

RESEARCH ARTICLE

Decoding molecular recognition of inhibitors targeting HDAC2 via molecular dynamics simulations and configurational entropy estimation

Suriya Tateing^{1,2,3}, Nuttee Suree^{1,2,4*}

1 Interdisciplinary Program in Biotechnology, Graduate School, Chiang Mai University, Chiang Mai, Thailand, **2** Division of Biochemistry and Biochemical Innovation, Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand, **3** Department of Plant and Soil Sciences, Faculty of Agriculture, Chiang Mai University, Chiang Mai, Thailand, **4** Center of Excellence in Materials Science and Technology, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand

* nuttee.suree@cmu.ac.th**OPEN ACCESS**

Citation: Tateing S, Suree N (2022) Decoding molecular recognition of inhibitors targeting HDAC2 via molecular dynamics simulations and configurational entropy estimation. PLoS ONE 17(8): e0273265. <https://doi.org/10.1371/journal.pone.0273265>

Editor: Hiroshi Fujisaki, Nippon Medical School, JAPAN

Received: May 1, 2022

Accepted: August 5, 2022

Published: August 18, 2022

Copyright: © 2022 Tateing, Suree. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting Information](#) files.

Funding: This work was supported by Center of Excellence in Materials Science and Technology, Faculty of Science, the Materials Science Research Center (MSRC), Chiang Mai University. It was also partially supported by the National Research Council of Thailand (Grant: 2556NRCT51390, 2558NRCT350269, and NRCT-PARB/39/2561). This research was also supported by the Program

Abstract

Molecular recognition by enzymes is a complicated process involving thermodynamic energies governing protein-ligand interactions. In order to aid the estimation of inhibitory activity of compounds targeting an enzyme, several computational methods can be employed to dissect this intermolecular contact. Herein, we report a structural dynamics investigation of an epigenetic enzyme HDAC2 in differentiating its binding to various inhibitors within the sub-sites of its active site. Molecular dynamics (MD) simulation was employed to elucidate the intermolecular interactions as well as the dynamics behavior of ligand binding. MD trajectories of five distinct HDAC2-inhibitor complexes reveal that compounds lacking adequate contacts with the opening rim of the active site possess high fluctuation along the cap portion, thus weakening the overall affinity. Key intermolecular interactions determining the effective binding of inhibitors include hydrogen bonds with Gly154, Asp181, and Tyr308; hydrophobic interactions between Phe155/Phe210 and the linker region; and a pi-stacking with Arg39 at the foot pocket. Decomposition of the binding free energy calculated per-residue by MM/PBSA also indicates that the interactions within the internal foot pocket, especially with residues Met35, Leu144, Gly305, and Gly306, can contribute significantly to the ligand binding. Additionally, configurational entropy of the binding was estimated and compared to the scale of the binding free energy in order to assess its contribution to the binding and to differentiate various ligand partners. It was found that the levels of entropic contribution are comparable among a set of structurally similar carbamide ligands, while it is greatly different for the set of unrelated ligands, ranging from 2.75 to 16.38 kcal/mol for the five inhibitors examined. These findings exemplify the importance of assessing molecular dynamics as well as estimating the entropic contribution in evaluating the ligand binding mechanism.

Management Unit for Human Resources & Institutional Development, Research and Innovation, Office of National Higher Education Science Research and Innovation Policy Council (NXPO) (Grant Number B16F640001).

Competing interests: The authors have declared that no competing interests exist.

Introduction

Several physiological functions and gene expressions are epigenetically controlled through the modification of chromatin structures and the alteration of histone acetylation and deacetylation levels in the nucleosomes [1–3]. Antagonistic enzymes involved in these processes include histone acetyltransferases (HATs) and histone deacetylases (HDACs). For histone acetylation, HATs acetylate the ϵ -NH₂ group of histone lysines, thereby neutralizing the charge on the protein surface, relaxing the chromatin structure, and thus allowing the access of transcription factors [4–7]. HDACs, on the other hand, deacetylate the histone lysine residues and reconstitute the positive charge, thus compacting the chromatin structure and preventing the gene expression from occurring [5, 6]. This important role in gene expression makes HDACs critical targets for drug development, as their dysfunction and inhibition have been associated with several diseases, including cancer [7–10], neurological, renal, lung, and cardiovascular diseases [11–16], as well as inflammatory immune disease [17–19].

HDAC enzymes can be subcategorized into four families based upon their structure and function, denoted as classes I, II, III, and IV HDACs [20]. HDAC2, a member of the class I HDAC family, is one of the most studied isoforms and it possesses a high enantioselectivity towards its substrates [21]. Aside from its structural homology and catalytic activity similar to other isoforms within the class I HDACs, HDAC2 also shares a highly conserved catalytic domain that comprises a lipophilic tube, a deep catalytic site where its Zn²⁺ cofactor is located, and an internal ‘foot pocket’ [7]. The lipophilic tube connecting the opening rim to the catalytic site is formed by Gly154, Phe155, His183, Phe210 and Leu276 [22, 23]. The catalytic site is surrounded by Asp181, His183, and Asp269, while the internal ‘foot pocket’ is formed by Tyr29, Met35, Arg39, Phe114, and Leu144 residues [22–24]. This elongated shape of the substrate binding site also provides high specificity to HDAC2 inhibitors, which usually contain a cap group, a hydrophobic linker portion, and a zinc-binding group (ZBG) [25–27]. In recent years, various novel HDAC inhibitors with *in vitro* and *in vivo* potencies have been reported for their potential therapeutic effects against chromatin-related diseases [28, 29].

In order to help accelerate the process of inhibitor development, computational methods are generally employed in order to dissect the molecular recognition by which the target protein confers activity, thus providing a powerful tool to virtually search for candidate compounds that potentially bind tightly to the binding site. Simulation of molecular motions and structural information of the binding at atomic details have also helped elucidate the thermodynamic energy terms involved in the intermolecular recognition [30].

Molecular dynamics (MD) simulation and molecular mechanics (MM) calculations have been used in the designing of HDAC2 inhibitors, especially focusing on the non-bonding energy contributed by the binding with the composite active site (CAS) [31]. Dewaker, et al. demonstrated the use of energy decomposition of residues contacting the ligand cap group, the ZBG, and the foot-pocket residues in the design of new inhibitor compounds potentially bind more effectively to the active site. However, a comparison between the calculated binding energy and the actual experimental data, e.g. IC₅₀ values of the ligands, is still necessary in order to derive a more reliable binding mechanism. Additionally, although MD simulation and MM calculations could help estimate the enthalpic contribution of the binding, the role of the entropic contribution and its approximation have remained elusive in many molecular systems. One of the prime interests here is the configurational entropy which is caused by the variation of conformational states of the protein. Currently, several methods have been devised in order to calculate entropy. One example is an empirical method that subcategorizes entropic contribution into two parts: solvation free entropy and configurational free entropy [32, 33]. The quasi-harmonic approximation (QH) is another method introduced by Karplus and

Kushick in 1981 [34] that estimates the configurational entropy based upon MD simulation by building a Gaussian distribution of protein conformations [35].

For most applications of MD simulation in biomolecular systems, entropic contributions to the binding free energy are generally neglected due to their high computational cost or complicated estimation processes. Nonetheless, insight into this particular energetic contribution could provide deeper understanding of the molecular system of interest since entropy can have different levels of impact on molecular recognition. For HDACs, the Christianson group has reported the role of entropy in driving selectivity for inhibitors targeting a class-IIb HDAC6 vs. a class-I HDAC8 enzyme [36]. Isothermal titration calorimetry experiments revealed that the binding entropy of HDAC6-specific inhibitors was more favorable for HDAC6 than for HDAC8 [36]. This difference in entropic contribution was believed to stem from differences in configurational entropy and/or desolvation entropy, particularly through the interactions with the aromatic cleft of the binding site [36].

In this study, we aimed to investigate a range of binding mechanisms and molecular dynamics properties involved in the ligand binding of HDAC2 in order to appraise the influential factors contributing to the inhibitory activity of various ligands. MD simulations of complexes formed between HDAC2 and several known inhibitors were performed and analyzed for the modes of binding in the sub-sites of its active site cleft. Molecular mechanics Poisson-Boltzmann Surface Area (MM/PBSA) calculations were also employed to estimate the energetic contributions to the inhibitor binding as well as to investigate the level of influential interaction of each important amino acid residue. Additionally, we quantified the entropic contribution to the overall binding energy, and made comparisons among the structurally diverse and similar ligands. The data from this study provide new insight into the significance of dynamics behavior and an entropic contribution in delineating inhibitor binding mechanisms. This approach could become a powerful tool in selecting and optimizing novel ligand candidates for HDACs.

Material and methods

Protein-ligand complex preparation

Crystal structures of human HDAC2 in complex with various inhibitors were retrieved from the Protein Data Bank [37, 38], which include complexes with either N-(2-aminophenyl)benzamide [22] (LLX, PDB ID: 3MAX, 2.05 Å resolution), octanedioic acid hydroxyamide phenylamide [39] (SAHA or vorinostat, PDB ID: 4LXZ, 1.85 Å resolution), 4-(acetylamino)-N-[2-amino-5-(thiophen-2-yl)phenyl]benzamide [39] (20Y, PDB ID: 4LY1, 1.57 Å resolution), N-(4-amino-4'-fluoro[1,1'-biphenyl]-3-yl)oxane-4-carboxamide [40] (IWX or BRD4884, PDB ID: 5IWG, 1.66 Å resolution), or (3-exo)-N-(4-amino-4'-fluoro[1,1'-biphenyl]-3-yl)-8-oxabicyclo[3.2.1]octane-3-carboxamide [40] (6EZ or BRD7232, PDB ID: 5IX0, 1.72 Å resolution). The five crystal structure complexes were prepared as input coordinate files for the molecular dynamics simulations using a combination software involving antechamber and LEaP programs. Atomic partial charges of inhibitors were obtained by fitting the electrostatic potentials (ESP), derived using Gaussian09 (G09), with HF/6-31G* as the basis set [41]. The protonation configurations of the titratable histidine residues were predicted by the APBS-PDB2PQR server [42]. The webserver was employed to predict the proper protonation states for each titratable residue, which results in relevant output files pqr, stdout.txt, and stderr.txt. Especially, the protonation states of HIS to be HID or HIE of protein were displayed in the pqr output file, while the calculated pKa values were shown in the log output file. Notably, the mass and charge of the metal ion (Zn^{2+}) were assigned via a parameter file `frmod.ions234lm_126_tip3p` (for divalent ions for the TIP3P water model), which yielded the mass and charge of Zn^{2+}

⁺ of 65.4 and 2.0, respectively. The results are summarized in S1 Table of [S1 File](#). For the following molecular mechanics (MM) minimizations and MD simulations, AMBER14SB force field [43, 44] and the general AMBER force field [45] (gaff) were employed to establish the potential of proteins and inhibitors, respectively. Water molecules were randomly replaced by chloride ions in order to neutralize the charge of the system. Finally, the entire system was solvated in a cubic periodic box of the TIP3P explicit solvent model [46], keeping the distances between the edges of the water box and the closest atom of the solutes at least 10 Å. To avoid any edge effects, periodic boundary conditions were applied during the entire MD simulations. All HDAC2-inhibitor complex systems were converted via a ParmED script to be compatible with the Gromacs format.

Molecular dynamics simulations

For each molecular system, energy minimization and MD simulations were performed using Gromacs 6.4.3 package [47]. For the HDAC2-inhibitor complexes, the protein-ligand modeling was performed using the LEaP from AmberTools20 [48]. Subsequently, for the force fields employed, leaprc.protein.ff14SB [38], leaprc.water.tip3p, and leaprc.gaff were sourced for protein, water, and ligand, respectively, which would load a frcmod.ions234lm_126_tip3p file that has a parameter set for Zn²⁺ ion for the TIP3P water model. This command also loaded an AMBER-format parameter set file and placed it in the variable 'frcmod.ionsjc_tip3p' that consists of monovalent ion parameters and a TIP3P water model. The force field parameter for Zn²⁺ ion adopted in our simulations were designed by Li and Merz [49] and employed according to the Amber 2020 protocol. The AMBER topology files were converted to Gromacs topology files via ParmED script. Each system was subjected to an energy minimization without restriction, which entailed a 5000-step steepest descent minimization. For the equilibration process, a position restraint was applied to the protein and the bound inhibitor both for NVT and NPT ensembles. Particle Mesh Ewald [49] algorithm [50, 51] was employed to handle the long-range electrostatic interactions. The cutoff distance for the van der Waals energy interactions was set at 14 Å and the LINCS algorithm for bond constraints was used [52]. The systems were gradually heated in the NVT ensemble from 0 to 300 K for over 100 ps. Subsequently, an isothermal isobaric ensemble (NPT) with periodic boundary conditions was performed for 100 ps. Finally, 100-ns MD simulations were carried out to extend the MD production for each system in an isothermal isobaric ensemble (NPT) [53] using periodic boundary conditions. The SHAKE method [54] was applied to constrain all covalent bonds involving hydrogen atoms. Each simulation was coupled to a 300 K thermal bath at 1.0 bar by applying Berendsen the thermostat and the Parrinello-Rahman barostat, respectively. The temperature and pressure coupling parameters were separately treated for more accuracy. During the sampling process, the coordinates were saved every 0.1 ps and the conformations generated from the simulations were used for further MD trajectory analyses. MD trajectories and structures of each system were visualized using Visual Molecular Dynamics [55] software [55] and Accelrys Discovery Studio Visualizer 4.0 (Accelrys Software Inc.). SigmaPlot 12.5 (Systat Software, San Jose, CA) was used for generating all plots of the various parameters.

