Discovery of novel hit compounds as potential HDAC1 inhibitors: The case of ligand and structure-based virtual screening

Hajar Sirous[a],*, Giuseppe Campiani[b], Vincenzo Calderone[c], Simone Brogi[c],*

[a] Bioinformatics Research Center, School of Pharmacy and Pharmaceutical Sciences, Isfahan University of Medical Sciences, Isfahan, Iran
[b] Department of Excellence of Biotechnology, Chemistry and Pharmacy, 2018-2022, University of Siena, via Aldo Moro 2, 53100, Siena, Italy
[c] Department of Pharmacy, University of Pisa, via Bonanno 6, 56126, Pisa, Italy

*Corresponding authors:

E-mail addresses: h_sirous@pharm.mui.ac.ir (H. Sirous), simone.brogi@unipi.it (S. Brogi).

This is the accepted version of an article published on *Computers in Biology and Medicine*. Version of record at https://doi.org/10.1016/j.compbiomed.2021.104808

Abstract

Histone deacetylases (HDACs) as an important family of epigenetic regulatory enzymes are implicated in the onset and progression of carcinomas. As a result, HDAC inhibition has been proven as a compelling strategy for reversing the aberrant epigenetic changes associated with cancer. However, non-selective profile of most developed HDAC inhibitors (HDACIs) leads to the occurrence of various side effects, limiting their clinical utility. This evidence provides a solid ground for ongoing research aimed at identifying isoform-selective inhibitors. Among the isoforms, HDAC1 have particularly gained increased attention as a preferred target for the design of selective HDACIs. Accordingly, in this paper, we have developed a reliable virtual screening process, combining different ligand- and structure–based methods, to identify novel benzamide-based analogs with potential HDAC1 inhibitory activity. For this purpose, a focused library of 736,160 compounds from PubChem database was first compiled based on 80% structural similarity with four known benzamide-based HDAC1 inhibitors, Mocetinostat, Entinostat, Tacedinaline, and Chidamide. Our inclusive in-house 3D-QSAR model, derived from pharmacophore-based alignment, was then employed as a 3D-query to discriminate hits with the highest predicted HDAC1 inhibitory activity. The selected hits were subjected to subsequent structure-based approaches (induced-fit docking (IFD), MM-GBSA calculations and molecular dynamics (MD) simulation) to retrieve potential compounds with the highest binding affinity for HDAC1 active site. Additionally, in silico ADMET properties and PAINS filtration were also considered for selecting an enriched set of the best drug-like molecules. Finally, six top-ranked hit molecules, CID_38265326, CID_56064109, CID_8136932, CID_55802151, CID_133901641 and CID_18150975 were identified to expose the best stability profiles and binding mode in the HDAC1 active site. The IFD and MD results cooperatively confirmed the interactions of the promising selected hits with critical residues within HDAC1 active site. In summary, the presented computational approach can provide

a set of guidelines for the further development of improved benzamide-based derivatives targeting HDAC1 isoform.

Keywords: Cancer, HDAC1 inhibitors, Drug design, virtual screening, induced-fit docking, Molecular dynamics simulation

1. Introduction

Histone deacetylases (HDACs) represent one of the valuable targets in antineoplastic drug design, mainly due to their potential for regulation of fundamental life phenomena [1-3]. They regulate gene expression and cell cycle progression by removing acetyl groups from lysine residues located on the N-terminal tail of the histone proteins. The deacetylation causes structural condensation of chromatin in a way that impedes the accessibility of transcription factors to nucleosomal DNA, suppressing the expression of key pro-apoptotic genes [4-6].

According to sequence homology and catalytic mechanism, the 18-known human HDAC isoforms are grouped into two main categories: the “classical” HDACs and the “sirtuin” proteins. This enzyme superfamily can be further subdivided into four classes as well. The classical HDACs are zinc-dependent metalloenzymes that comprise class I (HDAC1–3 and 8), II (HDACs 4–7, 9 and 10) and IV (HDAC11). On the contrary, sirtuins (SIRT1-7), also described as class III HDACs, require NAD+ as a cofactor for their enzymatic activity [4, 7, 8].

