACis can be divided into five main classes: hydroxamates, cyclic peptides, benzamides, short-chain fatty acids, ketones, and others. However, most hydroxamates are pan-HDACis, while the benzamides have increased class I selectivity. Three (Vorinostat, Belinostat, Panobinostat) out of the four US FDA-approved drugs are hydroxamate inhibitors that target the evolutionarily-conserved catalytic pocket, thus explaining why these inhibitors are nonspecific binders and potentially toxic[12].

For the structure-based design of specific inhibitors, there is a need to focus on the structural differences between HDACs [12, 14]. HDAC4 has been implicated in promoting tumor growth through the suppression of p21 expression in colon cancer, glioblastoma, ovarian cancer, and gastric cancer cells[15] and therefore is a potential drug target for anti-cancer therapy. However, there are no known inhibitors that target HDAC4 selectively. Here, we aimed to use structure-based drug discovery to identify novel inhibitors that selectively bind to HDAC4. Structure-based design is possible since there are two reported crystal structures of the HDAC4 catalytic domain, in open and closed conformations[8]. In the structure of the open conformation a small molecule pan inhibitor is bound in the catalytic pocket, whereas in the closed conformation there is a gain of function mutation of H332Y. The region bound to the structural zinc is closer to the active site in the closed conformation

Most of the enzymatic activity associated with HDAC4 expressed in mammalian cells is due to endogenous HDAC3. HDAC4 can shuttle between the cytoplasm and the nucleus and bind to NCoR, which is also bound to enzymatically active HDAC3[16]. It has been previously reported that HDAC4 binds to repression domain 3 of NCoR, whereas

HDAC3 binds to the SANT domain of NCoR[16, 17]. There are no structures of the HDAC4-NCoR complex, but based on mutational studies [18] 21 “hot spot” residues (C667, C669, C751, D759, T760, S767, A774, P799, P800, G801, H803, A804, F812, C813, H842, H843, G844, N845, G846, G868, and F871) present on the surface of HDAC4 have been identified as preventing binding to the NCoR-HDAC3 complex in vivo (**Figure 1**). These interfacial residues are conserved across all class IIa HDACs and include the residues in the opening of the catalytic pocket. In contrast, some of these residues (C667, C669, C751, D759, T760, and F871) are only found in class IIa HDACs and not in class I, which potentially forms the basis for interaction specificity with NCoR.

In an attempt to engineer selectivity our strategy focused on the HDAC4 interface with NCoR rather than targeting the conserved catalytic pocket. Protein-protein interfaces (PPIs) have emerged as major drug targets since a large number of proteins critical in biological pathways related to various diseases function after complex formation[19]. However, there are major challenges in targeting PPI interfaces since they often lack cavities for small molecules to bind or are intrinsically disordered[20, 21]. Nevertheless, small molecules interacting at interface “hot-spots” can compete with the binding of the cognate partner without necessarily covering the whole PPI surface[22-24].

We used classical molecular dynamics (MD) and enhanced sampling accelerated molecular dynamics (aMD) simulation[25] approaches to generate an ensemble of structures and to identify potential pockets in the experimentally characterized interface of HDAC4 and NCoR. Using these techniques, we identified novel pockets in the PPI which were not present in the crystal structures. We hypothesized that small molecules binding in these pockets could lead to disruption of complex formation. We performed high throughput *in silico* screening of small molecule libraries on the ensemble of structures generated from the simulation. Using this ensemble docking approach[26]., we identified 18 compounds that were tested in cell viability assays, five of which decreased cell viability to less than 60%. One of these compounds inhibited the catalytic activity of HDAC4 but not HDAC3 while, serendipitously, another inhibited the catalytic activity of HDAC3 but not HDAC4.

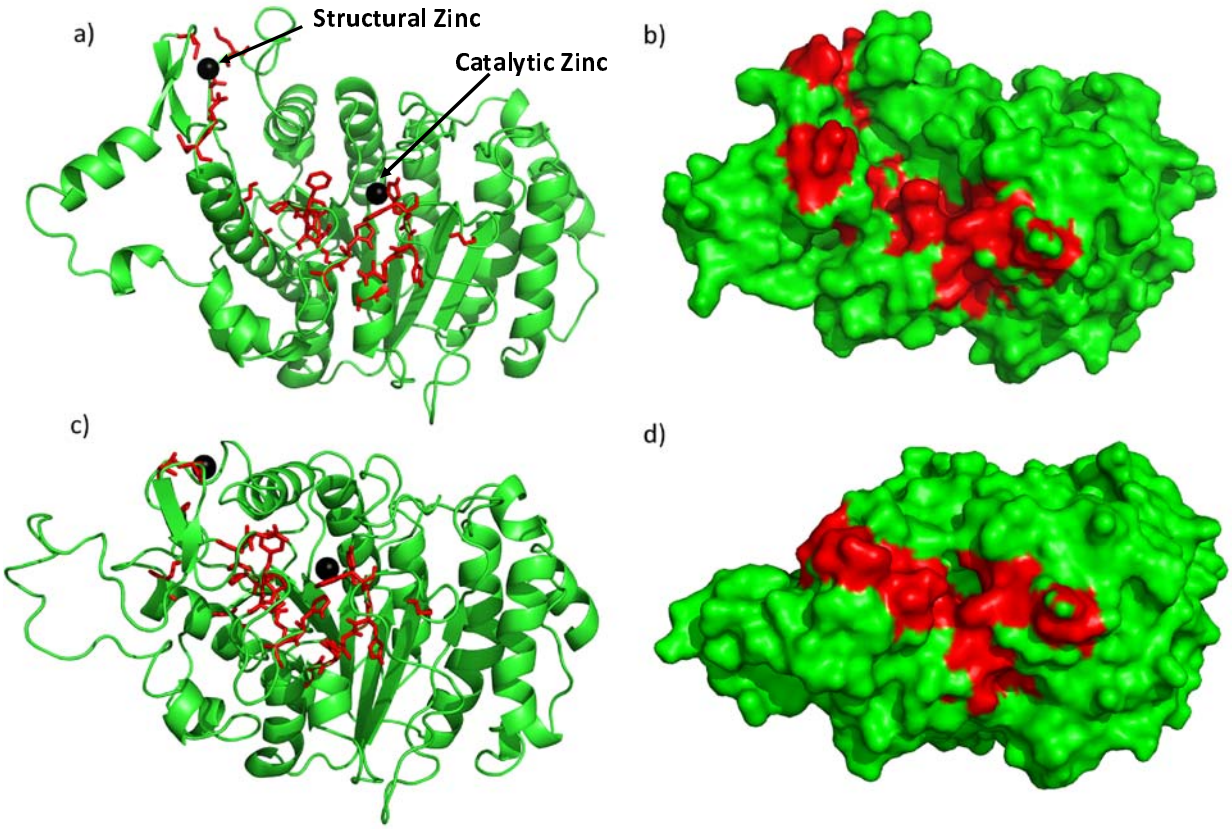


Figure 1: Crystal structure of the open conformation[8] (a & b) and the closed conformation[8] (c & d) of HDAC4. The red sticks and red region represent residues classified as the "hot spot" residues identified previously[18]; black spheres represent the two zincs (catalytic & structural). a & c are ribbon models of the crystal structures, and b & d are surface models.

Methods

System preparation

Crystal structures of HDAC4 in the open (PDB 2VQM)[8] and in the closed state (PDB 2VQW)[8] were used for this study. The protein structures was modeled using Amber ff14SB parameters[27, 28] and and water was modeled as TIP3P[29]. Using xleap, cysteines 667, 669, and 751 were deprotonated (-ve charge) and converted to CYM for the closed conformation, and cysteines 667 and 751 were converted to CYM for the open conformation as observed in the respective crystal structure. The open structure was crystallized in the presence of an inhibitor, which was deleted for the simulation. The closed structure was crystallized with a mutation H332Y, and so the tyrosine 332

was mutated to wild type histidine for the simulation. The protein was surrounded by an octahedral box of water model TIP3P[29] box of 10 Å around the protein in each direction.