Coordination number of the Zn²⁺ cofactor ion was calculated according to the concept of quantifying atomic interactions surrounding the central metal atom [56, 57] using the following equation:

$$n = \int_{R_1}^{R_2} 4\pi r^2 \rho g(r) dr \quad (1)$$

, where n is the number of specific atoms in a peak of radius distribution function (RDF), r is

the distance from the atom to the target metal atom, R_1 is the starting position (at 0 nm), R_2 is the position of the first shell (with a cutoff distance of 0.27 nm), ρ is the number density of the atom in the main phase, and $g(r)$ is the RDFs from the atom to the target atom.

Additionally, coordination of the Zn^{2+} metal cofactor was visualized from the MD simulations using the Accelrys Discovery Studio Visualizer. Briefly, from the MD trajectories of all complexes, eleven snapshots representing every 10 ns were extracted from the runs and superimposed to visualize the coordination number of the cofactor ion. Atomic distances between the Zn^{2+} cofactor of HDAC2 and a nearby heavy atom of amino acid residues or ligands were also measured.

Principal component analysis

Trajectories of the MD simulations were utilized to identify the concerted motions of HDAC2-ligand complexes. For the estimation of configurational entropy change in each protein-ligand complex, calculations were based upon the MD trajectories of the apo-protein, individual ligands, as well as the protein-ligand complexes. The last snapshots (80–100 ns) of MD trajectories of apo-proteins, ligands, or protein-ligand complexes were acquired. First, for constructing the covariance matrix, translational and rotational movements of the HDAC2 protein were removed. To do this, a command `g_trjconv`, a utility within the Gromacs suite, was employed. This command contains a command `-fit rot+trans` as an additional function to systematically remove rotation and translation motions by aligning the $C\alpha$ atoms of structures within a single MD run to those $C\alpha$ atoms of the initial structure. Of note, for the entropy calculation, the initial structure was the structure at 80 ns (as the range of calculation was 80–100 ns). However, for the density plot visualization, the initial structure was the structure at 10 ns (as the range for this observation was 10–100 ns).

The covariance matrix was calculated using Cartesian coordinates of the $C\alpha$ atoms of the protein. This method was also applied for the inhibitor ligands using all of their atoms. Principal component analysis (PCA) was employed for estimating the direction of molecular motions based upon the covariance matrix of the atomic fluctuations. Diagonalization of this matrix yields a set of eigenvectors and eigenvalues that describe collective modes of fluctuations of the molecules. Eigenvectors represent the direction of the motions, whereas the corresponding eigenvalues represent the amplitudes of those directions along the multidimensional space [58]. Displacement of atoms along each eigenvector represents concerted motions of the protein along each direction. Gromacs utilities, including `'g_covar'` and `'g_anaeig'`, were used to solve some functionally relevant motions. Particularly, the `'g_covar'` tool was used to perform the diagonalization of the computation elements from the covariance matrix of positional fluctuations found within the $C\alpha$ carbon atoms of the protein molecules. The `'g_anaeig'` tool was used to analyze the free energy landscape for the projections along the two eigenvectors, and the entropy term was then estimated by the quasi-harmonic approximation (QH) method. Differences between the entropies calculated in the bound and unbound forms of the protein and ligands were used to calculate the configurational entropy change.

Molecular mechanics calculations

For HDAC2-inhibitors complexes, 1,000 snapshots taken from the last 10 ns trajectories were obtained for calculating the binding free energy using the molecular mechanics Poisson-Boltzmann Surface Area method (`g_mmpbsa`) [59], as well as the APBS software suite [42], which is implemented in the Gromacs package. The binding free energy was calculated using Eq (2).

$$\Delta G_{\text{binding}} = \Delta G_{\text{complex}} - (\Delta G_{\text{protein}} + \Delta G_{\text{ligand}}) \quad (2)$$

$$\Delta G_{\text{binding}} = \Delta E_{\text{MM}} + \Delta G_{\text{solvation}} - T\Delta S \quad (3)$$

$$\Delta E_{\text{MM}} = \Delta E_{\text{bonded}} + \Delta E_{\text{vdW}} + \Delta E_{\text{elec}} \quad (4)$$

$$\Delta G_{\text{solvation}} = \Delta G_{\text{polar}} + \Delta G_{\text{nonpolar}} \quad (5)$$

$$\Delta G_{\text{nonpolar}} = \gamma \times \text{SASA} + b \quad (6)$$

, where $\Delta G_{\text{binding}}$ is an estimate of the binding free energy between the HDAC protein and the ligand. Each free energy term in Eq (2) is a sum of the gas phase molecular mechanical energy (ΔE_{MM}), the solvation free energy ($\Delta G_{\text{solvation}}$) and the entropic term ($-T\Delta S$). In Eq (4), ΔE_{MM} can be divided into contributions from the bonded term (ΔE_{bonded}), which includes bond, angle, and dihedral terms, and two non-bonded terms, which are the van der Waals energy (ΔE_{vdW}) and the electrostatic energy (ΔE_{elec}) in the gas phase. The solvation free energy ($\Delta G_{\text{solvation}}$) was estimated according to Eq (5) using continuum solvent methods and it can be divided into two parts: the polar contribution (ΔG_{polar}) and the nonpolar contribution ($\Delta G_{\text{nonpolar}}$). The electrostatic solvation energy was determined using the finite difference PB model. The dielectric constants used for the interior [59] and the exterior (water) were 1 and 80, respectively. Atomic radii and charges were the same as those used in the MD simulations. The nonpolar contribution ($\Delta G_{\text{nonpolar}}$) to the solvation free energy was calculated from the solvent-accessible surface area [59], which is expressed in Eq (6). The standard numerical values of γ is 0.00542 kcal/(mol·Å²), and β is 0.92 kcal/mol for the Poisson-Boltzmann method. The probe radius of the solvent was set to 1.4 Å.

Results and discussions

Molecular dynamics simulations of the HDAC2-inhibitors complexes

In order to gain a better perspective on the ligand binding of HDAC2 as well as the corresponding changes in their dynamics behavior, MD simulations were performed on the crystal structures of complexes involving HDAC2 and five known inhibitors. These selected ligands, which include LLX, SAHA, 20Y, IWX, or 6EZ, possess distinct structural scaffolds, and their reported inhibitory activities (IC_{50}) also range from 1.7 nM to 1,200 nM. Initially, coordinates from the crystal structures were used as input for our 100-ns MD simulations, after which the resulting dynamics parameters, as well as the three-dimensional spatial snapshots, were acquired and analyzed. In this particular study we focused on the structural motions observed within the active site cleft of HDAC2 and its neighboring areas, that could shed light onto the mechanisms by which the selected ligands convey various levels of inhibitory activity. Significant changes in dynamics and energetic contributions of these residues when interacting with the ligands could also imply their important roles in conferring selectivity for suitable ligands.

Fig 1 depicts an overview of how the five ligands were nested within the active site pocket. HDAC2 contains a lipophilic tube that links the opening surface to the catalytic triad (His145, His146, and Tyr308), and the adjacent Zn²⁺ cofactor binding region (Asp181, His183, and Asp269). Deep inside the cleft also lies an internal cavity or a ‘foot pocket’, which is formed by Tyr29, Met35, Arg39, Phe114, and Leu144, which could convey additional specificity to the ligand binding. As observed from the overlaid crystal structures, ligands that contain thin and elongated shapes could thus be threaded into the pocket and can extend their binding interactions, surpassing the cofactor binding area to this deeper site. The foot pocket of the active site

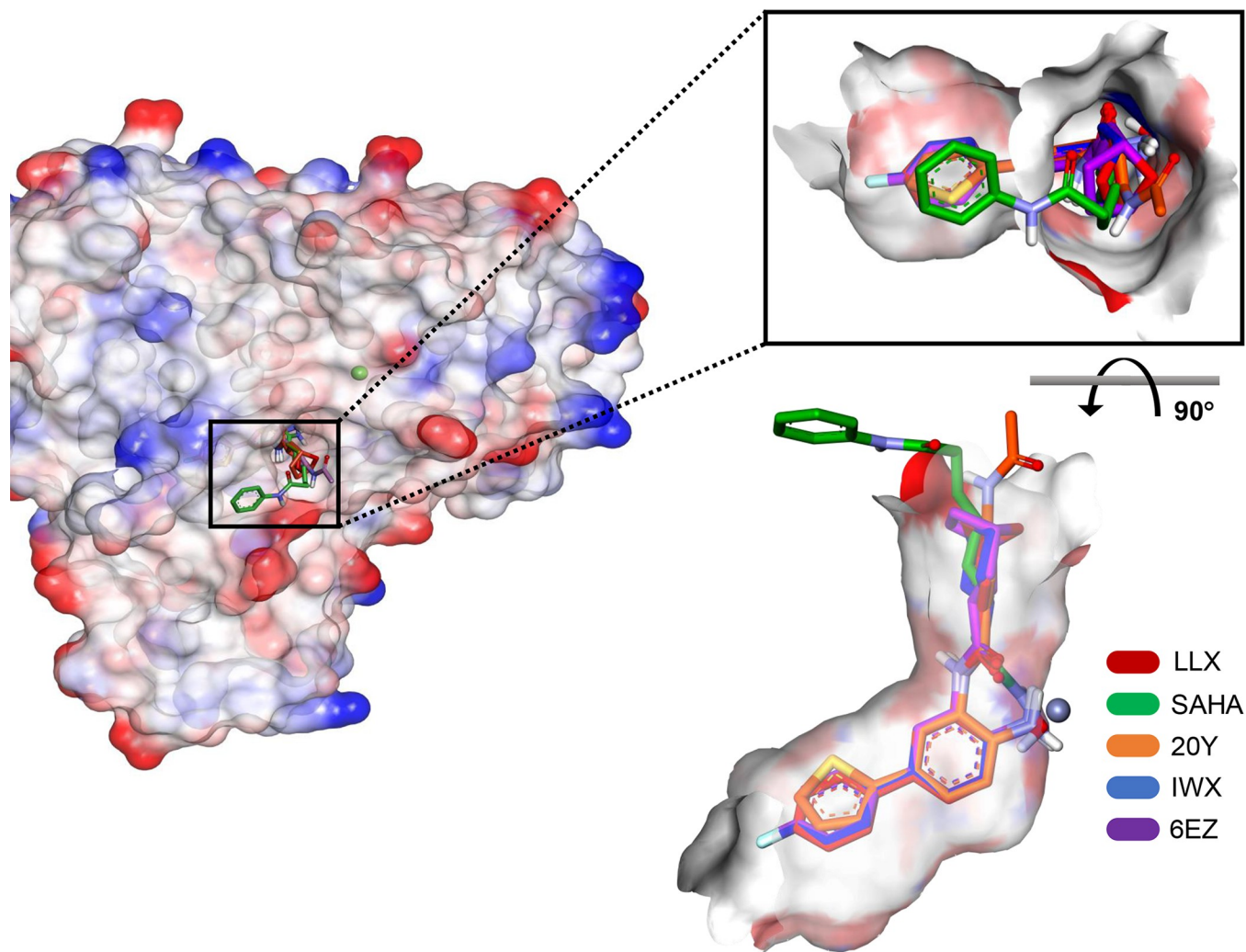


Fig 1. Structural overview of HDAC2 and its active site location. (Left) Electrostatic surface representation of HDAC2 showing the binding location of its inhibitors. (Right) Overlaid structures of inhibitor ligands when bound to the active site. The inhibitor molecules are illustrated in stick representations showing stabilized conformations of LLX (red), SAHA [60], 20Y (orange), IWX (blue), and 6EZ (purple) ligands, which are structurally overlaid using heavy atom coordinates. Zinc ion is shown as a gray ball.

<https://doi.org/10.1371/journal.pone.0273265.g001>

cleft generally contains multiple water molecules, which are commonly found within crystal structures of complexes with smaller-sized ligands [22]. Additionally, the opening rim of the active site cleft could also provide additional interactions to augment the ligand affinity and specificity. SAHA is a good example of ligands that take advantage of this binding mechanism since its aniline group partially acts as a ‘knob’ plugging the opening region. From these different regions of binding, it is plausible that any ligand that employs all areas of intermolecular interactions would create a stable complex, exhibiting a more stable dynamics pattern, and potentially contributing to stronger inhibitory activity.