The aberrant activation of HDACs and corresponding hypo-acetylation of histones has long been associated with the development of various types of human cancers. Consequently, HDAC inhibitors (HDACIs) were found to induce cell cycle arrest, differentiation, apoptosis, inhibition of proliferation and cytostasis in tumor cells through transcriptional re-activation of suppressed genes [9-11]. The FDA approval of some structurally diverse HDACIs for clinical use, including, Vorinostat [12], Belinostat [13], Panobinostat [14], Chidamide [15], and Romidepsin well validated HDAC family members as potential drug targets for cancer therapy [16]. Moreover, many HDACIs such as Mocetinostat [17], Entinostat [18], Tacedinaline [19], Givinostat [20], and Abexinostat [21] are currently undergoing clinical trials for treating various solid tumors and hematologic malignancies. The chemical structures of these inhibitors are shown in Figure 1.

Generally, a widely accepted pharmacophore model for most HDACIs is constituted of three distinct structural features: a) a zinc-binding group (ZBG), b) a surface recognition group (cap group), and c) a linker. The first motif, ZBG, plays a vital role in interacting with the catalytic zinc ion at the bottom of the active binding site of enzymes. The cap group was involved in hydrophobic interactions with residues on the outer surface of the enzyme, leading to block the entrance of the HDAC active pocket. The linker occupies the long and narrow channel between the catalytic pocket and the outer surface of the enzyme, thereby provide the tubular access to the active site [22-24].

Hydroxamates and benzamides are renowned ZBGs that have frequently been exploited in the design of most potential HDACIs (Figure 1) [2, 24]. Unfortunately, the vast majority of hydroxamate derivatives nonspecifically target multiple HDAC isoforms, behaving as

pan-inhibitors [25, 26]. It is now well proved that insufficient selectivity is main drawback of pan-inhibitors, which contributed in the occurrence of the undesired off-target activities leading to severe toxicity [27, 28]. Thus, increasing efforts have recently been devoted to discovery of non-hydroxamate HDACIs with improved selectivity toward a specific isoform due to reduced side effects, improved clinical efficacy and superior therapeutic index. Moreover, isoform-selective inhibitors would provide chemical tools for probing the precise functions of individual HDAC isoforms in human diseases [29, 30]. In this frame, benzamide derivatives have caught more attention as the most successful class of non-hydroxamate HDACIs in the isoform-selective HDACIs research owing to excellent selectivity for class I HDACs, particularly HDAC1 [31, 32]. In particular, HDAC1 was significantly overexpressed in various cancers, including prostate, gastric, breast, pancreatic and colon cancer [33-37]. Furthermore, the role of this isoform in tumorigenesis is not only linked with histone deacetylation, but also displays unique features through forming complexes with a variety of non-histone proteins. For example, the HDAC1 acts as a pivotal component in interaction with tumor suppressor protein p53, which is a controlling factor on cell proliferation and differentiation [38].

Nowadays, in silico screening methodologies significantly support the modern drug discovery process [39]. In this regard, academia and pharmaceutical companies extensively exploit high-throughput virtual screening (HTVS) strategy to accelerate the identification of potential bioactive molecules for a given drug target by searching in large molecular databases. The cardinal goal of such methods is to minimize failures in the final stages of development of a newfangled drug in a fast and cost-efficient way [40]. Furthermore, a combination of two complementary VS approaches (ligand- and structure-based) is often undertaken to compensate the shortcomings of each individual method, leading to better productivity of virtual screening outcomes (41-45).

Although various virtual screening and molecular modeling approaches have been assisted in the search for HDACIs [46-51], a limited number of chemical libraries have been surveyed for the discovery of potent and selective HDAC1 inhibitors [52]. Thus, this necessitates the demand for an extension of chemical space to achieve structural diverse selective HDAC1 inhibitors possessing less toxic and more potent profiles. Accordingly, we have recently reported a meaningful pharmacophore-based 3D-QSAR model using a large set of benzamide derivatives as valuable selective HDAC1 inhibitors [53]. The developed 3D-QSAR model represented a robust tool for predicting HDAC1 inhibitory activity, possessing a rationale for its employment in virtual screening campaign [53]. To complement these efforts, in the current study, we decided to exploit this model as a convenient filtering tool for further seeking benzamide-based chemotypes as selective HDAC1 inhibitors in online PubChem database. In particular, an exhaustive virtual screening protocol based on the mentioned 3D-QSAR model coupled with an extensive flexible molecular docking and MM-GBSA rescoring calculation was implemented in a stepwise-filtering approach. The sequential virtual screening applied in this study is illustrated in Figure 2, from which we can witness the number of hits reduced after each screening step. The 3D-QSAR model allowed us to discriminate “hits” that satisfy the chemical and the geometrical aspects governing HDAC1 inhibitory activity. Subsequently, structure-based searching was used to dock each “hit” to the HDAC1 active site and then, rank their binding affinities. Further, to enrich the final potential hits with desirable pharmacokinetics (Absorption, Distribution, Metabolism and Excretion; ADME) profiles, in