Classical molecular dynamics simulation

MD simulations were performed on both crystal structures. First, the HDAC4 structure was held fixed with force constant of 500 kcal mol^{-1} $Å^{-2}$ while the system was minimized with 500 steps of steepest descent followed by 500 steps with the conjugate gradient method. In the second minimization step, the restraints on HDAC4 were removed, and 1000 steps of steepest descent minimization were performed, followed by 1500 steps of a conjugate gradient. The system was heated to 300 K while holding the protein fixed with a force constant of 10 kcal mol^{-1} $Å^{-2}$ for 1000 steps. Then, the restraints were released, and 1000 MD steps were performed. The SHAKE algorithm[30] was used to constrain all bonds involving hydrogen in the simulations. A 50 ns MD was performed at 300 K using the NPT ensemble and a 2 fs time step. The temperature was fixed with the Langevin dynamics thermostat[31] and the pressure was fixed with a Monte Carlo barostat[32]. This procedure yielded a total of 250,000 snapshots for subsequent analyses.

Accelerated molecular dynamics simulation

Accelerated molecular dynamics (aMD)[25] samples conformational space more efficiently than classical MD, resulting in receptor conformers which would not have been otherwise sampled. Unlike other enhanced sampling methods, aMD does not require reaction coordinates or collective variables and it is well suited for creating an ensemble of structures to characterize the conformational flexibility of a receptor. The aMD simulation was performed by using a dual energy boost (i.e. dihedral energy and potential energy) with an acceleration parameter (α) of 0.2. The average total potential energy and average dihedral energy parameters for aMD were calculated from the classical MD simulation. We performed 100 ns of aMD simulation. This procedure yielded a total of 500,000 snapshots for subsequent analyses. Snapshots from classical MD and aMD were clustered by root mean square deviation (RMSD) of hot spot residues with the hierarchical agglomerate clustering algorithm present in the Cpptraj module[33].

In silico screening

We generated an ensemble of 4 representative snapshots from each of the classical MD and aMD clusters for both open and closed conformations. The NCI Diversity V set (~1600 compounds)[34] was docked to a total of 9 receptor conformations (4 snapshots from MD, 4 from aMD and 1 crystal structure) with a cubic box size of ~30 Å. VinaMPI[35], a parallel version of AutodockVina[36], was used to perform the *in silico* screening. The docked poses were then ranked by the AutodockVina scoring function[37]. The poses of the top 100 compounds were manually checked and 18 compounds were suggested for experimental testing.

Cell viability assay

Cell viability assays were determined using Methyl Thiazolyl Tetrazolium (MTT) assay (Travigen, Gaithersburg, MD, USA). Five × 10^3 SW780 (human transitional urinary bladder cancer cell line) or 1 x 10^4 (human triple negative breast cancer cell line MDA-MB231, American Type Culture Collection [ATCC], Rockville, MD) in 100 µL of culture media were seeded to each well of 96-well culture plates. At 24 hours after seeding, cells were treated with various concentrations of potential HDAC inhibiting compounds for 48 hours (100 µM for initial screening, 0.01, 0.01, 1, 10, 100, 250, and 500 µM for dose-response experiment). Subsequently,10 µL of MTT reagent were added to each well and incubated for 4 hours, followed by addition of 100 µL of detergent reagent. The plate was incubated overnight. The reduced MTT reagent in cultures was measured at 570 nm by using an ELISA reader (Bio-Tek, Winooski, VT, USA). The IC50 for each compound was calculated from fitted non-linear regression dose-response curves.

HDAC activity assay

HDAC3 and HDAC4 activity assays were performed using fluorogenic assay kits obtained from BPS Bioscience (San Diego, CA). Compounds were prepared as 10 mM stocks in DMSO and were screened at 100 µM in the assays. The activity of the enzymes in the absence and presence of compounds was measured on a BioTek Cytation 5 plate reader, exciting the samples at 350 nm and measuring emission at 450

nm. IC50 values were determined for the compounds that decreased HDAC activity by at least 25% relative to no inhibitor.

Results

Novel pockets from MD simulations and ensemble docking

The PPI residues on HDAC4 encompass the catalytic pocket and structural zinc and the region between them. In the crystal structures, this region is rather smooth and has no druggable cavity for a small molecule to bind, which is a common drawback in targeting PPIs[20, 21]. However, in the ensemble of receptor structures generated from classical MD and aMD simulations for both open and closed conformations, we identified novel cavities/pockets in the PPI which were not present in the crystal structures (**Figure 2**). These cavities are present in the “hot-spot” region also provide pockets for small molecules to bind. These pockets were generated by clustering of the trajectory, and hence are present for significant proportions of time. Therefore, are low energy states and not rare conformations. We observed some shallow cavities (**Figure 2**), and a deep pocket near the structural zinc, which contains PPI hot spot residues as shown in **Figure 2**. These pockets stretch throughout the PPI region. These results illustrate the usefulness of docking to an ensemble of structures rather than to only the crystal structure.

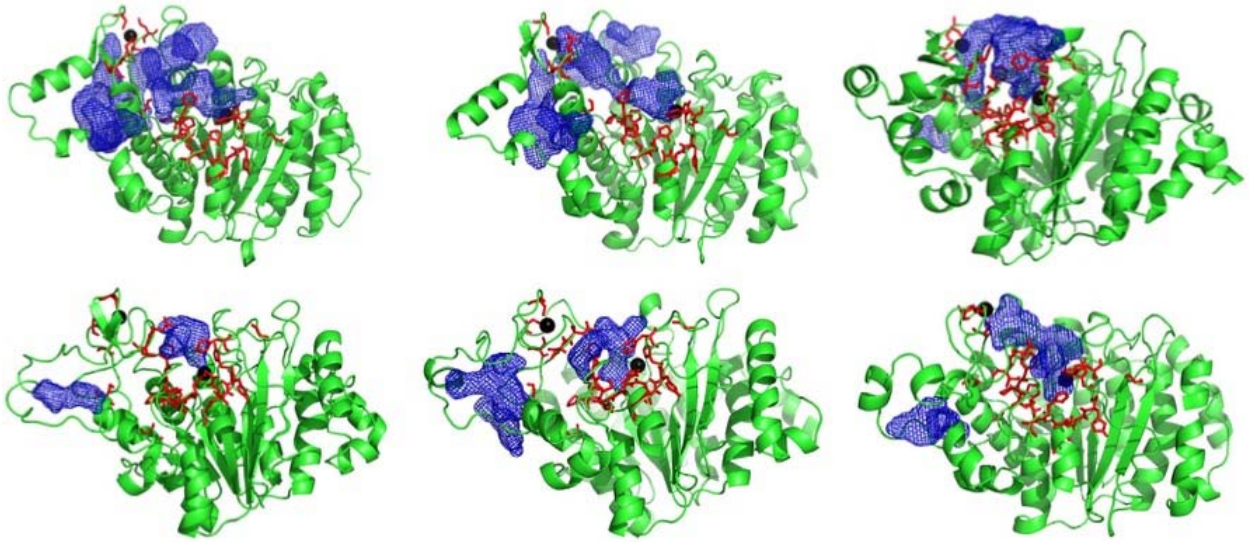


Figure 2: HDAC4 ensemble structures showing novel pockets/cavities (blue mesh) generated from simulations of crystal structures. Red sticks represent PPI "hot spot" residues identified experimentally.