To delineate the binding mechanism in a more comprehensive manner, MD simulations were employed to analyze changes in motions observed within the HDAC protein upon ligand binding. Fig 2 shows the dynamics parameters collected from our 100-ns MD simulations of the protein-ligand complexes. Simulation equilibria were evaluated by comparing the root-mean-square deviation (RMSD) of the backbone atoms relative to the initial coordinates

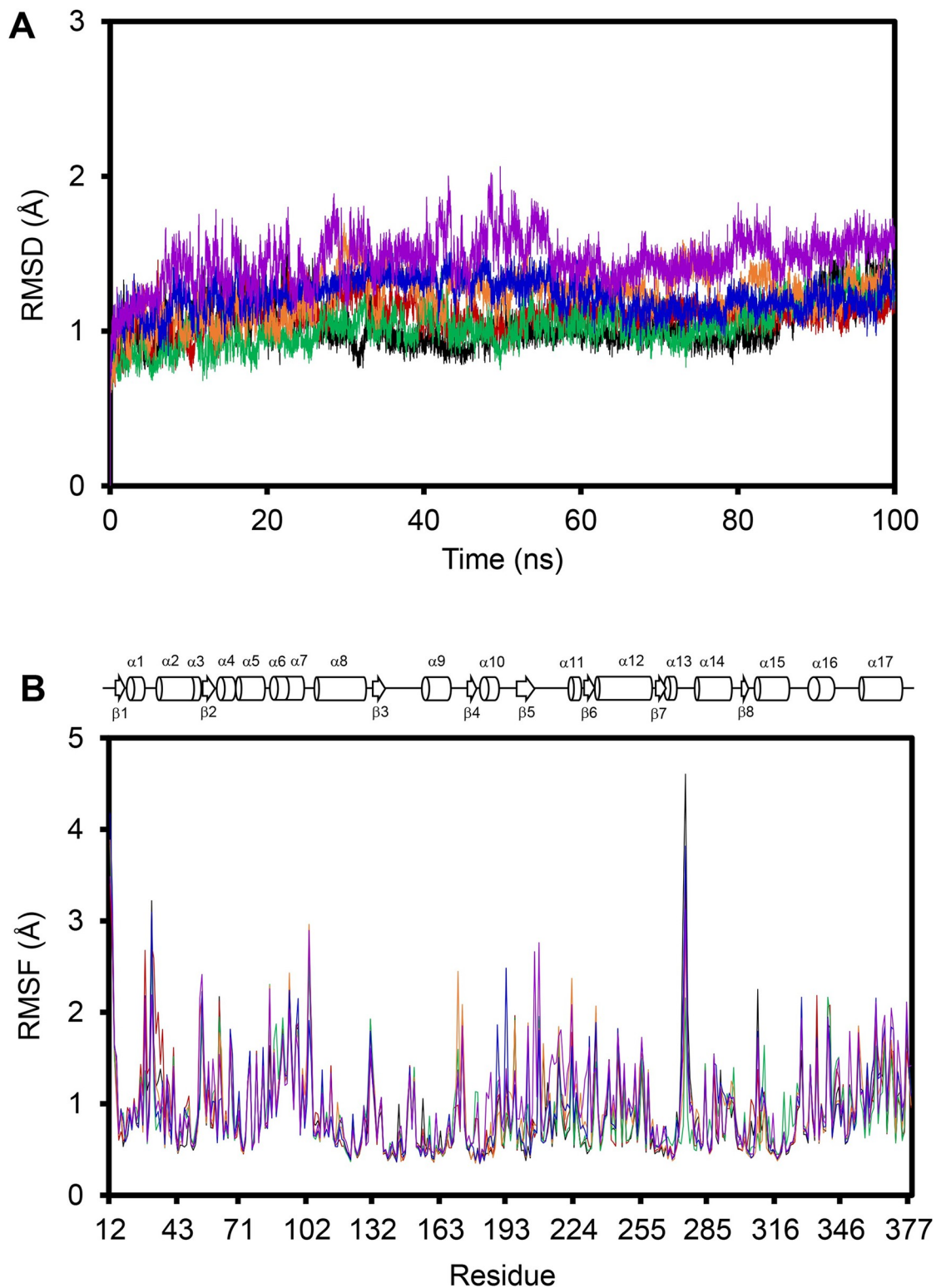


Fig 2. (A) Root-mean-square deviations (RMSD) and (B) root-mean-square fluctuations (RMSF) of backbone atoms on HDAC2 when bound or unbound by inhibitors. Unbound HDAC2 (black) is compared with its complex forms with LLX (red), SAHA [60], 20Y (orange), IWX (blue), and 6EZ (purple). Topology diagram of HDAC2 is also illustrated atop of the RMSF chart to indicate range of secondary structures along the residue sequence.

<https://doi.org/10.1371/journal.pone.0273265.g002>

(Fig 2A). Additionally, the MD simulations of the unbound HDAC2 and all complexes were repeated to observe the ranges of fluctuation comparing between the runs and throughout the 100-ns simulation period in order to justify if the simulation time used was sufficient. It was found that the RMSD levels fluctuated within a range of 1 Å (S1 Fig in S1 File). An extended MD run of HDAC2-20Y complex also showed similar range of fluctuation (S2 Fig in S1 File). Overlay of snapshots of the complex structures obtained at 100 ns and at 200 ns of the MD simulation shows similar conformations, with an RMSD between the heavy atoms of the two structures of 0.8187 Å (S3 Fig in S1 File). Thus, we selected 100 ns as the simulation time for this study. It is apparent that the RMSDs of all the ligand complexes fluctuate in slightly different ranges. The highest fluctuations were observed for the complexes of HDAC2-6EZ (purple) and HDAC2-20Y (orange), especially during the first 60 ns of simulation. On the other hand, the HDAC2-IWX complex (blue) fluctuates the least during the entire simulation. Nevertheless, all molecular systems reached a relatively steady period after 80 ns, indicating that the 100-ns simulation duration is sufficient for a comparative study. However, the information from the RMSD data might not be adequate to describe localized motions of the protein structures. Therefore, root-mean-square fluctuations (RMSF) data were also collected and compared among all complexes as well as with the free enzyme.

Residual structural fluctuations of the bound ligands

Fig 2B shows spatial fluctuations of the backbone heavy atoms of all residues along the protein sequence, comparing the free protein with the protein-ligand complexes. As anticipated, the regions with most fluctuation are located among the loops connecting secondary structure motifs within the protein, particularly at the $\alpha 13/\alpha 14$ loop that exhibits the highest fluctuation. However, when focusing only on the active-site residues (His145, His146, Asp181, His183, Asp269, and Tyr308) or the internal cavity residues (Tyr29, Met35, Arg39, Phe114, Leu144, Gly305, and Gly306), fluctuations among these regions do not significantly differ, and no obvious trend could thus be derived from the different ligand complexes.

Since the observed spatial changes within the protein structure did not yield any complete differentiation among the ligands, we then asked whether there could be any differences observable within the ligand molecules, possibly over various stages of time. To this end, snapshots of all ligand structures taken from the different time points of simulation, including 0, 20, 40, 60, 80, and 100 ns, were overlaid and compared for their molecular motions. As illustrated in Fig 3, all five ligands have different levels of movement relative to the active-site location. When mapped onto the estimated position of ligand binding, it is apparent that LLX, 20Y, IWX, and 6EZ insert themselves deep inside the pocket and reach the internal cavity. SAHA, on the other hand, extends to only the cofactor binding region but uses its aniline moiety as a cap group in order to interact with the opening surface. This pattern of cap binding at the opening region can also be seen, though to a lesser extent, in the binding of 20Y, 6EZ, and LLX, but not in the case of IWX. However, it is apparent that the cap portion of the ligands may not have sufficient intermolecular interaction to restrict the overhang structures as they fluctuate highly in the region protruding from the active-site canal.

Intermolecular interactions restraining the ligand binding

Key binding interactions between the active site residues and the ligands are also summarized in Fig 4. Although many hydrophobic interactions collectively contribute to binding affinity, a few hydrogen bonds are prominently found within the active site, particularly near the catalytic triad and the cofactor binding center. These include the backbone carbonyl of Gly154 and the side-chain carboxyl of Asp181, which is also a part of the Zn^{2+} -binding residues (Asp181,

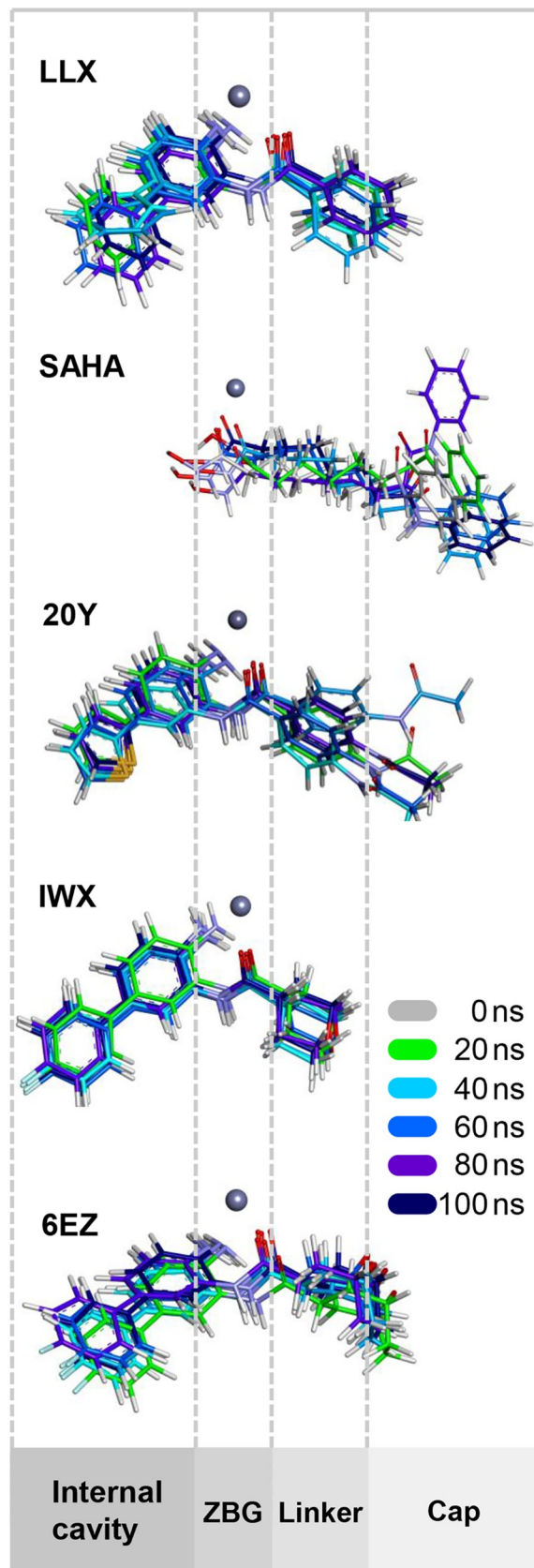


Fig 3. Structural overlays of ligand conformations observed at different time points during the 100-ns MD runs. Relative positioning of the bound ligand structures along the active site canal is also indicated with dashed lines. Ligand positions are aligned using the coordinates of the Zn^{2+} cofactors, shown as gray spheres, as a reference point. ZBG denotes the ligand's zinc-binding group.

<https://doi.org/10.1371/journal.pone.0273265.g003>

His183, and Asp269), and the side-chain hydroxyl of Tyr308. Additionally, histidine residues located near the catalytic triad, His145 and His146, can form hydrogen bonds with the ligands, particularly at the hydroxyl or the amine moieties of the ligand structures. These interactions near the Zn^{2+} cofactor could further anchor the inhibitors to remain within the active site (S4 Fig in S1 File). For the ligands with high inhibitory potencies (LLX, 20Y, and IWX, exhibiting IC_{50} values of 27, 56.3, and 62 nM, respectively) (S3 and S5 Tables in S1 File), two or three of these hydrogen bonds are observable within the complexes. On the other hand, ligands with lower levels of potency (SAHA and 6EZ, exhibiting IC_{50} values of 251 and 168 nM, respectively) form only one hydrogen bond. The correlation between the number of hydrogen bonds formed and the inhibitor potency level coincides with several studies [61–64] suggesting the role of hydrogen bonds as one of the key indicators for a successful inhibitor. Of note, the reported IC_{50} values for HDAC2 inhibitors can be varied based upon different studies. Therefore, an extensive literature review of IC_{50} values was conducted and summarized in S5 Table of S1 File. Values selected for use in our calculation are the results from the most similar methods of determination. Additionally, we did not convert the IC_{50} values into K_i due to the lack of the information of substrate concentration and K_m , which were not explicitly stated in most studies reported.