silico drug-likeness and physicochemical filters have been incorporated in this screening. Finally, to examine stabilities of ligand binding modes and validate screening outcomes, 30 ns molecular dynamics (MD) simulations of HDAC1 complexes with respective best-ranked ligands was performed. The in silico screening workflow described here, could aid in expediting and streamline discovery and development campaign of selective HDAC1 inhibitors for cancer treatment.

2. Material and methods

2.1. Hardware and Software Specifications

All computations in this study were performed using molecular modeling package from Schrödinger's Drug Discovery Suite 2015 (Schrodinger, Inc., LLC, New York, USA) installed on an Intel Xeon CPU E5-2620 v2 @ 3.30 GHz, 64 GB RAM with 12 processors, and a 2 GB graphics card of NVIDIA Quadro K2200 running Ubuntu 10.04 LTS (long- term support) as the operating system. Access to the Schrödinger modules as well as the capability to organize and analyze data was provided by Maestro as a portal interface of Schrödinger [54].

2.2 Dataset and library preparation

The PubChem is a free online server comprising more than 96 billion unique chemical structures and over 246 billion substances which was supported by U.S. National Library of Medicine [55]. This database was used in the current in silico framework for virtual screening the molecules, which could act as HDAC1 inhibitors. For this purpose, similarity search was used to select all compounds containing benzamide sub-structure included within PubChem database. In particular, we used the similarity search tool implemented in PubChem database. This tool used the fingerprint Tanimoto-based 2-dimensional similarity search method. In this regard, the structures of four well-known benzamide-based HDAC1 inhibitor, Mocetinostat, Entinostat, Tacedinaline, and Chidamide were employed as templates to afford the most similar compounds with a Tanimoto threshold of 80% similarity considering the whole structure of compounds. This latter resulted in the generation of a library composed by 736,160 candidate hits. The structures of all achieved molecules were saved in Structure-Data File (SDF) format and subjected to the next screening filters.

2.3 Ligand preparation

Accurate 2D to 3D conversion and proper optimization of ligand structures is an important precursor step toward virtual screening studies. Thus, all the members of obtained library

were prepared by means of Macromodel [56] and LigPrep [57] modules of Schrodinger suite 2015, as previously described by us [58, 59]. In particular, molecular energy minimization of the structures was performed in MacroModel environment using the OPLS-AA 2005 as force field [60, 61]. The generalized born/surface area (GB/SA) solvation model was used for simulating the solvent effects with “no cut-off” for non-bonded interactions [62]. The Polak-Ribiere conjugate gradient (PRCG) method with 5,000 maximum iterations and 0.001 gradient convergence threshold was employed. All the compounds were then treated by the LigPrep application to accomplish chemical correctness along with generate the most probable ionization state at the cellular pH value (7.4 ± 0.5).

2.4. Pharmacophore-based 3D-QSAR dataset screening

The five-feature pharmacophore hypothesis (ADDRR) and its corresponding 3D-QSAR model, recently developed by us [53], was used as a 3D-query for screening against 736,160 hits derived through similarity search. This dataset screening was conducted using advanced pharmacophore-screening option of software package Phase 4.2, implemented in Schrödinger suite [63, 64]. Conformers of each molecule were generated in Phase employing the option “generated conformers during search.” The maximum number of conformers per structure and the number of conformers per rotatable bond was retained at 100 and 10, respectively. The Conformational searches were performed by applying the rapid torsional sampling method. The conformers were filter out by specifying a maximum energy of 10 kcal/mol relative to the lowest conformation so that the high-energy conformers discarded. The dataset was then searched to match the pharmacophore hypothesis ADDRR using the following criteria: a) at least four out of the five Pharmacophore features should be satisfied; b) no site should be considered as mandatory; c) do not prefer partial matches involving more sites; d) tolerance level for matching the inter-feature distances was set to the default value (3.0 Å). At the end of the search, in order to narrow down the number of potential hits, resulting compounds were filtered using the 3D-QSAR model to predict HDAC1 inhibitory activity. Hit compounds with estimated pIC50 values more than 0.7 were selected for further screening.