The docking search box was set to be large enough to include both the catalytic pocket and the novel pockets identified, as both are in the PPI region. Since the catalytic pocket itself is large and hydrophobic, we found that many of the compounds ended up in the catalytic pocket. Therefore, we filtered out the compounds which were within 10 Å of the catalytic zinc. The aim of this approach was to identify compounds that bind with a high score in the PPI. We found small molecules bound in the novel pockets discovered (**Figure 3**). These docked poses were further refined in their respective pockets using the 'induced fit' tool in MOE[38]. After manually examining the docked poses, we picked 18 high-scoring compounds chosen to maximize the pocket coverage (**Figure S2**), for chemical diversity (**Figure S1**), and for the availability of compounds at NCI. These compounds were then suggested for experimental testing.

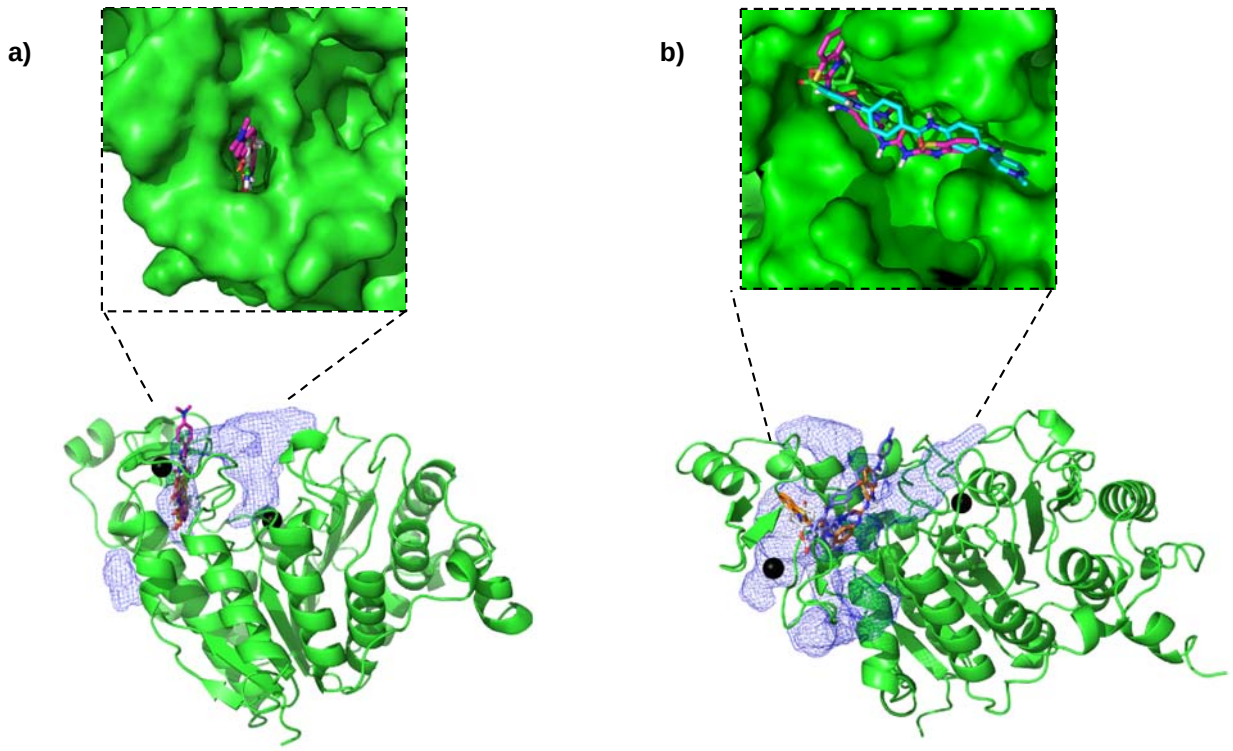


Figure 3: Some of the top ranked compounds bound in the novel pockets identified in a) closed conformation and b) open conformation ensemble structure

Cell susceptibility to identified small molecules

To test the in silico derived compounds in vivo, we selected 18 NCI compounds (NCI IDs: 117922, 134199, 319435, 34488, 36425, 373535, 4135, 51936, 55172, 60034, 79887, 88402, 42231, 11926, 195327, 299968, 44584, 67436) for testing on human breast carcinoma MDA-MB231 and transitional urothelial carcinoma SW780 cell lines. Both these cell lines have been shown to significantly upregulate HDACs[39, 40]. These 18 compounds were tested in a cell viability assay at 100 µM, with DMSO as a negative control and SAHA (Vorinostat, a US FDA-approved drug)[41] as a positive control. 5 compounds: 67436, 11926, 299968, 44584 and 195327, were able to reduce cell viability of MDA-MB231 to lower than 60% (**Figure 4a**) and were considered as active. Moreover, 67436, 11926, and 195327 were able to reduce SW780 to lower than 50%, indicating that these three compounds are effective against both cancer cell lines. Next, compounds (195327, 44548, 299968, 67436, 11926) that reduce cell viability to less than 60% were further examined by dose-response exposure of each compound at

0.01, 0.01, 1, 10, 100, 250, and 500 µM, and the IC_{50} of each compound calculated. A concentration-dependent decrease of cell viability was observed for all the selected compounds (**Figure 4b**). The most potent compound is 195327, which has an IC_{50} comparable to SAHA at 37±17 µM vs. 35±10 µM for MDA-MB231 and 89±48 vs. 90±45 for the SW780 cell line. Data also showed that 299968 and 44584 were effective against SW780 but at higher doses (IC50 = 300±17 and 213±13, respectively). Compound 11926 reduced cell viability of both MDA-MB231 and SW780, but the efficacy against SW780 was lower than MDA-MB231(Cell viability = 45% at 500 µM in SW780, compared to 9% at 500 µM in MDA-MB231). These results indicated that the compounds are effective against both cancer cell lines, but with different levels of response.