Notably, a moderate level of hydrophobic interactions at the opening surface of the active site cleft are also identified, namely with aromatic rings of Phe155 and Phe210. These two residues interact with the linker portion of the ligand structures and can provide particular pi-pi stabilization for the ligands that contain aromatic moieties in this region such as LLX and 20Y, both of which exhibit higher levels of potency. At the deeper side at the foot pocket, Arg39 provides an additional cation-pi intermolecular interaction further anchoring the ligands reaching this site. All the ligands, except SAHA, contain an aromatic moiety at this particular location. The role of Arg39 contributing to the ligand affinity observed here agrees well with previous studies [29, 65, 66]. For the Zn^{2+} metal chelation, key residues Asp181, His183, and Asp269 predominantly act as chelators for the Zn^{2+} cofactor. In the absence of ligand, a water molecule fulfills the coordination of Zn^{2+} . For ligand binding, however, this water molecule is replaced by a carbonyl oxygen atom from the ligand inhibitor participating in the chelation (S4 Fig in S1 File) and the coordination number of Zn^{2+} remained the same. Atomic distance ranges of metal chelation are summarized in S5 Fig and S2 Table of S1 File. For the calculation of coordination number of the Zn^{2+} cofactor ion, the number of specific atoms in a peak of radius distribution function (RDF) was computed according to Eq (1) via a command `g_rdf`, a utility function implemented in the GROMACS package. The initial calculation resulted in distributions of atomic partners around the ion, as illustrated in S6 Fig of S1 File. It was observed that the distance trends of the first peaks, which indicate the first hydration shell, were similar among the complexes and centered around 0.2 nm. To distinguish between the first and the second hydration shells, a cutoff distance was placed at 0.27 nm. Since the coordination of the ion is mainly contributed by the first shell, coordination numbers can be calculated from the integral function. The integration yielded an overall running coordination number for the Zn^{2+} ion as a result of interactions with its atomic partners lining the catalytic site of HDAC2 (S7 Fig in S1 File). From the entire 100-ns MD simulations, the calculated coordination numbers of the Zn^{2+} were 6.00, 5.86, 5.29, 5.13, and 5.01 for the complexes between HDAC2 and LLX,

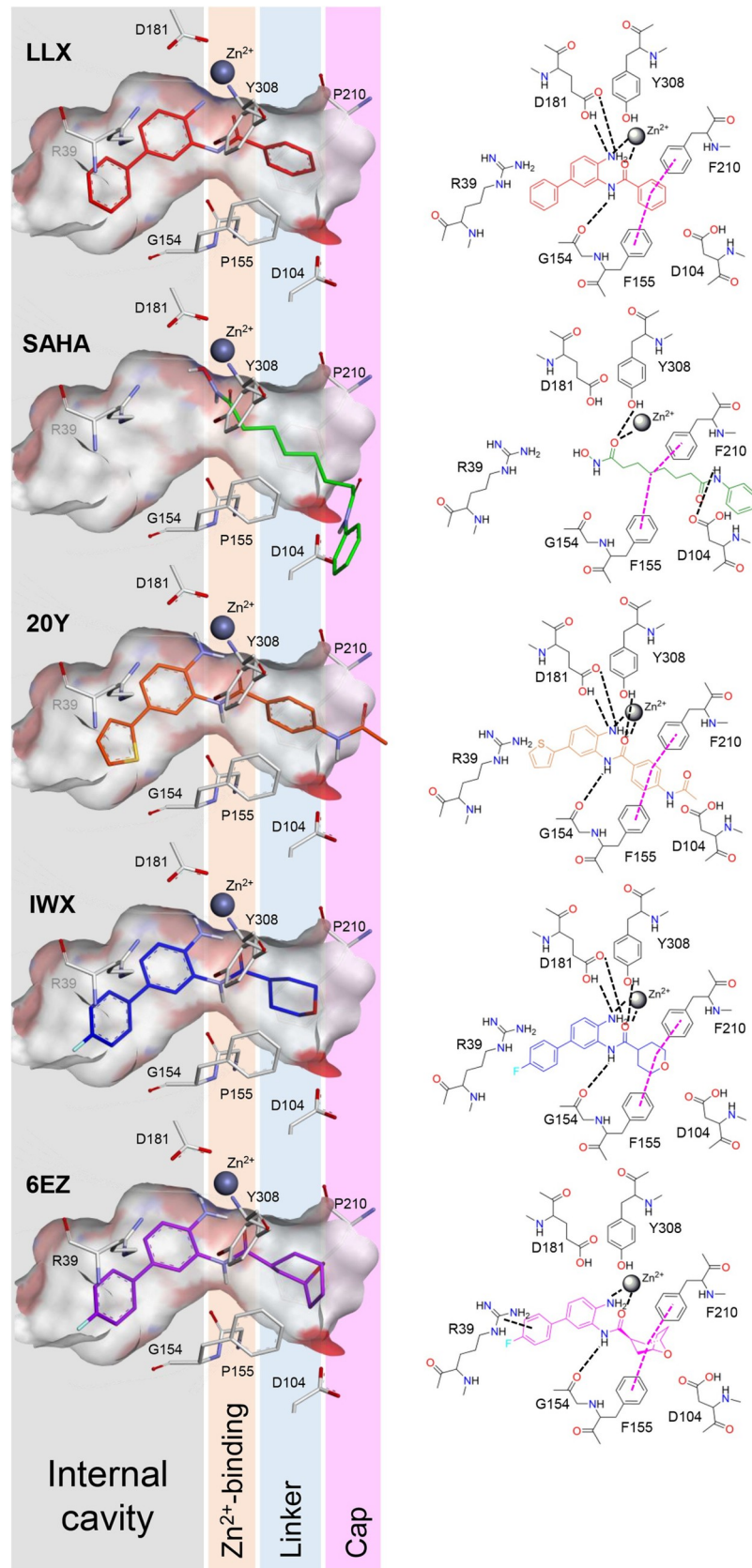


Fig 4. Binding interactions between inhibitor ligands and the interacting amino acid residues lining the active site of HDAC2. (Left) three-dimensional representations of the ligands when nested inside the active site canal. Inhibitor ligands LLX (red), SAHA [60], 20Y (orange), IWX (blue), and 6EZ (purple) are shown in stick representations. Zn^{2+} cofactors are shown as gray spheres. Positions of the ligands along the active site are compared for their structural regions of internal-cavity binding, Zn^{2+} binding, linkers, and cap groups. (Right) two-dimensional diagrams summarizing the protein-ligand binding interactions. Hydrogen bonding networks and cofactor coordination are highlighted in black dashed lines, whereas hydrophobic interactions are indicated with purple dashed lines.

<https://doi.org/10.1371/journal.pone.0273265.g004>

SAHA, 20Y, IWX, and 6EZ, respectively. Of note, these coordination numbers were determined using the same set of the R_1 and R_2 values for Eq (1).

Notably, the atomic partners coordinating the Zn^{2+} ion are different between the event of substrate catalysis of HDAC2 and the event of inhibitor binding. HDACs employ a promoted-water mechanism for catalysis, in which the active site metal ion (Zn^{2+}) and a histidine general base activate a metal-bound water molecule for the nucleophilic attack of the native substrate [67]. For the inhibitor binding, however, HDAC inhibitors often exploit the chelation interactions with the catalytic Zn^{2+} ion, thus replacing this water molecule from the active site. For example, hydroxamate and benzamide derivatives ionize to form exceedingly stable 5- and 6-membered rings that efficiently chelate the Zn^{2+} [68].

Taken together, it can be summarized that the key intermolecular interactions may include: 1) hydrogen bonds with His145, His146, Gly154, Asp181, and Tyr308; 2) hydrophobic interactions between Phe155/Phe210 and the linker region; 3) a pi-stacking with the side chain of Arg39 at the foot pocket and 4) metal chelation contributions from the ligand and the Zn^{2+} -binding residues (Asp181, His183, and Asp269). Different combinations of these key interactions could provide a spectrum of ligand affinities and thus inhibitory activities. They could also be perceived as determining factors to predict the inhibitory potency of a ligand. For example, 6EZ that interacts with both the foot pocket and the linker regions, but forms only one hydrogen bond with the active site, exhibits only a moderate level of potency ($IC_{50} = 168$ nM) when compared to other ligands that have robust interactions with the foot pocket. On the other hand, LLX that forms three hydrogen bonds with D181 and G154 and interacts with both the foot pocket and the linker regions, exhibits a higher level of inhibitory potency ($IC_{50} = 27$ nM).

Binding energetic contribution by the active-site residues

In order to pinpoint energetic contributions from each amino acid residue to the ligand binding in a more quantitative manner, we employed molecular mechanics Poisson-Boltzmann Surface Area (MM/PBSA) calculations to estimate the binding free energy per-residue. Fig 5A shows a decomposition plot of binding free energy contributed from individual binding residues of HDAC2 when bound by the five inhibitor ligands. It is apparent that most of the binding residues show favorable contribution energies (negative values) to the ligand binding as anticipated. However, a few charge-bearing residues within the active site give rise to unfavorable (positive values) contribution energies. These include the Zn^{2+} -binding residues Asp181, His183, and Asp269, as well as the neighboring residues Asp160 and Arg39. Due to the fact that, these residues normally form a charge relay network with the cofactor in the absence of any ligand, it is possible that binding to a ligand that disrupts or interferes with this charge network could result in unfavorable energetic contributions. Interestingly, when looking at the trends of contribution energies from the two main Zn^{2+} -binding residues Asp181 and Asp269 (Fig 5B), the least unfavorable values are found for the binding of LLX and IWX ligands, both of which are highly potent inhibitors with IC_{50} values of 27 nM and 62 nM, respectively. Of note, decomposition of non-bonding energies, namely electrostatic energy and van der Waals

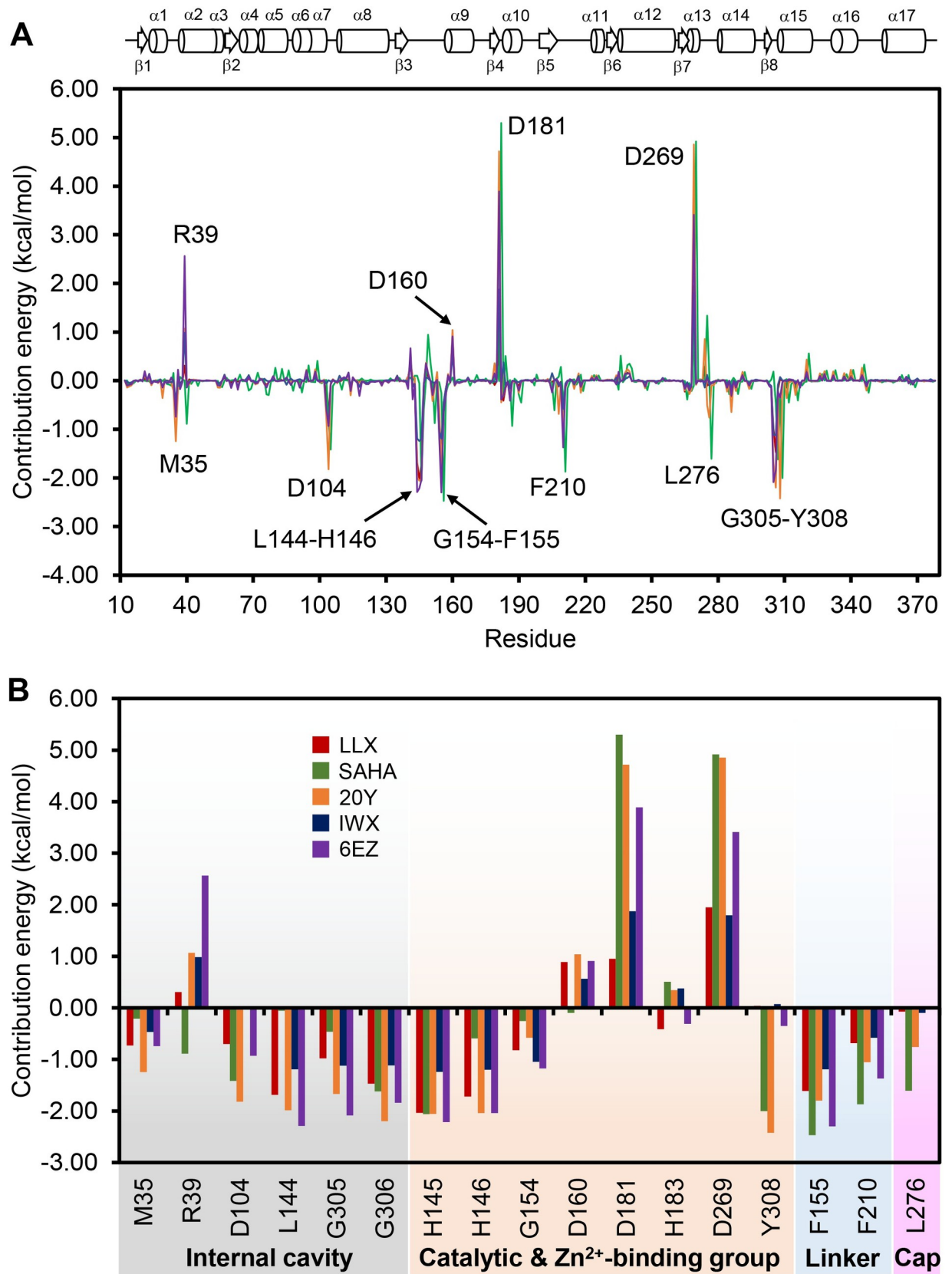


Fig 5. Decomposition of binding free energy per-residue of HDAC2 when bound by various inhibitors. (A) Contribution energies of all residues on HDAC2 are indicated when bound by either LLX (red), SAHA [60], 20Y (orange), IWX (blue), or 6EZ (purple) ligands. Topology diagram of HDAC2 is also depicted atop of the chart indicating the range of secondary structures along the residue sequence. Consensus of key residues that have significant contribution to the binding energy are identified by residue number and are specifically illustrated in more detail in a bar graph (B). Regions of the ligands those the key residues interact with are also indicated.

<https://doi.org/10.1371/journal.pone.0273265.g005>

interaction, have been shown to correlate with the total binding energy of HDAC2 ligand binding [31], although without a consideration of entropic term. Our observation of the correlation between the energy decomposition and the actual potency thus confirms the role of per-residue energy as determinant factors for predicting the actual inhibitory activity as previously described [69–71].

Additionally, Fig 5B summarizes the contribution energies of amino acid residues found at sub-sites along the active-site cleft ranging from the internal foot pocket, catalytic residues (His145, His146, and Tyr308) and the Zn^{2+} -binding group, the linker-binding, and the cap-binding regions. It is apparent that, in addition to the catalytic residues, several residues within the foot pocket, as well as the linker- and cap-binding sites, also contribute significantly to the ligand binding. Trends of contribution energies for each ligand correlate well with the binding mechanism outlined in the previous section describing intermolecular interactions. For example, binding of SAHA that primarily uses its aniline moiety to plug into the opening surface, but lacks the foot-pocket binding group, gives rise to the contribution energies that predominate in the linker- and cap-binding residues but less so in the internal cavity. On the other hand, binding of a potent ligand 20Y ($IC_{50} = 56.3$ nM) shows strong contribution energies in all regions. This coincides with the intermolecular interactions identified by a spatial visualization from the MD analysis (*vide supra*). Taken together, the results from this experiment suggest that the contribution energy per residue may be used in combination with the intermolecular analysis in order to explain a ligand binding mechanism and ultimately forecast its inhibitory potency. However, evaluation of additional thermodynamic parameters may also be required in order to construct a quantitative model to predict the inhibitory activity of a ligand family.