2.5. Protein preparation

rystal structure of human HDAC1 (PDB ID: 4BKX) was retrieved from the protein data bank [65], and subjected to the protein preparation wizard (PPW) protocol available in Schrödinger suite 2015 [66]. This protocol allowed us to obtain a reasonable starting structure of the target protein for structure-based screening experiments using a series of computational steps [67]. As previously reported [68, 69] ELM2-SANT domain of MTA1 used for crystallizing HDAC1 (PDB ID 4BKX) was manually removed and the HDAC1 enzymes was treated using PPW protocol. At first, the protein structure was preprocessed as follows: 1) removal of crystallographic water molecules and all irrelevant heteroatoms except for the Zn^{2+} ion within the active site; 2) adding all hydrogen atoms to the structure;

3) assigning bond orders; 4) generating metal binding state for the enzyme; 5) creating disulfide bonds; and 6) filling missing side chains and loops. Moreover, in order to optimize the hydrogen bond network, the protonation state of His, Asp, and Glu residues were predicted, 180° rotations of the terminal angle of Asn, Gln, and His residues were assigned, and hydroxyl and thiol hydrogens were sampled. Finally, OPLS-2005 force field was implemented with the convergence RMSD cut-off of 0.30 Å for energy minimization of resultant protein structure to ensure its stability and quality for the next studies [60].

2.6. Receptor grid generation

Receptor grid file for the prepared HDAC1 protein was generated using Receptor Grid Generation option of Glide module (Grid-based Ligand Docking with Energetic) implemented in Maestro suite [70]. Energy grids were prepared using the default value of protein atom scaling factor (1.0 Å) within a cubic box. Since the position of the catalytic Zn^{2+} ion within the binding site of crystallized HDAC1 (PDB ID: 4BKX) was known, we centered the grid on the zinc ion, roughly representing the center of active site (grid box center coordinate: X: -46.756; Y: 16.29; Z: -7.785). Moreover, the grid box was chosen to be sufficiently large to include not only the active site of the HDAC1 but also significant regions of the surrounding surface [68, 69]. The grid box was adjusted based on a size capable of accommodating ligands with a length of 16 Å. As a requirement for the grid generation procedure for metal-dependent enzymes, metal constraints for receptor grids were also considered to generate possible geometries for the metal-coordination bonds.

2.7. Structure-based virtual screening

The virtual screening workflow (VSW) module implemented in the Schrödinger suite was used to identify potential inhibitors, amongst the 28,252 top-ranked hits predicted as most active ones through the 3D-QSAR model, against the putative active site of prepared HDAC1 protein [71]. As part of this multi-step workflow, two pre-filtering choices, Lipinski's Rule of 5 (RO5) and Reactive filters, were set to remove ligands with undesirable drug-likeness properties and reactive functional groups. The survivors from the specified filtering criteria proceeded to the subsequent docking steps in the workflow. The VSW exploits Glide docking protocol to rank the best compounds, providing three different levels of docking precision: high-throughput virtual screening (HTVS), standard precision (SP), and extra precision (XP) [70, 72, 73]. The pre-generated grid file for the prepared receptor was used for docking calculations. In this study, docking calculations were first performed in HTVS mode. Subsequently, the retained ligands from this stage are passed to the next stage, which performs XP docking calculations. The scaling factor for the van der Waals radii of the docked ligands was set to 0.80 Å, with a partial charge cut-off 0.15 Å. In every step of glide docking, 10% of the ligands were retrieved with the higher docking score. The XP GlideScore scoring function was used to order the best ranked compounds and the specific interactions were assessed by using ligand-interaction diagram implemented in maestro and PyMOL [74]. Finally, the 105 top-ranked compounds, displaying GlideScore values lower than -7.0 kcal/mol, and a satisfactory binding mode consistent with the key

interactions observed for the active inhibitors within the HDAC1 active site, were retrieved for the next filtering stage.