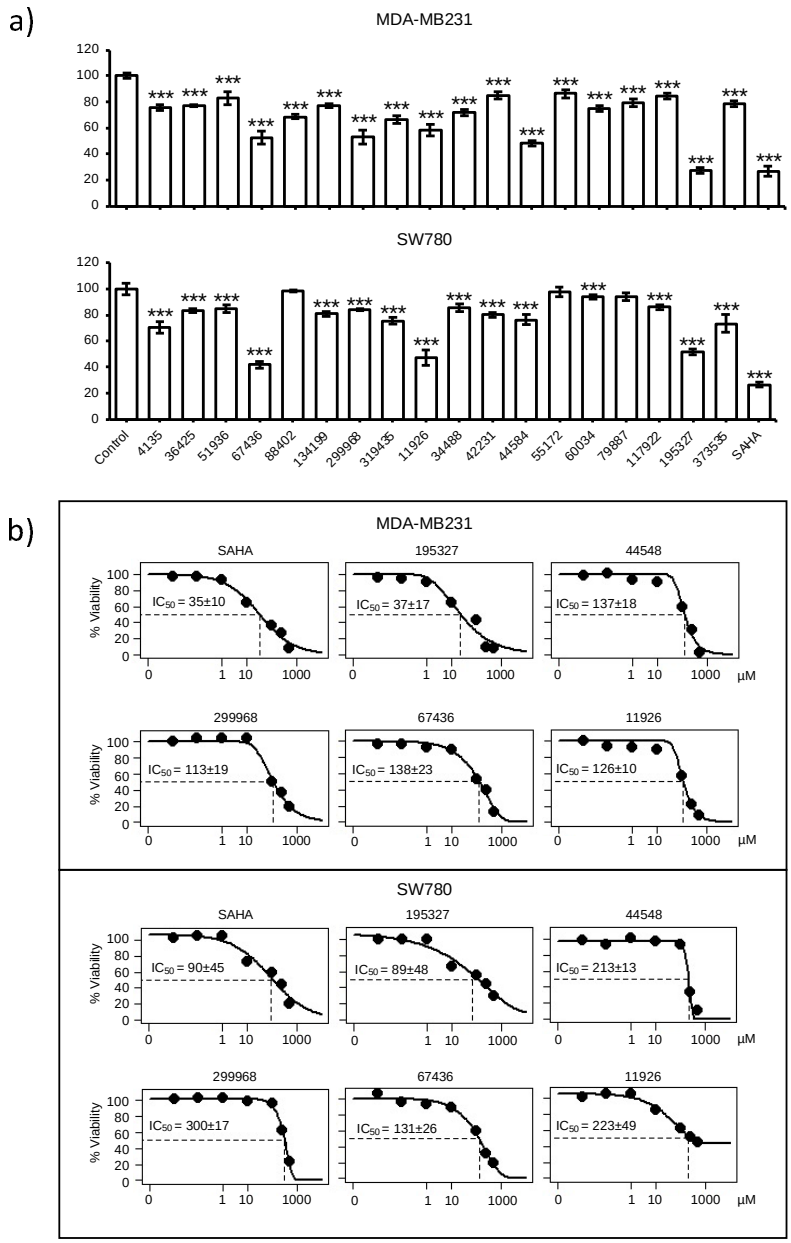


Figure 4. Effect of compounds on cell viability of cancer cell lines. a) Relative cell viability of MDA-MB231 and SW780 cell lines 48 hours after treatment with candidate compounds determined by MTT assay. Relative cell viability was normalized by the value determined in untreated counterpart cells set as 100%. Columns, mean of triplicates; bars, SD. Statistical significance was determined using student's t-test (a, e, & f), indicated by ***p < 0.001. b) Dose-response exposure of compounds that reduced cell viability to less than 60%. IC_{50} of each compound was calculated.

HDAC Activity

With the results from the cell viability assays, we decided to dig deeper and performed enzyme assays. Since the cell lines used also contain upregulated class I HDACs, the 18 compounds were screened for their ability to inhibit HDAC4 and HDAC3 (from class I HDAC subfamily) (**Figure 5**) *in vitro*. At 100 μM, two compounds (67436 and 134199) decreased both HDAC3 and HDAC4 activities to 40%, or less, and a number of other compounds were specific to one of the enzymes. HDAC4 was significantly inhibited by 88402 (<40% activity retained), but the compound only mildly perturbed HDAC3 activity (75% activity); however, that compound did not show activity in the in-vivo assay (**Figure 4a**). Compounds 34488, 44584, and in vivo active compound 195327 decreased HDAC3 activity by 50%, 60%, and 80%, respectively, whereas HDAC4 activity remained unchanged in the presence of those compounds (**Figure 5**).

The compounds that decreased the activity of HDAC4 by at least 25% were further screened for IC50 values (**Table 1, Figure S3, S4, S5**). Compounds 67436 and 134199 had IC50 values in the low micromolar range for both HDACs, which is again comparable to the known inhibitor of HDAC4, SAHA. However, both compounds were significantly weaker inhibitors of HDAC3 than SAHA and thus more selective. Compounds 88402, 319435, and in vivo active compound 299968 had IC50 values at least 2-fold lower for HDAC4 than HDAC3, showing that these compounds are potentially also HDAC4-specific inhibitors. Both 299968 and 319435 decreased the activity of HDAC3 and HDAC4 to similar levels in the initial compound screen (**Figure 5)**. However, these compounds did not inhibit HDAC3 up to a concentration of 500 μM in the IC50 experiments, indicating they most likely do not inhibit HDAC3 **(Figure S3, S4, S5**). The higher errors in the screening assay for HDAC3 (**Figure 5**) with compounds 299968 and 319435 may have masked differences in inhibition of the two enzymes by these compounds. The other compounds that weakly inhibited HDAC4 (4135, 36425, and 51936) had IC50 values not much lower than the highest concentration used in the assay (500 μM) and may not be significant inhibitors of either enzyme. Also, while compound 195327 yielded promise as an inhibitor of HDAC3 in the initial compound screen, no inhibition was noted in the IC50 experiments. Also, 195327 did not inhibit HDAC4 in either assay suggesting the inhibition of HDAC3 by the

compound noted in the initial screen may have been artefactual. Interestingly, compound 134199 was one of the better inhibitors of both HDACs in the enzyme assay but had little potency in the in vivo screen. From these results, we have identified 8 compounds that bind to HDAC4 or HDAC3 and were able to inhibit the deacetylase activity in the assay. Out of these 8, 3 compounds (88402, 299968, and 319435) bind selectively to HDAC4 according to IC50 data, and one compound (67436) was slightly more for HDAC3.

Table 1. IC50 values for compounds that inhibited HDAC3 and HDAC4

Compound	IC50 (μM) for HDAC3	IC50 (μM) for HDAC4
SAHA	0.024 ± 0.006	25±13
4135	>500	410 ± 240
11926	>500	ND
34488	>500	ND
36425	>500	480 ± 260
44584	>500	ND
51936	>500	550 ± 350
55172	>500	ND
67436	9.2 ± 5.2	47±28
79887	>500	ND
88402	210 ± 60	100±40
134199	25 ± 7	33±12
195327	>500	ND
299968	>500	250 ± 170
319435	>500	150 ± 50

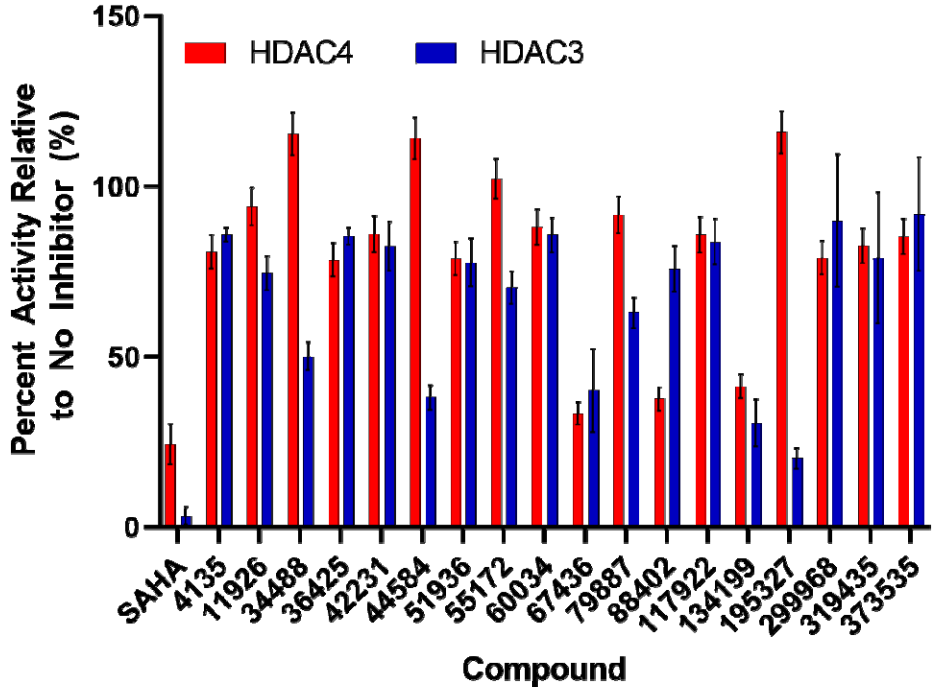


Figure 5. Effect of compounds on the activities of HDACs. The relative percent activity of HDAC3 (blue bars) and HDAC4 (red bars) in the presence of 100 μM inhibitor compound compared to the absence of inhibitor. Error bars are the standard deviation of at least two assays.