Estimation of the entropic term

In order to extract additional information from the simulations, namely the entropic term of protein-ligand complex formation, the directions of molecular motions and fluctuations were first analyzed. The “essential dynamics method” was employed to examine conformational changes observed within the HDAC2 protein when bound by ligands of interest. PCA was performed via Gromacs inbuilt tool ‘g_covar’ in order to determine the concerted molecular motions, that resulted in a covariance matrix of backbone $C\alpha$ atoms within the HDAC2-ligand complexes. Fig 6 shows two-dimensional projections of trajectory motions of the HDAC2 proteins when bound by the five ligands. Positive and negative limits of each dimension represent, respectively, correlated and anti-correlated motions of the backbone atoms in the same and opposite directions relative to the average positions. Subsequently, the covariance matrix was diagonalized to obtain the eigenvectors and eigenvalues, where the eigenvector with the highest eigenvalue is assigned as the first principal component, and the one with the second highest eigenvalue is considered the second principal component. The amplitude of eigenvectors along the spatial dimensions thus represents the conformational changes observed within the protein and the ligands.

The results show that conformational changes observed within the HDAC2 protein when bound by different ligands are similar in their amplitudes along the multidimensional spaces. However, HDAC2 when bound by 6EZ has slightly larger changes when compared to the

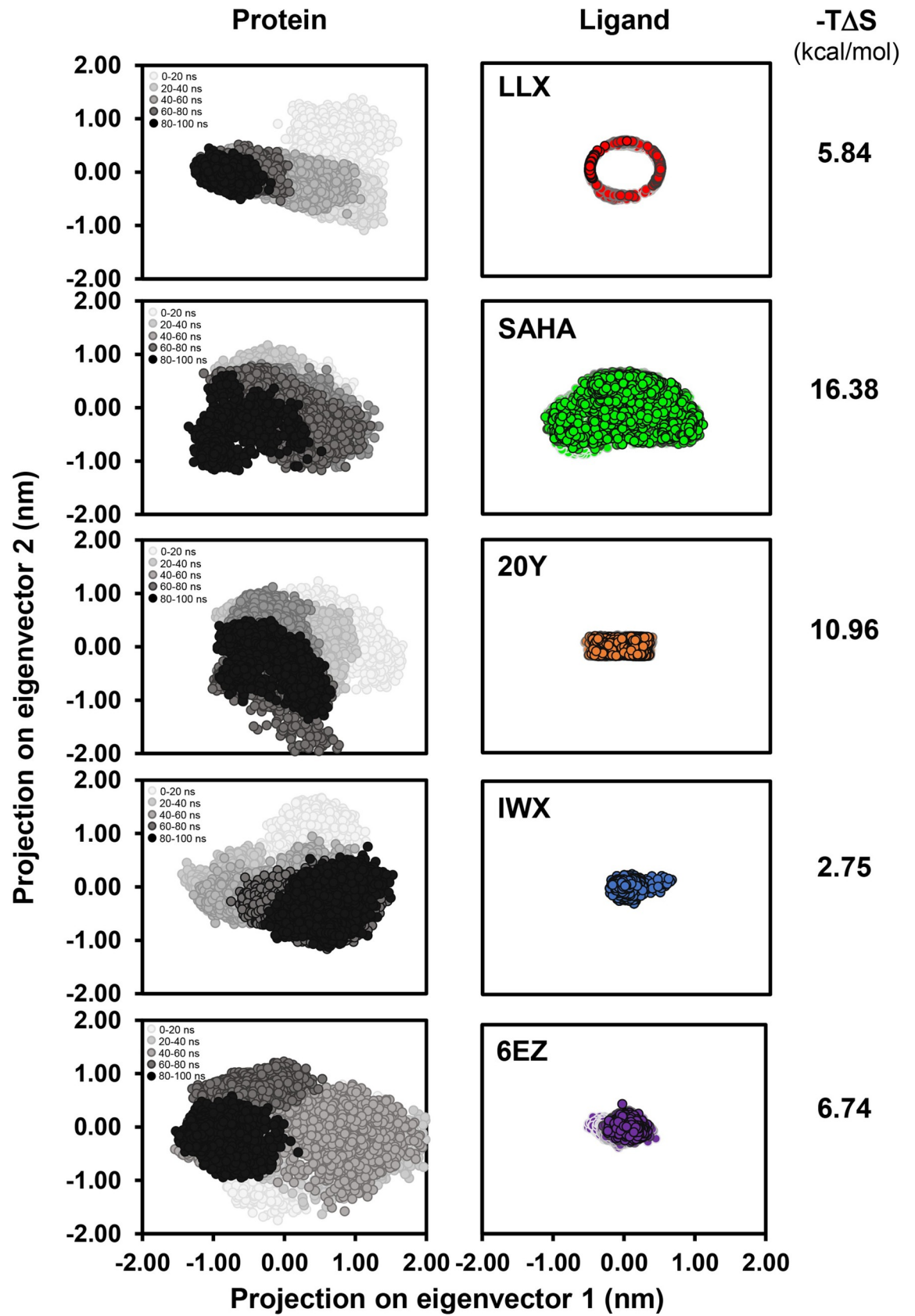


Fig 6. Two-dimensional projections of trajectory motions of the HDAC2 protein (left panels) and the bound ligands (right panels) in multidimensional space over the first two (PC1 and PC2) principal eigenvectors. Calculated $-T\Delta S$ values in kcal/mol of HDAC2-ligand complexes are shown on the right.

<https://doi.org/10.1371/journal.pone.0273265.g006>

others. On the other hand, when considering only ligand motions, SAHA was found to move around the most as the amplitudes on both directions are apparently larger than those found in other inhibitors. These observations could suggest that both HDAC2-6EZ and HDAC2-SAHA complexes may be less energetically stable. Interestingly, inhibitory activities of these two ligands are also the poorest among the five ligands, with IC_{50} values of 168 nM for 6EZ, and 251 nM for SAHA.

To further analyze the conformational results in a more quantitative manner, an entropic term was also derived and compared. To this end, an inbuilt Gromacs function 'g_anaeig' was employed to analyze the energy landscape for the two eigenvector projections and a quasi-harmonic approximation method was applied to derive the entropic term for each protein-ligand binding. The $-T\Delta S$ term was determined and displayed in the right column of Fig 6. It is apparent that the entropic term of all ligand complexes examined are rather different, with the highest being that of the HDAC2-SAHA complex. The vast difference in this term among the complexes piqued our curiosity: if the value of entropic term was added to the $\Delta G_{\text{binding}}$ equation (Eq 2), would improve the prediction of ligand binding and inhibitory activity?

Comparison of the entropic contributions estimated among the different and similar ligand derivatives

To help assess the level of energetic contribution from the entropic term to the overall binding free energy, additional energetic parameters were then collected. Aside from the binding free energy per-residue, the aforementioned MM/PBSA calculation also yielded other additional energetic terms that can be beneficial to our in-depth analysis. Table 1 summarizes the energetic parameters calculated from the MM/PBSA and the quasi-harmonic approximation of all HDAC2-inhibitor complexes. These parameters include ΔE_{vdW} , ΔE_{elec} , ΔG_{polar} , $\Delta G_{\text{non-polar}}$, ΔG_{PBSA} , and $-T\Delta S$, all of which amounted to the ΔG_{total} ($\Delta G_{\text{binding}}$). The related equations of each term are described in the Material and Methods section above. Fig 7A also compares the ΔG_{total} and the $T\Delta S$ terms for all HDAC2-ligand complexes in a bar graph. When considering the sizes of both energetic terms determined in each complex, it is apparent that the entropic term contributes greatly to the total free energy of binding and including this term into the calculation could also change the trend of the ΔG_{total} .

Table 1. Binding free energy of complexes between HDAC2 and known HDAC2 inhibitors based on 100-ns MD simulations and MM-PBSA calculations. The entropic contributions were calculated using the quasi-harmonic (QH) approximation. All values are presented in kcal/mol.

Energy	HDAC2-inhibitor complexes				
	LLX	SAHA	20Y	IWX	6EZ
ΔE_{vdW}	-33.30	-26.27	-38.34	-39.34	-42.82
ΔE_{elec}	-112.99	-103.15	-117.77	-101.20	-101.46
ΔG_{polar}	80.06	58.62	84.07	80.91	79.20
$\Delta G_{\text{non-polar}}$	-3.86	-3.37	-4.37	-3.98	-4.36
ΔG_{PBSA}	-70.08	-74.16	-76.42	-63.62	-69.44
$-T\Delta S$	5.84	16.32	10.96	2.75	6.74
ΔG_{total}	-64.24	-57.84	-65.46	-60.87	-62.70

<https://doi.org/10.1371/journal.pone.0273265.t001>

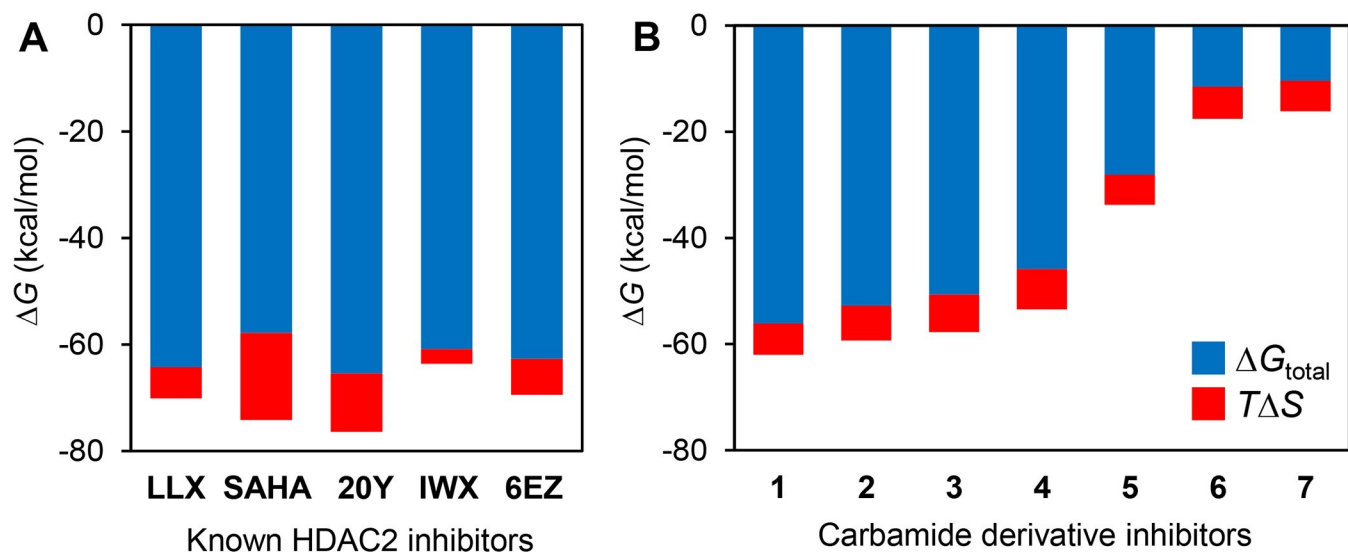


Fig 7. Binding free energies of either known inhibitors (A) or carbamide derivative inhibitors (B) comparing the scale of the entropic term contribution.

<https://doi.org/10.1371/journal.pone.0273265.g007>

From the observation that the entropic term is rather important in a comparison of diverse ligands, we also wondered whether this pattern could also be seen among a set of structurally similar ligands. To this end, we performed the same set of experiments on seven additional carbamide derivative inhibitors of HDAC2 (chemical structures and energetic parameters shown in S4 Table of S1 File). These include molecular docking, MD simulations, MM/PBSA calculations, as well as the approximations of the entropic term. Table 2 summarizes the energetic parameters contributing to the total binding free energy of the complexes of carbamide scaffolds. Interestingly, the entropic terms found within the protein complexes for this structurally similar inhibitor molecules do not appear to be significantly different (Fig 7B). When accounting for both ΔG_{total} and $T\Delta S$ terms of this set of ligands, it is obvious that addition of the entropic term does not change the trend of overall free energy of binding. Therefore, it is possible that, for deriving a trend among structurally similar ligands, including the entropic term may not be very critical. This possibility is being investigated in our laboratory with a larger sample size, and will be reported elsewhere.

Table 2. Binding free energy of complexes between HDAC2 and seven carbamate derivative inhibitors based on 100-ns MD simulations and MM-PBSA calculations. The entropic contributions were calculated using the quasi-harmonic (QH) approximation. All values are presented in kcal/mol.