2.8. ADMET properties prediction and PAINS filter analysis

Early determination of problematic candidates possessing poor pharmacokinetic profiling, can streamline the overall drug discovery process through a dramatic reduction in wasted time and money. Nowadays, in silico approaches are widely used to predict ADMET properties (absorption, distribution, metabolism, excretion and toxicity) of new chemical entities during lead identification and optimization [75]. Therefore, QikProp module implemented in the Maestro suite was used as an extra filter for in silico evaluation of ADMET properties and drug-like behavior of the potential hits selected from VSW [76]. In these calculations, Mocetinostat, Entinostat, Chidamide, and Tacedinaline were employed as standard drugs. This step was performed to select only hits that to be predicted to possess satisfactory physicochemical properties in the appropriate range recommended by QikProp. Especially, membrane permeability, lipophilicity, human oral absorption, cardiotoxicity or potential interaction with hERG channels were amongst important criteria investigated for filtering [77]. Default settings were employed for these calculations.

Subsequently, the set of selected compounds from Qikprop, was further inspected for behaving as “Pan Assay Interference Compounds” (PAINS) by using FAFDrugs 4.0 tool [78-80]. The PAINS compounds tend to nonspecifically react with a wide range of biological targets rather than specifically affecting one desired target, providing false positive compounds. Accordingly, these hits cannot be considered appropriate starting points for the drug discovery trajectory and should be removed from the screening process.

2.9. Induced Fit Docking Studies

In order to improve the reliability of the screening workflow, the induced-fit docking (IFD) protocol implemented in the Schrödinger suite was applied on the extracted drug-like compounds from the last filter [81]. This protocol allows the incorporation of receptor flexibility by adjusting the receptor structure in the binding site based on the docked ligand, offering more accurate method for predicting ligand binding mode [82, 83]. The protocol was conducted in three consecutive steps. During the first stage, ligands were docked into a rigid receptor using softened-potential Glide docking with van der Waals (vdW) radii scaling of 0.5 for both the protein and ligand non-polar atoms. The Glide SP mode was used for the initial docking, and maximum 20 ligand poses were retained for protein structural refinements [72]. The grid box was centered on the zinc ion within the active site and the box size was set to “Auto”. In the next step, the Prime module was employed to generate the induced-fit protein–ligand complexes. In this way, each of the 20 initial ligand poses from the previous step was subjected to the side-chain and backbone refinements. Amino acid residues having at least one atom located within 5 Å of each corresponding ligand pose were included in the conformational search and refinement while residues

outside this range were fixed [84, 85]. The refined complexes were ranked by Prime energy, and 10 receptor structures within 30 kcal/mol of the minimum energy structure were passed through for a subsequent round of Glide docking and scoring. Finally, a rigorous re-docking of the ligands was conducted into their respective low-energy receptor structure that produced in the second step using Glide XP mode at default settings. An IFD score that accounts for both the protein–ligand interaction energy and the total energy of the system was calculated employing OPLS-2005 force field and used to rank the IFD poses. Since, the more negative IFD score characterizes the more favorable binding, the potential hits with top IFD score were selected for the next screening steps. The final ligand–protein complexes were also visualized using PyMOL [74] and Ligand Interactions diagram embedded in the Maestro suite.

2.10. Ligand binding affinity estimation based on MM-GBSA technique

The approximate scoring functions like GlideScore are untrustworthy predictive criteria to rank compounds with respect to their binding affinities [86]. Thus, to obtain further accuracy for our protocol, the best ligand-HDAC1 complexes obtained from the IFD studies were subjected to the subsequent analysis with MM-GBSA process provided in the Prime module of Schrödinger suite 2015 [87]. This method offers an efficient and worthwhile post-scoring approach for prioritizing the screened hits with lower ΔGbind values through improvement in prediction of relative binding free energy between ligands and receptor. The MM-GBSA approach combines the molecular mechanical (MM) energies with the continuum solvent generalized Born (GB) model for polar solvation as well as a solvent-accessible surface area (SASA) for non-polar solvation term [88, 89]. Accordingly, the representative docked pose of ligands was first minimized using the local optimization feature in the Prime and then energies of ligand-protein complexes were calculated using the OPLS-2005 force field and Generalized-Born/Surface Area continuum solvent model. During this process, the ligand strain energy was also considered, as previously reported [42, 58]. The binding free energy of each ligand was determined by the following equations [90]:

$$\Delta G_{bind} = G_{complex} - (G_{protein} + G_{ligand})$$

$$= \Delta E_{MM} + \Delta G_{SOL}$$

$$\Delta G_{SOL} = \Delta G_{GB} + \Delta G_{SA}$$

Where ΔE_{MM} is the difference in energy minimization between the protein-ligand complex and the sum of the energies of free protein and ligand. This energy is calculated by summing contributions from the internal energies, electrostatic and van der Waals interaction energies. ΔG_{SOL} is the corresponding difference in the solvation free energies calculated by summing contributions from polar (ΔG_{GB}) and non-polar (ΔG_{SA}) solvation energies.