Discussion

HDACs are cancer drug targets, but most of the FDA-approved HDAC drugs are pan HDAC inhibitors and selectivity towards a specific sub-family is difficult to achieve. Hence, in this study, we moved away from the common strategy of targeting the catalytic pocket and focused on a protein-protein interface. HDAC4 complexation with NCoR and HDAC3 is critical for scaffolding function and disrupting the complex can lead to inhibition. However, the PPI interfaces present in the crystal structures, like most PPIs, are smooth and do not have any cavities/pockets that can be targeted by a small molecule. Therefore, we performed classical and enhanced sampling MD simulations to identify different conformations of the PPI. Indeed, we were able to identify pockets in the PPI region that have not been characterized before and are not present in either of the crystal structure. Using ensemble docking, we targeted these novel pockets present in the PPI region. To further improve the hit rate and reduce the number of false-

positives, we performed an ‘induced fit’ refinement. The aim of the study was to find a novel hit molecule with a novel scaffold that can be optimized to make a selective inhibitor.

By combining computational and experimental methods, we were able to identify compounds that reduce cell viability in human breast carcinoma MDA-MB231 and transitional urothelial carcinoma SW780 cell lines as well as in in vitro assays. NCI compounds 195327, 44548, 299968, 67436, and 11926 showed a reduction in cell viability to below 50%. These five compounds have IC50s comparable the FDA-approved drug SAHA (Vorinostat) but very different scaffolds than both SAHA and the other 4 known drugs when compared using MACCS fingerprinting (**Figures S6, S7)**. These 18 compounds were then tested in vitro for HDAC activity. In vitro activity assays indicate that 67436 inhibits the activity of both HDAC4 and HDAC3 by ~50% inhibition at 100 μM, whereas both 44584 and 195327 show higher inhibition for HDAC3 than HDAC4. Interestingly, some compounds like 134199 had very high inhibition (~40%) for both HDAC4 and HDAC3, while 34488 and 88402 had selective inhibition for HDAC3 and HDAC4, respectively. Dose-dependent assays suggested that 88402, 299968, and 319435 binds selectively to HDAC4 with IC50s of 100±40 μM, 250 ± 170 μM, and 150 ± 50 μM, respectively, while 67436 was selective for HDAC3 with an IC50 of 9.2 ± 5.2 μM. 195327, 44548, and 11926 showed high activity in cell viability assay but were not as active in the in vitro assay. We hypothesize that these molecules potentially bind in the PPI region of HDAC4, a hypothesis that could not be tested using the activity assay.

The results highlight the usefulness of ensemble docking in early-stage drug discovery and identifying molecules with novel chemical matter. Starting with the goal of finding selective HDAC4 inhibitors we serendipitously also found some HDAC3 selective inhibitors. These compounds are a good starting point for hit expansion, lead optimization steps, or degrader (PROTACS) strategy[42]. Additional characterization using structural biological techniques such as X-ray crystallography would be useful in guiding lead optimization and associated efficacy improvement.

References:

1. Lawlor, L.; Yang, X. B., Harnessing the HDAC–histone deacetylase enzymes, inhibitors and how these can be utilised in tissue engineering. *International journal of oral science* **2019,** *11* (2), 20-11.
2. Berger, S. L., The complex language of chromatin regulation during transcription. *Nature* **2007,** *447* (7143), 407-12.
3. Di Gennaro, E.; Bruzzese, F.; Caraglia, M.; Abruzzese, A.; Budillon, A., Acetylation of proteins as novel target for antitumor therapy: review article. *Amino acids* **2004,** *26* (4), 435-41.
4. Bolden, J. E.; Peart, M. J.; Johnstone, R. W., Anticancer activities of histone deacetylase inhibitors. *Nature reviews. Drug discovery* **2006,** *5* (9), 769-84.
5. Stengel, K. R.; Hiebert, S. W., Class I HDACs Affect DNA Replication, Repair, and Chromatin Structure: Implications for Cancer Therapy. *Antioxid Redox Signal* **2015,** *23* (1), 51-65.
6. RUIJTER, A. J. M. d.; GENNIP, A. H. v.; CARON, H. N.; KEMP, S.; KUILENBURG, A. B. P. v., Histone deacetylases (HDACs): characterization of the classical HDAC family. **2003**.
7. Finkel, T.; Deng, C. X.; Mostoslavsky, R., Recent progress in the biology and physiology of sirtuins. *Nature* **2009,** *460* (7255), 587-91.
8. Bottomley, M. J.; Lo Surdo, P.; Di Giovine, P.; Cirillo, A.; Scarpelli, R.; Ferrigno, F.; Jones, P.; Neddermann, P.; De Francesco, R.; Steinkuhler, C.; Gallinari, P.; Carfi, A., Structural and functional analysis of the human HDAC4 catalytic domain reveals a regulatory structural zinc-binding domain. *The Journal of biological chemistry* **2008,** *283* (39), 26694-704.
9. Di Giorgio, E.; Gagliostro, E.; Brancolini, C., Selective class IIa HDAC inhibitors: myth or reality. *Cellular and molecular life sciences : CMLS* **2015,** *72* (1), 73-86.
10. Minucci, S.; Pelicci, P. G., Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. *Nature reviews. Cancer* **2006,** *6* (1), 38-51.
11. Suresh, P. S.; Devaraj, V. C.; Srinivas, N. R.; Mullangi, R., Review of bioanalytical assays for the quantitation of various HDAC inhibitors such as vorinostat, belinostat, panobinostat, romidepsin and chidamine. *Biomedical chromatography : BMC* **2017,** *31* (1).
12. Zhao, C.; Dong, H.; Xu, Q.; Zhang, Y., Histone deacetylase (HDAC) inhibitors in cancer: a patent review (2017-present). *Expert Opin Ther Pat* **2020,** *30* (4), 263-274.
13. Xu, W. S.; Parmigiani, R. B.; Marks, P. A., Histone deacetylase inhibitors: molecular mechanisms of action. *Oncogene* **2007,** *26* (37), 5541-5552.
14. Ho, T. C. S.; Chan, A. H. Y.; Ganesan, A., Thirty Years of HDAC Inhibitors: 2020 Insight and Hindsight. *Journal of medicinal chemistry* **2020,** *63* (21), 12460-12484.
15. Guerriero, J. L.; Sotayo, A.; Ponichtera, H. E.; Castrillon, J. A.; Pourzia, A. L.; Schad, S.; Johnson, S. F.; Carrasco, R. D.; Lazo, S.; Bronson, R. T.; Davis, S. P.; Lobera, M.; Nolan, M. A.; Letai, A., Class IIa HDAC inhibition reduces breast tumours and metastases through anti-tumour macrophages. *Nature* **2017,** *543* (7645), 428-432.
16. Hudson, G. M.; Watson, P. J.; Fairall, L.; Jamieson, A. G.; Schwabe, J. W. R., Insights into the Recruitment of Class IIa Histone Deacetylases (HDACs) to the SMRT/NCoR Transcriptional Repression Complex. **2015**.
17. Fischle, W.; Dequiedt, F.; Hendzel, M. J.; Guenther, M. G.; Lazar, M. A.; Voelter, W.; Verdin, E., Enzymatic Activity Associated with Class II HDACs Is Dependent on a Multiprotein Complex Containing HDAC3 and SMRT/N-CoR. *Molecular Cell* **2002,** *9* (1), 45-57.