Energy	Carbamate derivative compounds						
	1	2	3	4	5	6	7
ΔE_{vdw}	-35.80	-34.62	-34.78	-33.53	-23.39	-21.77	-25.85
ΔE_{elec}	-95.61	-88.20	-86.39	-90.26	-98.22	-97.43	-95.54
ΔG_{polar}	73.21	66.92	66.96	70.51	90.89	104.25	108.19
$\Delta G_{\text{non-polar}}$	-3.8	-3.45	-3.54	-3.72	-3.03	-2.62	-2.98
ΔG_{PBSA}	-62.00	-59.34	-57.75	-53.46	-33.75	-17.57	-16.17
$-T\Delta S$	5.88	6.57	7.08	7.60	5.64	6.07	5.75
ΔG_{total}	-56.12	-52.77	-50.67	-45.86	-28.11	-11.49	-10.42

<https://doi.org/10.1371/journal.pone.0273265.t002>

Potential impact of the entropic term towards the inhibitory activity prediction

To further analyze the effect of entropic contribution, we then hypothesized that including the entropic term to calculate the $\Delta G_{\text{binding}}$ may alter a correlation plot between the binding free energy and the actual activity of the ligands. To test this hypothesis, we created a plot between the actual pIC_{50} values of HDAC2 inhibitors and their calculated binding free energy, with (ΔG_{total}) or without the entropic term (ΔG_{PBSA}) (Fig 8). Fig 8A and 8B compare the trends, respectively, before and after adding the values of $-T\Delta S$ term. Although the data were analyzed with only these five different ligands, a trend can be observed when the entropic term was applied to the estimation of binding free energy. However, when considering the other set of structurally similar carbamide ligands before (Fig 8C) and after (Fig 8D) applying the entropic term, the correlation trends were not significantly different. It is also known in the literature that entropic contribution could be neglected when comparing a set of chemically related ligands, but not so for the unrelated ligands [72]. This agrees well with our previous observation that, for comparing derivative ligands, estimations of entropic terms may not be critical.

Additionally, it is noteworthy that the trend between experimental and computational data found within the similar ligands are already highly correlated. This could also imply a possibility that a prediction model created by energetic parameter calculation is more effective among

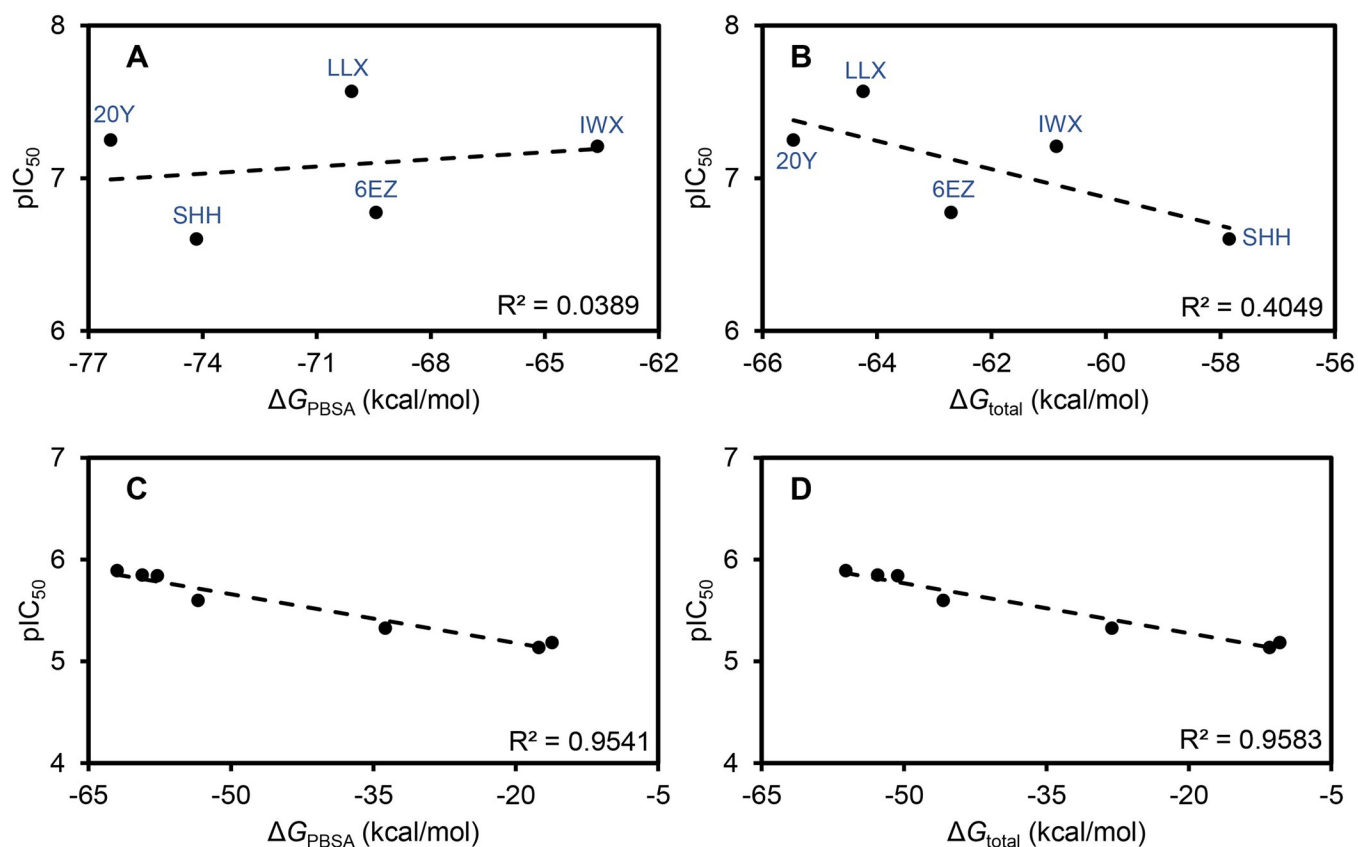


Fig 8. Correlation plots between experimental pIC_{50} values of HDAC2 inhibitors and the binding free energies of their complexes with HDAC2 calculated from MM-PBSA and additional entropic term estimations. Actual pIC_{50} potencies of known HDAC2 inhibitor ligands (A and B) or a set of carbamide derivatives (C and D) are plotted against the binding free energies of their complexes calculated from MM/PBSA alone (ΔG_{PBSA}) (A and C) or calculated as total binding free energies by the addition of entropic term estimations (ΔG_{total}) (B and D).

<https://doi.org/10.1371/journal.pone.0273265.g008>

similar ligands, but not so for diverse ligands. A larger pool of sample is needed, preferably from the same set of experiment, to fully examine this effect in the future.

Noticeably, the order of binding energies calculated from the ligand set changed slightly upon introducing the entropic term. Without the entropy, ΔG_{PBSA} energies ranks from low to high as $20Y < SHH < LLX < 6EZ < IWX$, and with the entropy, the order of ΔG_{total} energies became $20Y < LLX < 6EZ < IWX < SHH$. The change in the ranking of SHH was largely due to the consideration of entropic term, thus signifying the importance of the entropy calculation. However, as the experimental IC_{50} values range from low to high as $LLX < 20Y < IWX < 6EZ < SHH$, the calculated ranking does not correlate well with the experimental data. Additionally, the correlation remains relatively poor even when the entropic term was considered. This could suggest missing factors contributing to the accurate binding energy calculation.

Intriguingly, the correlation from the closely related ligands is relatively high even before the entropy was taken into account (Fig 8C and 8D), while the one from unrelated ligands remains poor. A potential explanation may include the fact that, for the unrelated ligands, MD simulations were performed based upon different crystal structures of HDAC2-ligand complexes, as we wanted to employ the most reliable structures available. This essentially differs from the MD simulations of related ligands as a single protein structure was used as a reference structure, and the ligands were docked into the same active site to generate initial structures for the MD simulations. Discrepancies in the protein structural coordinates among the crystal structure could contribute to the difference in the resulted MD trajectories and thus the calculation of entropic terms. This hypothesis will be tested further in our laboratory with a larger pool of related and unrelated ligands. A potential caveat may include the selection of a single crystal structure to be used as initial coordinates that could represent the entire population of legitimate crystal structures.

Another avenue for investigating the factors contributing to the total binding energy calculation is to consider solvation entropy in addition to the configurational entropy. Several studies have highlighted the role of solvation entropy in protein-ligand complex formation, namely the water molecules around the binding pockets [73–76]. However, a recent study also suggests that configurational entropy dominates over solvation entropy in the overall entropic contribution to ligand binding [77].

As aforementioned, the IC_{50} values reported from different laboratories, though some were from the same isozyme and using similar methods of determination, could varied greatly and thus resulted in uncorrelated trends (S8 Fig in S1 File). Future experiments could also benefit from using K_i values instead of IC_{50} as it is an absolute value that would not vary between experiments. However, for our study, K_i values could not be properly calculated using a theoretical conversion equation as the information of substrate concentration and the K_m values were not explicitly reported in most experiments we reviewed. Therefore, we used a set of reported IC_{50} values those determined by the most similar method of determination. Undoubtedly, for the purpose of prediction model construction in the future, a large number of K_i values, or IC_{50} values determined from the same experiment, should give rise to a more accurate prediction of inhibitory potency.

Employing molecular modeling to predict binding free energies currently helps accelerate drug development by lowering the time and effort required to generate a new set of candidate compounds. To date, there are several successful computational tools to calculate binding free energies in protein-ligand interactions. One of the state-of-the-art strategies is alchemical free energy methods [78–85] that showed a lot of promise in terms of facilitating the computation of binding free energies. Accurate relative free energies could identify whether ligand alterations have boosted affinity and selectivity, especially in lead optimization efforts. Generally, due to their higher rigor in evaluating interactions with explicit solvent and configurational entropy, alchemical approaches are recommended in later stages of drug optimization when

fewer simulations of a restricted range of candidates are required [86]. On the other hand, MM/PBSA is recommended for the initial phases of virtual screening when dealing with a large number of chemical candidates. Although MM/PBSA provides less accurate results when compared with the alchemical methods, the technique requires less computational demand while providing higher accuracy than molecular docking, thus rendering it suitable for a preliminary estimation of binding free energies [87].

The modular nature of MM/PBSA and MM/GBSA, as well as the fact that these techniques do not require a training set in the calculations, make them appealing practices that have been successfully implemented to replicate and interpret experimental findings, as well as to improve virtual screening and docking results [87]. Nonetheless, the techniques contain several problematic approximations for estimating binding free energies, such as the lack of configurational entropy and data on the amount and free energy of water molecules in the binding site. Therefore, in our study, we further employed an estimation of the entropic term via the essential dynamics method in order to examine the concerted molecular motions, which can be translated to configuration entropy. Although this is not merely a method to obtain accurate values of both ΔG_{total} and $T\Delta S$ terms, this simple technique could provide a trend of entropic contribution to the binding of a protein with structurally diverse or similar ligands, which could be beneficial to the initial screening of drug candidates. This is particularly critical for targeting proteins with multiple isoforms like the HDAC enzyme family because entropic contribution has been shown to be a selectivity driver for inhibitor binding [36]. It was found that the binding of several cycloalkenyl hydroxamates selective to HDAC6 exhibited significant entropic gain, whereas the binding of these compounds to HDAC8 showed substantial entropic loss, suggesting that entropy is a crucial contributor to the binding selectivity [36]. Hence, it is essential to consider the entropic term in the approximation of binding free energies even in an early stage of the compound screening.

Conclusion

This study aimed to investigate the binding mechanism that the HDAC2 enzyme employs in differentiating various types of inhibitor ligands. MD simulation data provided a spatial visualization of how different ligands can interact with the sub-sites of the active site cleft. The level of energetic contribution per amino acid residues were estimated using MM/PBSA calculations, which resulted in a quantitative view of the binding behavior. It was found that the interactions within the internal foot pocket can contribute significantly to the ligand binding. Residues lining the opening rim of the active site also provide additional interactions to restrict the movement of ligand molecules. Protruding regions of the ligands lacking these contacts could cause a high fluctuation along the cap portion and thus weaken the overall ligand affinity. To further quantify this dynamics aspect, the entropic term was estimated and compared with the scale of $\Delta G_{\text{binding}}$. It was found that the values of $-T\Delta S$ term correlate with the level of disorder found within the ligand complexes. Furthermore, accounting for the entropic term in the estimation of ΔG_{total} could provide additional insight into its contribution to the binding energy. In particular, we found that the $-T\Delta S$ term may be necessary for a comparison of binding energy among a set of diverse ligands as the levels of their entropic contribution could greatly differ. However, for structurally similar ligands, this term appears within a comparable range and would unlikely change the ranking trend. This study has demonstrated the significance of molecular dynamics and an estimation of entropic contribution in describing inhibitor binding mechanism. Further development of the model to predict ligand affinities and inhibitory activity could thus be pursued in order to help guiding an optimization of ligand candidates to be developed in future drug discovery campaigns.