Eventually, the results were ranked based on the obtained ΔG_{bind} values. A satisfactory ΔG_{bind} (estimated value lower than −40 kcal/mol) was considered to repossess final hits with the highest binding affinity to HDAC1 active site.

2.11. Molecular dynamics (MD) simulation study

The best ranked candidates based on MM-GBSA results were subjected to MD simulation experiments for evaluating the stability of their binding mode within the HDAC1 active site and elucidation of the ligand-protein interactions in details [91-93]. These studies were performed by means of Desmond 4.1 academic version, using Maestro as graphical interface [54, 94, 95]. In order to obtain reasonable ligand-protein complexes as starting points for the MD simulation protocol, the complexes derived from MM-GBSA calculations, were initially prepared using protein preparation wizard workflow as described above. The prepared complexes were solved into a cubic-shaped box full of water molecules represented by known TIP3P model [96], using Desmond system builder. Force field parameters for the protein-ligand systems were assigned using the OPLS-2005 force field [60]. An appropriate number of Na+/Cl− counter ions were added to obtain a fixed salt concentration of 0.15 M representing the physiological concentration of monovalent ions. Before the MD simulation, a series of restrained minimization and short MD simulations were performed to slowly relax the model system without deviating considerably from the initial protein coordinates. The stages of pre-relaxation process are the following: 1) 12 ps simulation in the NVT ensemble (constant number of particles, volume and temperature) restrained with non-hydrogen solute atoms (temperature 10 K); 2) 12 ps simulation in the NPT ensemble (constant number of particles, pressure and temperature) restrained with non-hydrogen solute atoms (temperature 10 K); 3) 12 ps simulation in the NPT ensemble (temperature 300 K) restrained with solute non-hydrogen atoms and 4) 24 ps simulation in the NPT ensemble (temperature 300 K) with no restraints. The temperatures and pressures in the short initial simulations were controlled using Berendsen thermostat and barostat, respectively. Finally, the equilibrated system was simulated for 30 ns at the constant temperature of 300 K and pressure of 1.01325 bar, employing the NPT as ensemble class. RESPA integrator was applied in order to integrate the equations of motion, with an inner time step of 2.0 fs for bonded interactions and non-bonded interactions within the short-range cut-off [97]. Nose-Hoover thermostats [98] were used to keep constant the simulation temperature, and the Martyna-Tobias-Klein method [99] was used to control the pressure. Long-range electrostatic interactions were calculated by particle-mesh Ewald method (PME) [100]. The cut-off for van der Waals and short-range electrostatic interactions was set at 9.0 Å. Consequently, a single trajectory of 30 ns for each complex was obtained. The trajectory files were analyzed using simulation event analysis, simulation quality analysis, and simulation interaction diagram tools provided in the Desmond package. Moreover, the mentioned tools were employed to generate all plots regarding MD simulation presented in this study.

3. Results and discussion

This study aims to develop an attentively designed a virtual screening workflow that can efficiently be used to identify novel and promising hits as selective HDAC1 inhibitors. Recently, our group successfully developed a predictive 3D-QSAR model derived from

pharmacophore-based alignment method using a large set of benzamide derivatives as noteworthy selective HDAC1 inhibitors [53]. Encouraged by the results obtained from the validation of model, to exploit the available data concerning both ligands and HDAC1 structure, the virtual screening strategy was centered on the parallel application of both 3D-QSAR model and conceptually different structure-based methods, in a stepwise-filtering approach (Figure 2). Below, the results of the different in silico experiments used in the current screening are presented and discussed.