18. Kim, G. S.; Jung, H. E.; Kim, J. S.; Lee, Y. C., Mutagenesis Study Reveals the Rim of Catalytic Entry Site of HDAC4 and -5 as the Major Binding Surface of SMRT Corepressor. *PLoS One* **2015,** *10* (7), e0132680.
19. Jubb, H.; Blundell, T. L.; Ascher, D. B., Flexibility and small pockets at protein-protein interfaces: New insights into druggability. *Progress in biophysics and molecular biology* **2015,** *119* (1), 2-9.
20. Laraia, L.; McKenzie, G.; Spring, D.; Venkitaraman, A.; Huggins, D., Overcoming Chemical, Biological, and Computational Challenges in the Development of Inhibitors Targeting Protein-Protein Interactions. *Chem Biol* **2015,** *22* (6), 689-703.
21. Shin, W.-H.; Kumazawa, K.; Imai, K.; Hirokawa, T.; Kihara, D., Current Challenges and Opportunities in Designing Protein–Protein Interaction Targeted Drugs . *Advances and Applications in Bioinformatics and Chemistry* **2020,** *Volume 13*, 11-25.
22. Ran, X.; Gestwicki, J. E., Inhibitors of protein–protein interactions (PPIs): an analysis of scaffold choices and buried surface area. *Current Opinion in Chemical Biology* **2018,** *44*, 75-86.
23. Wang, X.; Ni, D.; Liu, Y.; Lu, S., Rational Design of Peptide-Based Inhibitors Disrupting Protein-Protein Interactions. *Front Chem* **2021,** *9*, 682675.
24. Lu, H.; Zhou, Q.; He, J.; Jiang, Z.; Peng, C.; Tong, R.; Shi, J., Recent advances in the development of protein–protein interactions modulators: mechanisms and clinical trials. *Signal Transduction and Targeted Therapy* **2020,** *5* (1).
25. Hamelberg, D.; Mongan, J.; McCammon, J. A., Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. *The Journal of Chemical Physics* **2004,** *120* (24), 11919-11929.
26. Amaro, R. E.; Baudry, J.; Chodera, J.; Demir, O.; McCammon, J. A.; Miao, Y.; Smith, J. C., Ensemble Docking in Drug Discovery. *Biophysical journal* **2018,** *114* (10), 2271-2278.
27. Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., The Amber biomolecular simulation programs. *Journal of computational chemistry* **2005,** *26* (16), 1668-88.
28. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. **2015**.
29. Price, D. J.; Brooks, C. L., 3rd, A modified TIP3P water potential for simulation with Ewald summation. *J Chem Phys* **2004,** *121* (20), 10096-103.
30. Miyamoto, S.; Kollman, P. A., Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of computational chemistry* **1992,** *13* (8), 952-962.
31. Zwanzig, R., Nonlinear generalized Langevin equations. *Journal of Statistical Physics* **1973,** *9* (3), 215-220.
32. Åqvist, J.; Wennerström, P.; Nervall, M.; Bjelic, S.; Brandsdal, B. O., Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm. *Chemical physics letters* **2004,** *384* (4-6), 288-294.
33. Roe, D. R.; Cheatham, T. E., 3rd, PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput* **2013,** *9* (7), 3084-95.
34. http://dtp.nci.nih.gov/branches/dscb/diversity_explanation.html.
35. Ellingson, S. R.; Tennessee, U. o. T. G. S. a. T. K.; Tennessee, O. R. N. L. U. O. C. f. M. B. O. R.; Smith, J. C.; Tennessee, O. R. N. L. U. O. C. f. M. B. O. R.; Tennessee, U. o. T. D. o. B. K.; Baudry, J.; Tennessee, O. R. N. L. U. O. C. f. M. B. O. R.; Tennessee, U. o. T. D. o. B. K., VinaMPI: Facilitating multiple receptor high-throughput virtual docking on high-performance computers. *Journal of Computational Chemistry* **2013,** *34* (25), 2212-2221.
36. Trott, O.; Department of Molecular Biology, T. S. R. I., La Jolla, California; Olson, A. J.; Department of Molecular Biology, T. S. R. I., La Jolla, California; Department of Molecular Biology, T. S. R. I., La Jolla, California, AutoDock Vina: Improving the speed and accuracy of docking with a new scoring

function, efficient optimization, and multithreading. *Journal of Computational Chemistry* **2009,** *31* (2), 455-461.
37. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry* **2009,** *30* (16), 2785-91.
38. Molecular Operating Environment (MOE), C. C. G. U., 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7 2020.
39. SY, P.; JA, J.; KJ, J.; HJ, H.; JS, S.; HY, L.; CG, P.; J, K., Histone deacetylases 1, 6 and 8 are critical for invasion in breast cancer. *Oncology reports* **2011,** *25* (6).
40. Buckwalter, J. M.; Chan, W.; Shuman, L.; Wildermuth, T.; Ellis-Mohl, J.; Walter, V.; Warrick, J. I.; Wu, X.-R.; Kaag, M.; Raman, J. D.; Degraff, D. J., Characterization of Histone Deacetylase Expression Within In Vitro and In Vivo Bladder Cancer Model Systems. *International Journal of Molecular Sciences* **2019,** *20* (10), 2599.
41. Marks, P. A., Discovery and development of SAHA as an anticancer agent. *Oncogene* **2007,** *26* (9), 1351-1356.
42. Sun, X.; Gao, H.; Yang, Y.; He, M.; Wu, Y.; Song, Y.; Tong, Y.; Rao, Y., PROTACs: great opportunities for academia and industry. *Signal Transduction and Targeted Therapy* **2019,** *4* (1).
43. Backman, T. W. H.; Cao, Y.; Girke, T., ChemMine tools: an online service for analyzing and clustering small molecules. *Nucleic acids research* **2011,** *39* (suppl), W486-W491.

Supplementary information

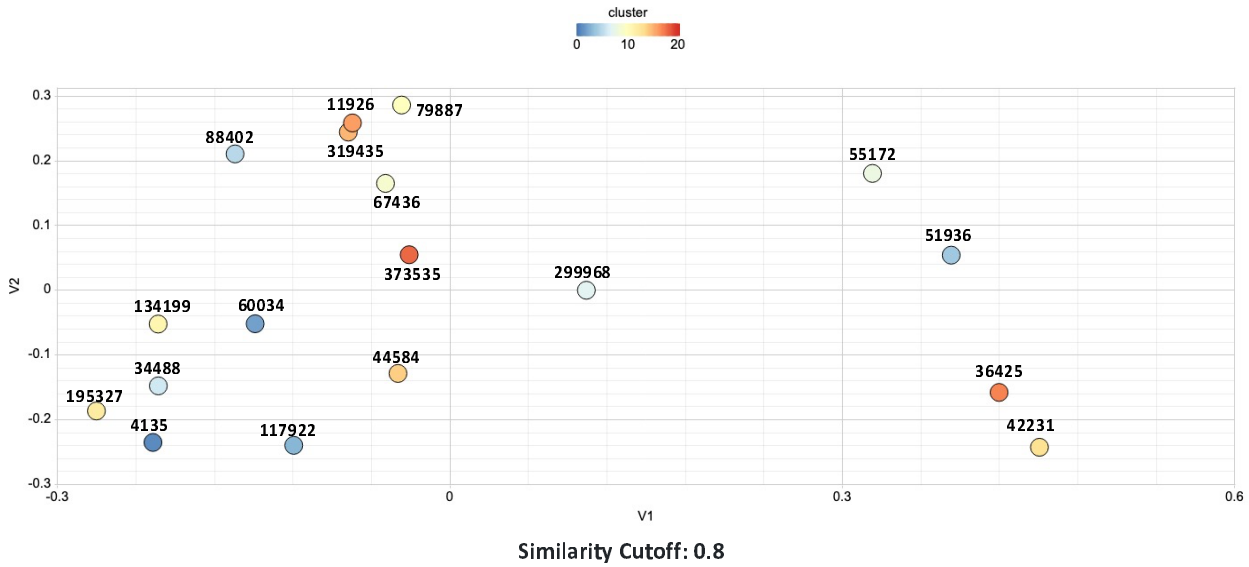


Figure S1: Clustering of compounds by structural and physicochemical similarities by Multidimensional Scaling (MDS) using ChemmineTools[43]. Similarity cutoff of 0.8 is used for coloring only.