Supporting information

S1 File. S1 Fig. Root-mean-square deviations (RMSD) of backbone atoms on HDAC2 when bound or unbound with inhibitors, comparing between two independent MD repeats. Unbound HDAC2 (A) or HDAC2 in complex with LLX (B), SAHA (C), 20Y (D), IWX (E), and 6EZ (F) are separately illustrated. S2 Fig. Root-mean-square deviations (RMSD) of backbone atoms on HDAC2 when bound with 20Y ligand, comparing the ranges of fluctuation during the original 100-ns MD run, as used in this study, and the extended 200-ns MD run. S3 Fig. Overlay of snapshot structures of HDAC2 when bound with 20Y ligand obtained at 100 ns and at 200 ns of the MD simulation. RMSD between the heavy atoms of the two structure is 0.8187 Å. S4 Fig. Binding interactions between inhibitor ligands and the active site of HDAC2, highlighting metal chelation with the Zn^{2+} cofactors. Inhibitor ligands LLX (A), SAHA (B), 20Y (C), IWX (D), and 6EZ (E) are shown in stick representations. Zn^{2+} cofactors are shown as purple spheres. S5 Fig. The distribution of distances between the Zn^{2+} cofactor of HDAC2 and a nearby heavy atom of amino acid residues or ligands. S6 Fig. Radius distribution function (RDF), detected during the 100-ns simulations, for the Zn^{2+} ion as a result of interactions with its atomic partners within the catalytic site of HDAC2, which include Asp181-OD1 (blue), Asp181-OD2 [88], His183-ND1 (orange), Asp269-OD1 [60], Asp269-OD2 (brown), and the bound inhibitors (blue and red). S7 Fig. Running coordination number, detected during the 100-ns simulations, for the Zn^{2+} ion as a result of interactions with its atomic partners within the catalytic site of HDAC2, which include Asp181-OD1 (blue), Asp181-OD2 [88], His183-ND1 (orange), Asp269-OD1 [60], Asp269-OD2 (brown), and the bound inhibitors (blue and red). S8 Fig. Correlation plots between experimental pIC_{50} values of HDAC2 inhibitors, collected from previous studies indicated in S5 Table, and the binding free energies of their complexes with HDAC2 calculated from MM-PBSA and additional entropic term estimations. Actual pIC_{50} potencies of known HDAC2 inhibitor ligands are plotted against the binding free energies of their complexes calculated from MM/PBSA alone (ΔG_{PBSA})(A) or calculated as total binding free energies by the addition of entropic term estimations (ΔG_{total})(B). Grey dots represent the pIC_{50} values reported in the various studies, while black dots represent the average value for each inhibitor. S1 Table Protonation states of histidine residue of HDAC2 in pH 7.0 as calculated from the PropKa server (<https://www.ddl.unimi.it/vegaol/propka.htm>). A histidine can adopt three protonation states: HIP (+1 charged, both δ - and ϵ -nitrogens protonated), HID (neutral, δ -nitrogen protonated), and HIE (neutral, ϵ -nitrogen protonated). S2 Table Atomic distances between the Zn^{2+} cofactor of HDAC2 and a nearby heavy atom of amino acid residues or ligands. Values are shown as average distance \pm standard deviation. S3 Table Comparison between biochemical activity (IC_{50}) and *in silico* information of known HDAC2 inhibitors. S4 Table Comparison between biochemical activity (IC_{50})^a and *in silico* information of HDAC2 carbamide derivative inhibitors. S5 Table Comparison of reported IC_{50} values of HDAC inhibitors and the experimental methods used in the determination of inhibitory potency. Values labeled with an asterisk are the values selected for the use in our calculation, which are results from the most similar methods of determination. (PDF)

Acknowledgments

ST and NS conceived the work and directed the experiments. ST performed the MD simulations, MM/PBSA calculations, estimation of entropy, collected and confirmed the data and performed the analysis. ST and NS drafted the first and second versions of the manuscript. Both authors read, edited, and approved the final version of manuscript. We thank the Center

of Excellence in Materials Science and Technology, Faculty of Science, the Materials Science Research Center (MSRC), Chiang Mai University, for the support. We also thank the National Research Council of Thailand, and the Program Management Unit for Human Resources & Institutional Development, Research and Innovation, Office of National Higher Education Science Research and Innovation Policy Council (NXPO). The authors would like to thank Prof. Richard L. Deming from California State University, Fullerton, USA for helpful discussions and English language editing.

Author Contributions

Conceptualization: Suriya Tateing, Nuttee Suree.

Data curation: Suriya Tateing.

Formal analysis: Suriya Tateing, Nuttee Suree.

Funding acquisition: Nuttee Suree.

Investigation: Nuttee Suree.

Methodology: Suriya Tateing, Nuttee Suree.

Project administration: Suriya Tateing, Nuttee Suree.

Resources: Nuttee Suree.

Software: Suriya Tateing.

Supervision: Nuttee Suree.

Validation: Suriya Tateing, Nuttee Suree.

Visualization: Suriya Tateing.

Writing – original draft: Suriya Tateing, Nuttee Suree.

Writing – review & editing: Suriya Tateing, Nuttee Suree.

References

1. Cai C, Yu H, Huang G, Du X, Yu X, Zhou Y, et al. Histone modifications in fatty acid synthase modulated by carbohydrate responsive element binding protein are associated with non-alcoholic fatty liver disease. *Int J Mol Med*. 2018; 42(3):1215–28. <https://doi.org/10.3892/ijmm.2018.3702> PMID: 29786745
2. Ferriero R, Nusco E, De Cegli R, Carissimo A, Manco G, Brunetti-Pierrri N. Pyruvate dehydrogenase complex and lactate dehydrogenase are targets for therapy of acute liver failure. *J Hepatol*. 2018; 69(2):325–35. <https://doi.org/10.1016/j.jhep.2018.03.016> PMID: 29580866
3. Kanyal A, Rawat M, Gurung P, Choubey D, Anamika K, Karmodiya K. Genome-wide survey and phylogenetic analysis of histone acetyltransferases and histone deacetylases of *Plasmodium falciparum*. *FEBS J*. 2018; 285(10):1767–82. <https://doi.org/10.1111/febs.14376> PMID: 29284196
4. Berger SL. The complex language of chromatin regulation during transcription. *Nature*. 2007; 447(7143):407–12. <https://doi.org/10.1038/nature05915> PMID: 17522673
5. Khangura RK, Bali A, Jaggi AS, Singh N. Histone acetylation and histone deacetylation in neuropathic pain: An unresolved puzzle? *Eur J Pharmacol*. 2017; 795:36–42. <https://doi.org/10.1016/j.ejphar.2016.12.001> PMID: 27916557
6. Kouzarides T. Chromatin modifications and their function. *Cell*. 2007; 128(4):693–705. <https://doi.org/10.1016/j.cell.2007.02.005> PMID: 17320507
7. Haberland M, Montgomery RL, Olson EN. The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nat Rev Genet*. 2009; 10(1):32–42. <https://doi.org/10.1038/nrg2485> PMID: 19065135
8. Mahady GB. Medicinal plants for the prevention and treatment of bacterial infections. *Curr Pharm Des*. 2005; 11(19):2405–27. <https://doi.org/10.2174/1381612054367481> PMID: 16026296

9. Lin CL, Tsai ML, Lin CY, Hsu KW, Hsieh WS, Chi WM, et al. HDAC1 and HDAC2 Double Knockout Triggers Cell Apoptosis in Advanced Thyroid Cancer. *Int J Mol Sci*. 2019; 20(2). <https://doi.org/10.3390/ijms20020454> PMID: 30669676
10. Stojanovic N, Hassan Z, Wirth M, Wenzel P, Beyer M, Schäfer C, et al. HDAC1 and HDAC2 integrate the expression of p53 mutants in pancreatic cancer. *Oncogene*. 2017; 36(13):1804–15. <https://doi.org/10.1038/onc.2016.344> PMID: 27721407
11. Noh H, Oh EY, Seo JY, Yu MR, Kim YO, Ha H, et al. Histone deacetylase-2 is a key regulator of diabetes- and transforming growth factor-beta1-induced renal injury. *Am J Physiol Renal Physiol*. 2009; 297(3):F729–39. <https://doi.org/10.1152/ajprenal.00086.2009> PMID: 19553350
12. Yang FQ, Liu M, Yang FP, Che J, Li W, Zhai W, et al. VPA inhibits renal cancer cell migration by targeting HDAC2 and down-regulating HIF-1 α . *Mol Biol Rep*. 2014; 41(3):1511–8. <https://doi.org/10.1007/s11033-013-2996-2> PMID: 24390319
13. Fritzsche FR, Weichert W, Röske A, Gekeler V, Beckers T, Stephan C, et al. Class I histone deacetylases 1, 2 and 3 are highly expressed in renal cell cancer. *BMC Cancer*. 2008; 8:381. <https://doi.org/10.1186/1471-2407-8-381> PMID: 19099586
14. Shang L, Pin L, Zhu S, Zhong X, Zhang Y, Shun M, et al. Plantamajoside attenuates isoproterenol-induced cardiac hypertrophy associated with the HDAC2 and AKT/ GSK-3 β signaling pathway. *Chem Biol Interact*. 2019; 307:21–8. <https://doi.org/10.1016/j.cbi.2019.04.024> PMID: 31009642
15. Datta M, Staszewski O, Raschi E, Frosch M, Hagemeyer N, Tay TL, et al. Histone Deacetylases 1 and 2 Regulate Microglia Function during Development, Homeostasis, and Neurodegeneration in a Context-Dependent Manner. *Immunity*. 2018; 48(3):514–29.e6. <https://doi.org/10.1016/j.immuni.2018.02.016> PMID: 29548672
16. Bin YF, Wu LJ, Sun XJ, Liang Y, Bai J, Zhang JQ, et al. Expression of GR- α and HDAC2 in steroid-Sensitive and steroid-Insensitive interstitial lung disease. *Biomed Pharmacother*. 2019; 118:109380. <https://doi.org/10.1016/j.biopha.2019.109380> PMID: 31545224
17. Tang W, Zhou W, Xiang L, Wu X, Zhang P, Wang J, et al. The p300/YY1/miR-500a-5p/HDAC2 signaling axis regulates cell proliferation in human colorectal cancer. *Nat Commun*. 2019; 10(1):663. <https://doi.org/10.1038/s41467-018-08225-3> PMID: 30737378
18. Lai T, Wu M, Zhang C, Che L, Xu F, Wang Y, et al. HDAC2 attenuates airway inflammation by suppressing IL-17A production in HDM-challenged mice. *Am J Physiol Lung Cell Mol Physiol*. 2019; 316(1):L269–L79. <https://doi.org/10.1152/ajplung.00143.2018> PMID: 30407865
19. Barnes PJ. Corticosteroid resistance in patients with asthma and chronic obstructive pulmonary disease. *J Allergy Clin Immunol*. 2013; 131(3):636–45. <https://doi.org/10.1016/j.jaci.2012.12.1564> PMID: 23360759
20. Wang Z, Zang C, Cui K, Schones DE, Barski A, Peng W, et al. Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell*. 2009; 138(5):1019–31. <https://doi.org/10.1016/j.cell.2009.06.049> PMID: 19698979
21. Brunmeir R, Lagger S, Seiser C. Histone deacetylase HDAC1/HDAC2-controlled embryonic development and cell differentiation. *Int J Dev Biol*. 2009; 53(2–3):275–89. <https://doi.org/10.1387/ijdb.082649rb> PMID: 19412887
22. Bressi JC, Jennings AJ, Skene R, Wu Y, Melkus R, De Jong R, et al. Exploration of the HDAC2 foot pocket: Synthesis and SAR of substituted N-(2-aminophenyl)benzamides. *Bioorg Med Chem Lett*. 2010; 20(10):3142–5. <https://doi.org/10.1016/j.bmcl.2010.03.091> PMID: 20392638
23. Hou J, Feng C, Li Z, Fang Q, Wang H, Gu G, et al. Structure-based optimization of click-based histone deacetylase inhibitors. *Eur J Med Chem*. 2011; 46(8):3190–200. <https://doi.org/10.1016/j.ejmech.2011.04.027> PMID: 21621883
24. Zhou H, Wang C, Ye J, Chen H, Tao R. Design, virtual screening, molecular docking and molecular dynamics studies of novel urushiol derivatives as potential HDAC2 selective inhibitors. *Gene*. 2017; 637:63–71. <https://doi.org/10.1016/j.gene.2017.09.034> PMID: 28939339
25. Xie R, Yao Y, Tang P, Chen G, Liu X, Yun F, et al. Design, synthesis and biological evaluation of novel hydroxamates and 2-aminobenzamides as potent histone deacetylase inhibitors and antitumor agents. *Eur J Med Chem*. 2017; 134:1–12. <https://doi.org/10.1016/j.ejmech.2017.03.038> PMID: 28391133
26. Liu J, Zhou J, He F, Gao L, Wen Y, Wang P, et al. Design, synthesis and biological evaluation of novel indazole-based derivatives as potent HDAC inhibitors via fragment-based virtual screening. *Eur J Med Chem*. 2020; 192:112189. <https://doi.org/10.1016/j.ejmech.2020.112189> PMID: 32151834
27. Fournier JF, Bhurruth-Alcor Y, Musicki B, Aubert J, Aurelly M, Bouix-Peter C, et al. Squaramides as novel class I and IIB histone deacetylase inhibitors for topical treatment of cutaneous t-cell lymphoma. *Bioorg Med Chem Lett*. 2018; 28(17):2985–92. <https://doi.org/10.1016/j.bmcl.2018.06.029> PMID: 30122227