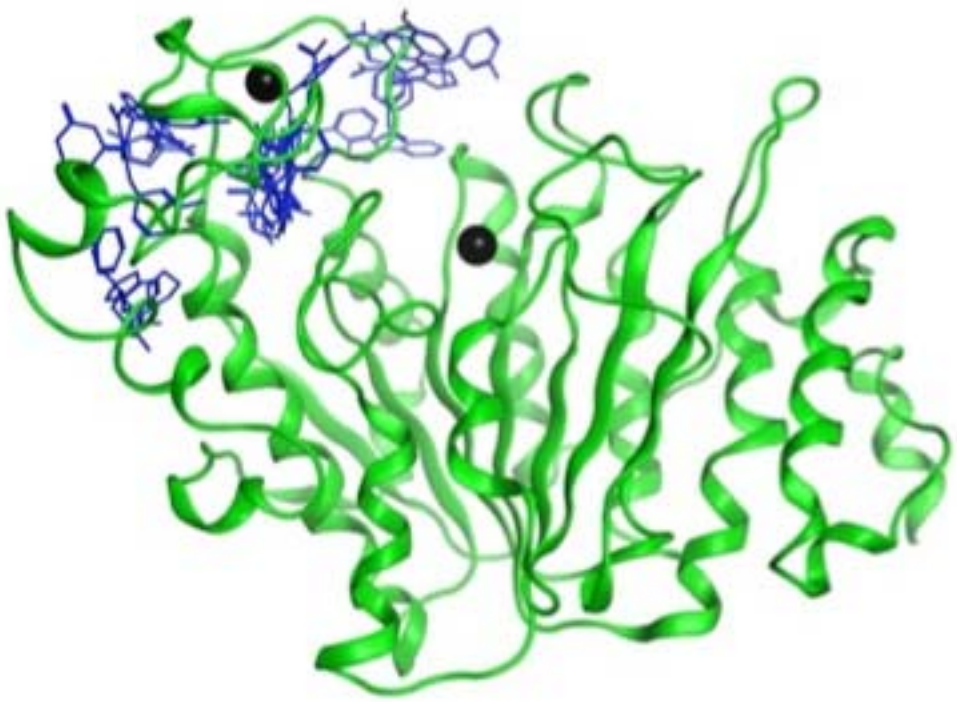


Figure S2: Some of the top hit molecules (in Blue) and there poses on HDAC4 (in green ribbon) with Zinc atoms (as black sphere) to shown binding site coverage. The figure was made after superimposition of the hits with their respective ensemble structure. All the ensemble structures and hits are not shown for clarity.

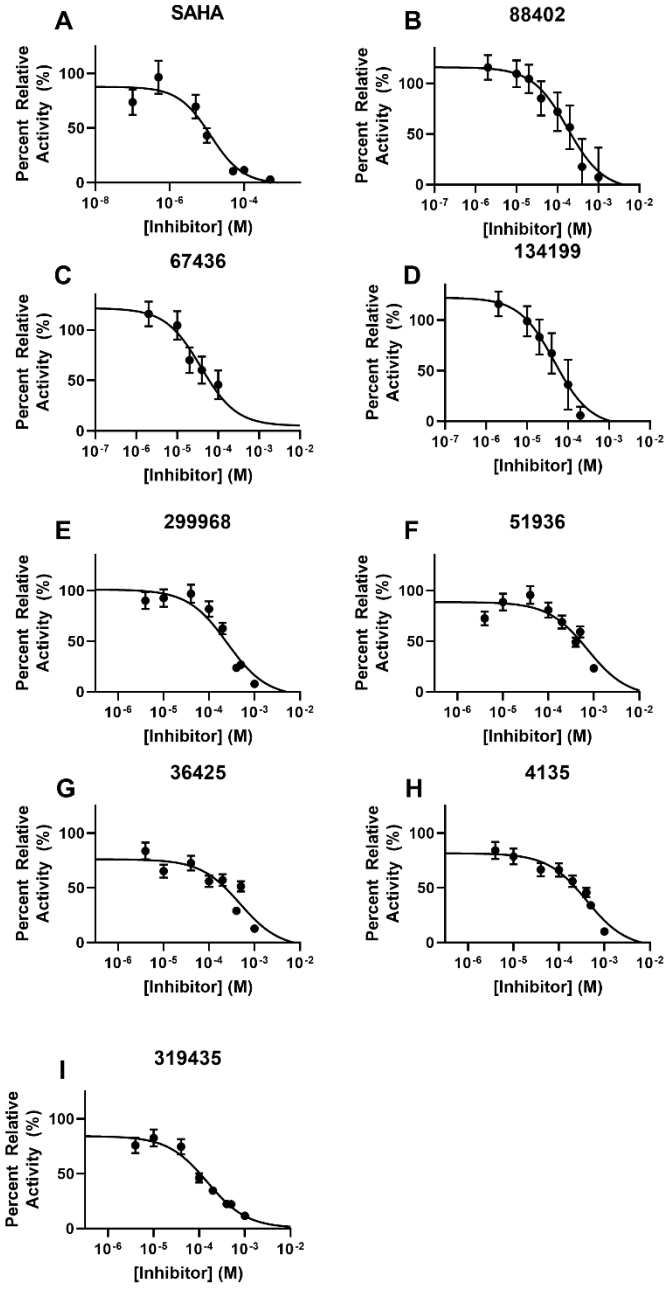


Figure S3. IC50 plots for compounds that inhibit HDAC4.

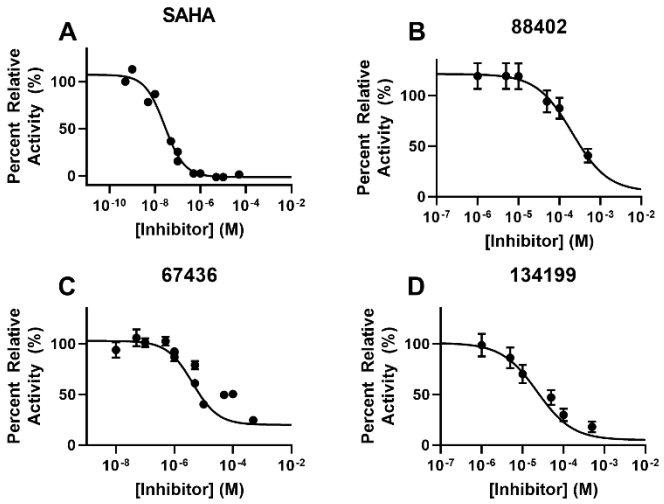


Figure S4. IC50 plots for compounds that inhibit HDAC3.

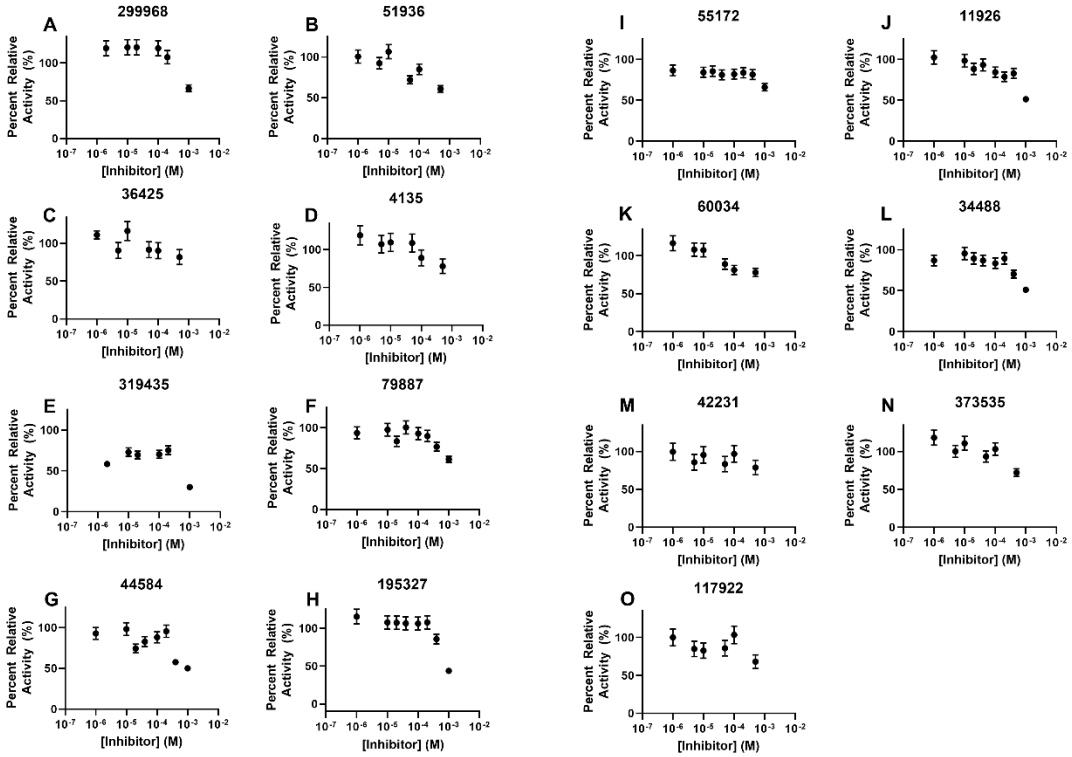


Figure S5. IC50 plots for compounds that yielded no inhibition of HDAC3 below 500 μM.

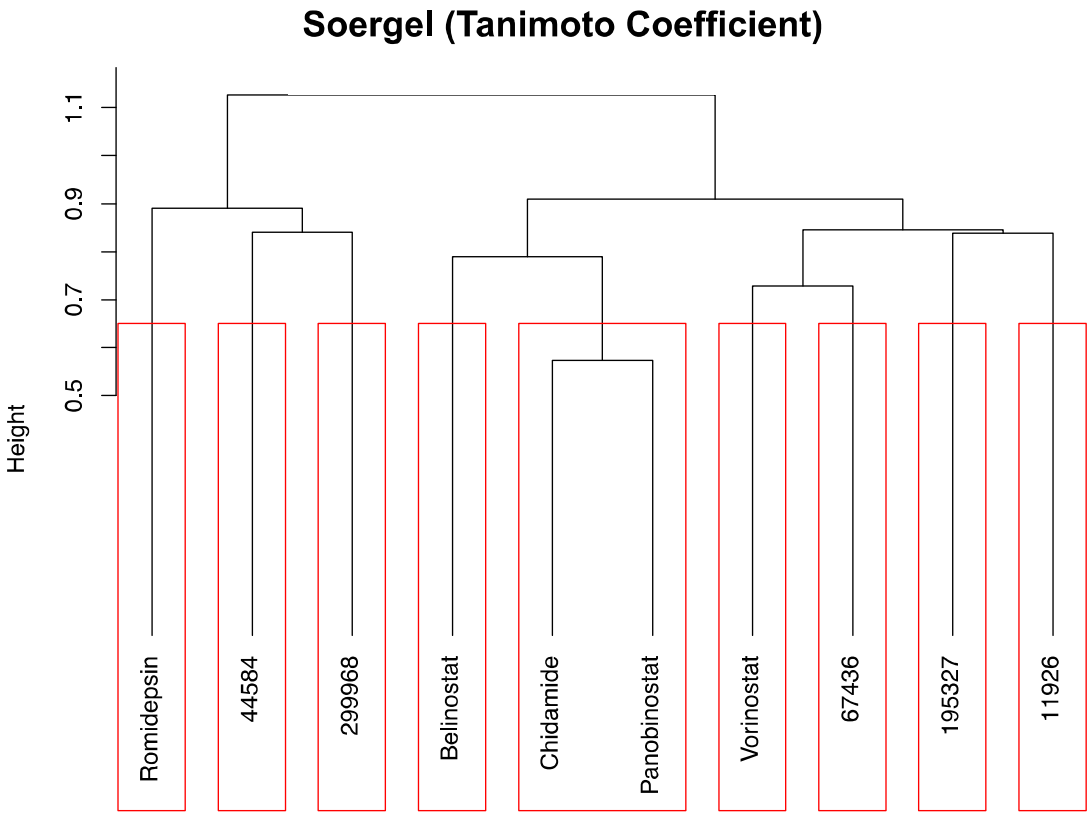
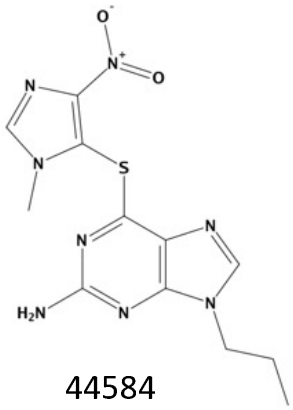
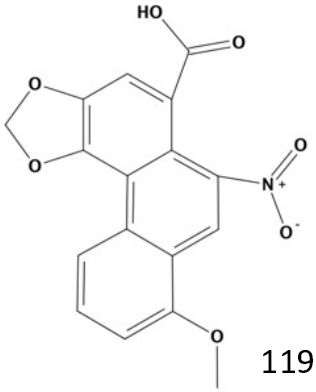


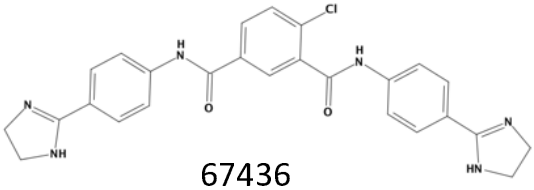
Figure S6: Hierarchical Clustering using MACCS fingerprinting. Distance Method Soergel (Tanimoto Coefficient); Clustering Method: Ward Linkage; Clustering Threshold: 0.7



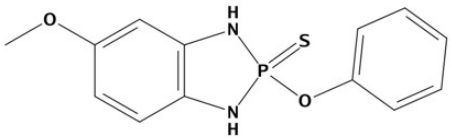
44584



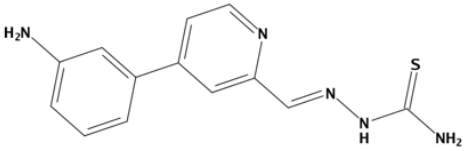
11926



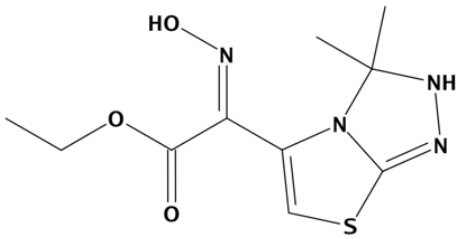
67436



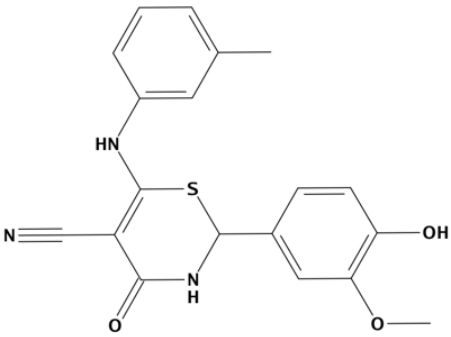
88402



195327



299968



319435

Figure S7: Two-dimensional structure of top hits.