28. Yun F, Cheng C, Ullah S, He J, Zahi MR, Yuan Q. Thioether-based 2-aminobenzamide derivatives: Novel HDAC inhibitors with potent in vitro and in vivo antitumor activity. *Eur J Med Chem.* 2019; 176:195–207. <https://doi.org/10.1016/j.ejmech.2019.05.007> PMID: 31103900
29. Alsawalha M, Rao Bolla S, Kandakatla N, Srinivasadesikan V, Veeraraghavan VP, Surapaneni KM. Molecular docking and ADMET analysis of hydroxamic acids as HDAC2 inhibitors. *Bioinformation.* 2019; 15(6):380–7. <https://doi.org/10.6026/97320630015380> PMID: 31312074
30. Caro JA, Harpole KW, Kasinath V, Lim J, Granja J, Valentine KG, et al. Entropy in molecular recognition by proteins. *Proc Natl Acad Sci U S A.* 2017; 114(25):6563–8. <https://doi.org/10.1073/pnas.1621154114> PMID: 28584100
31. Dewaker V, Srivastava PN, Verma S, Srivastava AK, Prabhakar YS. Non-bonding energy directed designing of HDAC2 inhibitors through molecular dynamics simulation. *J Biomol Struct Dyn.* 2021:1–24. <https://doi.org/10.1080/07391102.2021.1989037> PMID: 34662251
32. Hao GF, Wang F, Li H, Zhu XL, Yang WC, Huang LS, et al. Computational discovery of picomolar Q(o) site inhibitors of cytochrome bc1 complex. *J Am Chem Soc.* 2012; 134(27):11168–76. <https://doi.org/10.1021/ja3001908> PMID: 22690928
33. Pan Y, Gao D, Zhan CG. Modeling the catalysis of anti-cocaine catalytic antibody: competing reaction pathways and free energy barriers. *J Am Chem Soc.* 2008; 130(15):5140–9. <https://doi.org/10.1021/ja077972s> PMID: 18341277
34. Karplus M, Kushick JN. Method for estimating the configurational entropy of macromolecules. *Macromolecules.* 1981; 14(2):325–32.
35. Chang CE, Chen W, Gilson MK. Evaluating the Accuracy of the Quasiharmonic Approximation. *J Chem Theory Comput.* 2005; 1(5):1017–28. <https://doi.org/10.1021/ct0500904> PMID: 26641917
36. Porter NJ, Wagner FF, Christianson DW. Entropy as a Driver of Selectivity for Inhibitor Binding to Histone Deacetylase 6. *Biochemistry.* 2018; 57(26):3916–24. <https://doi.org/10.1021/acs.biochem.8b00367> PMID: 29775292
37. Berman HM. The protein data bank and the challenge of structural genomics. *Nature Struct Biol.* 2000; 7:957–9. <https://doi.org/10.1038/80734> PMID: 11103999
38. Tian C, Kasavajhala K, Belfon KAA, Raguette L, Huang H, Miguels AN, et al. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J Chem Theory Comput.* 2020; 16(1):528–52. <https://doi.org/10.1021/acs.jctc.9b00591> PMID: 31714766
39. Lauffer BE, Mintzer R, Fong R, Mukund S, Tam C, Zilberleyb I, et al. Histone deacetylase (HDAC) inhibitor kinetic rate constants correlate with cellular histone acetylation but not transcription and cell viability. *J Biol Chem.* 2013; 288(37):26926–43. <https://doi.org/10.1074/jbc.M113.490706> PMID: 23897821
40. Wagner FF, Weiwer M, Steinbacher S, Schomburg A, Reinemer P, Gale JP, et al. Kinetic and structural insights into the binding of histone deacetylase 1 and 2 (HDAC1, 2) inhibitors. *Bioorg Med Chem.* 2016; 24(18):4008–15. <https://doi.org/10.1016/j.bmc.2016.06.040> PMID: 27377864
41. Jakalian A, Jack DB, Bayly CI. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem.* 2002; 23(16):1623–41. <https://doi.org/10.1002/jcc.10128> PMID: 12395429
42. Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, et al. Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* 2018; 27(1):112–28. <https://doi.org/10.1002/pro.3280> PMID: 28836357
43. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput.* 2015; 11(8):3696–713. <https://doi.org/10.1021/acs.jctc.5b00255> PMID: 26574453
44. Zeng J, Li Y, Zhang JZ, Mei Y. Examination of the quality of various force fields and solvation models for the equilibrium simulations of GA88 and GB88. *J Mol Model.* 2016; 22(8):177. <https://doi.org/10.1007/s00894-016-3027-8> PMID: 27392746
45. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. *J Comput Chem.* 2004; 25(9):1157–74. <https://doi.org/10.1002/jcc.20035> PMID: 15116359
46. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics.* 1983; 79:926–35.
47. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. *J Comput Chem.* 2005; 26(16):1701–18. <https://doi.org/10.1002/jcc.20291> PMID: 16211538
48. Case DA, Aktulga HM, Belfon K, Ben-Shalom I, Brozell SR, Cerutti DS, et al. Amber 2021: University of California, San Francisco; 2021.
49. Li P, Merz KM. Taking into Account the Ion-induced Dipole Interaction in the Nonbonded Model of Ions. *J Chem Theory Comput.* 2014; 10(1):289–97. <https://doi.org/10.1021/ct400751u> PMID: 24659926

50. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. *Journal of Chemical Physics*. 1995; 103:8577–93.
51. Darden T, York D, Pedersen L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *Journal of Chemical Physics*. 1993; 98:10089–92.
52. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A Linear Constraint Solver for Molecular Simulations. *Journal of Computational Chemistry*. 1997; 18:1463–72.
53. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *Journal of Chemical Physics*. 1984; 81:3684–90.
54. Kräutler V, Van Gunsteren W, Hünenberger P. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J Comput Chem*. 2001; 22(5):501–8.
55. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *Journal of Molecular Graphics*. 1996; 14(1):33–8. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) PMID: 8744570
56. Xu J, Zhang Y, Chen H, Wang P, Xie Z, Yao Y, et al. Effect of surfactant headgroups on the oil/water interface: An interfacial tension measurement and simulation study. *Journal of Molecular Structure*. 2013; 1052:50–6.
57. Gao S, Kang Z, Yuan R, Liu N, Zhu P, Wang B. Molecular dynamics study of nonylphenol-substituted dodecyl sulfonate at air/water interface: Role of steric effect of surfactant headgroups. *Journal of Molecular Structure*. 2019; 1192:35–41.
58. Mesentean S, Fischer S, Smith JC. Analyzing large-scale structural change in proteins: comparison of principal component projection and Sammon mapping. *Proteins*. 2006; 64(1):210–8. <https://doi.org/10.1002/prot.20981> PMID: 16617427
59. Kumari R, Kumar R, Lynn A, Consortium OSDD. g_mmpbsa—a GROMACS tool for high-throughput MM-PBSA calculations. *J Chem Inf Model*. 2014; 54(7):1951–62. <https://doi.org/10.1021/ci500020m> PMID: 24850022
60. Jordan A, Defechereux P, Verdin E. The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. *EMBO J*. 2001; 20(7):1726–38. <https://doi.org/10.1093/emboj/20.7.1726> PMID: 11285236
61. Yang CY, Phillips JG, Stuckey JA, Bai L, Sun H, Delproposito J, et al. Buried Hydrogen Bond Interactions Contribute to the High Potency of Complement Factor D Inhibitors. *ACS Med Chem Lett*. 2016; 7(12):1092–6. <https://doi.org/10.1021/acsmchemlett.6b00299> PMID: 27994744
62. Chen D, Oezguen N, Urvil P, Ferguson C, Dann SM, Savidge TC. Regulation of protein-ligand binding affinity by hydrogen bond pairing. *Sci Adv*. 2016; 2(3):e1501240. <https://doi.org/10.1126/sciadv.1501240> PMID: 27051863
63. Itoh Y, Nakashima Y, Tsukamoto S, Kurohara T, Suzuki M, Sakae Y, et al. *N. Sci Rep*. 2019; 9(1):767.
64. Schmidtke P, Luque FJ, Murray JB, Barril X. Shielded hydrogen bonds as structural determinants of binding kinetics: application in drug design. *J Am Chem Soc*. 2011; 133(46):18903–10. <https://doi.org/10.1021/ja207494u> PMID: 21981450
65. Kalal BS, Pai VR, Behera SK, Somashekarappa HM. HDAC2 Inhibitor Valproic Acid Increases Radiation Sensitivity of Drug-Resistant Melanoma Cells. *Med Sci (Basel)*. 2019; 7(3). <https://doi.org/10.3390/medsci7030051> PMID: 30909413
66. Kandakatla N, Ramakrishnan G. Ligand Based Pharmacophore Modeling and Virtual Screening Studies to Design Novel HDAC2 Inhibitors. *Adv Bioinformatics*. 2014; 2014:812148. <https://doi.org/10.1155/2014/812148> PMID: 25525429
67. Lombardi PM, Cole KE, Dowling DP, Christianson DW. Structure, mechanism, and inhibition of histone deacetylases and related metalloenzymes. *Curr Opin Struct Biol*. 2011; 21(6):735–43. <https://doi.org/10.1016/j.sbi.2011.08.004> PMID: 21872466
68. Holmes MA, Matthews BW. Binding of hydroxamic acid inhibitors to crystalline thermolysin suggests a pentacoordinate zinc intermediate in catalysis. *Biochemistry*. 1981; 20(24):6912–20. <https://doi.org/10.1021/bi00527a026> PMID: 7317361
69. Chaudhary N, Aparoy P. Application of per-residue energy decomposition to identify the set of amino acids critical for. *Heliyon*. 2020; 6(10):e04944. <https://doi.org/10.1016/j.heliyon.2020.e04944> PMID: 33083581
70. Cele FN, Ramesh M, Soliman ME. Per-residue energy decomposition pharmacophore model to enhance virtual screening in drug discovery: a study for identification of reverse transcriptase inhibitors as potential anti-HIV agents. *Drug Des Devel Ther*. 2016; 10:1365–77. <https://doi.org/10.2147/DDDT.S95533> PMID: 27114700
71. da Costa KS, Galúcio JM, da Costa CHS, Santana AR, Dos Santos Carvalho V, do Nascimento LD, et al. Exploring the Potentiality of Natural Products from Essential Oils as Inhibitors of Odorant-Binding Proteins: A Structure- and Ligand-Based Virtual Screening Approach To Find Novel Mosquito

- Repellents. *ACS Omega*. 2019; 4(27):22475–86. <https://doi.org/10.1021/acsomega.9b03157> PMID: 31909330
72. Winkler DA. Ligand Entropy Is Hard but Should Not Be Ignored. *J Chem Inf Model*. 2020; 60(10):4421–3. <https://doi.org/10.1021/acs.jcim.0c01146> PMID: 33100015
 73. Syme NR, Dennis C, Bronowska A, Paesen GC, Homans SW. Comparison of entropic contributions to binding in a "hydrophilic" versus "hydrophobic" ligand-protein interaction. *J Am Chem Soc*. 2010; 132(25):8682–9. <https://doi.org/10.1021/ja101362u> PMID: 20524663
 74. Haider K, Wickstrom L, Ramsey S, Gilson MK, Kurtzman T. Enthalpic Breakdown of Water Structure on Protein Active-Site Surfaces. *J Phys Chem B*. 2016; 120(34):8743–56. <https://doi.org/10.1021/acs.jpcc.6b01094> PMID: 27169482
 75. Fenley AT, Muddana HS, Gilson MK. Entropy-enthalpy transduction caused by conformational shifts can obscure the forces driving protein-ligand binding. *Proc Natl Acad Sci U S A*. 2012; 109(49):20006–11. <https://doi.org/10.1073/pnas.1213180109> PMID: 23150595
 76. Young T, Abel R, Kim B, Berne BJ, Friesner RA. Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding. *Proc Natl Acad Sci U S A*. 2007; 104(3):808–13. <https://doi.org/10.1073/pnas.0610202104> PMID: 17204562
 77. Verteramo ML, Stenström O, Ignjatović MM, Caldaru O, Olsson MA, Manzoni F, et al. Interplay between Conformational Entropy and Solvation Entropy in Protein-Ligand Binding. *J Am Chem Soc*. 2019; 141(5):2012–26. <https://doi.org/10.1021/jacs.8b11099> PMID: 30618244
 78. Kirkwood JG. Statistical Mechanics of Fluid Mixtures. *J Chem Phys*. 1935; 3:300–13.
 79. Zwanzig RW. High-Temperature Equation of State by a Perturbation Method I. Nonpolar Gases. *J Chem Phys* 1954; 22:1420–6.
 80. Bennett CH. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J Comput Phys*. 1976; 22:245–68.
 81. Straatsma TP, and Mccammon J. A. Multiconfiguration Thermodynamic Integration. *J Chem Phys* 1991; 95:1175–88.
 82. Gilson MK, Given J. A., Bush B. L., and Mccammon J. A. The Statistical-Thermodynamic Basis for Computation of Binding Affinities: a Critical Review. *Biophysical J*. 1997; 72:1047–69. [https://doi.org/10.1016/S0006-3495\(97\)78756-3](https://doi.org/10.1016/S0006-3495(97)78756-3) PMID: 9138555
 83. Boresch S, Tettinger F., Leitgeb M., and Karplus M. Absolute Binding Free Energies: A Quantitative Approach for Their Calculation. *J Phys Chem B* 2003; 107:9535–51.
 84. Shirts MR. Best Practices in Free Energy Calculations for Drug Design. *Methods Mol Biol* 2012; 819:425–67. https://doi.org/10.1007/978-1-61779-465-0_26 PMID: 22183551
 85. Chodera JD, Mobley DL, Shirts MR, Dixon RW, Branson K, Pande VS. Alchemical free energy methods for drug discovery: progress and challenges. *Curr Opin Struct Biol*. 2011; 21(2):150–60. <https://doi.org/10.1016/j.sbi.2011.01.011> PMID: 21349700
 86. King E, Aitchison E, Li H, Luo R. Recent Developments in Free Energy Calculations for Drug Discovery. *Front Mol Biosci*. 2021; 8:712085. <https://doi.org/10.3389/fmolb.2021.712085> PMID: 34458321
 87. Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov*. 2015; 10(5):449–61. <https://doi.org/10.1517/17460441.2015.1032936> PMID: 25835573
 88. Wightman F, Lu HK, Solomon AE, Saleh S, Harman AN, Cunningham AL, et al. Entinostat is a histone deacetylase inhibitor selective for class 1 histone deacetylases and activates HIV production from latently infected primary T cells. *AIDS*. 2013; 27(18):2853–62. <https://doi.org/10.1097/QAD.000000000000067> PMID: 24189584