3.1. Dataset and Library preparation

Given that structural similarities of ligands can indicate similarities in their pharmacological features, a similarity search method (ligand-based) was used to efficiently screen PubChem database to cast an initial library for our screening process. This approach leads to the selection of a narrower chemical space of the database, and then increases the chance of success. For this purpose and according to a Tanimoto similarity metric of 80% with Mocetinostat, Entinostat, Tacedinaline, and Chidamide, a starting virtual library of 736,160 structurally similar compounds (227,638 Mocetinostat-like, 24,919 Entinostat-like, 239,276 Tacedinaline-like, and 244,327 Chidamide-like molecules) was identified from the PubChem database. These hits were prepared at the cellular pH value (7.4 ± 0.5) using LigPrep considering possible ionization states and guaranteed that all ligands are in the lowest energy conformation. Prepared ligands were subjected to the following hierarchical filtering steps.

3.2. 3D-QSAR Pharmacophore-based virtual screening

Pharmacophore hypothesis alignment is one of the most widely used ligand-based approaches for screening existing databases to prioritize potent compounds for experimental evaluation. The main advantage of this method is the high screening speed and rapid elimination of compounds lacking fundamental structural features. Thus, the five-point pharmacophore hypothesis (ADDRR), supplemented with a highly predictive 3D-QSAR model, previously developed by us for selective HDAC1 inhibitors [53], was used as a tool for discriminating potential inhibitors from virtual library of 736,160 hits obtained in the previous step. Various conformers of these hits were thoroughly searched for matching at least four out of the five features. Pharmacophore searching led to the identification of 473,919 hits that met the specified requirements in the pharmacophore model. Lastly, further filtration of matches based on estimated pIC50 values > 0.7 in turn allowed the selection of 28,252 top-ranked compounds with the highest predicted inhibitory activity for the next structure-based studies.

3.3. Structure-based virtual screening

Ligand-based approaches suffer from high false positives, i.e., hits that match the pharmacophore hypothesis but do not show activity, due to unsuitable fitting in the target binding site. In contrast, structure-based methods consider the receptor structure capable of predicting the accurate binding mode of ligands within the active site of the target protein and, thereby preferentially ranking active ligands ahead of inactive ones. Thus, to reduce the risk of false positives, the identified 28,252 potential hits were employed to execute structure-based virtual screening against the putative active site of HDAC1, using the VSW procedure available in Maestro molecular modeling environment. This workflow allows to perform a fast preliminary screening (HTVS), followed by a more accurate docking protocol (SP or XP). The molecular docking protocol necessitates prudently optimized 3D-structure of protein for accurate binding affinity prediction. Thus, the X-ray structure of HDAC1 (PDB ID: 4BKX) was optimized in the protein preparation wizard environment, applying OPLS-2005 as force field to remove bad atomic contacts in the 3D-structure and to obtain a structure with a lower energy state. Using “Lipinski’s rule of five” (RO5) and "Reactive" pre-filters at the dawn of the screening workflow, the dataset was trimmed to exclude less pharmacokinetically suitable ligands. In fact, Lipinski’s RO5 is commonly used as criteria for evaluation of drug-like behavior of hits, which is necessary for rational drug design campaigns. This rule describes four ranges for important physicochemical properties of small-molecule drugs, including 1) a partition coefficient (log P) ≤ 5; 2) a molecular weight (MW) ≤ 500 g/mol; 3) a number of hydrogen bond acceptors ≤ 10; 4) a number of hydrogen bond donors ≤ 5. Any value differing from these values was considered as a violation. The acceptable value of RO5 violations for a drug-like molecule is 1 [77]. According to results, 24,025 out of 28,252 hits were identified to satisfy Lipinski's properties and thereby signified to possess appropriate drug-like potential. Moreover, after applying Reactive filtering, 23,030 hits lacked any reactive functional groups that were survived for proceeding to the next steps. Afterwards, the two modes of structure-based virtual screening were then applied for the docking of the retrieved 23,030 drug-like hits targeting the key residues of HDAC1 active site and subsequent scoring of their binding affinities. In the first stage, the top 10% hits in terms of docking score comprising 2,303 molecules, were retained employing Glide HTVS method. Further re-docking of these compounds using XP Glide mode leads to 230 hits. Since XP docking method is highly accurate and penalizes severe steric clashes, the final hit selection was conducted based on XP GlideScore threshold lower or equal to -7.0 kcal/mol. This filter rendered 105 top-ranked hits as potential HDAC1 inhibitors to continue further examination.

3.4. ADMET prediction and PAINS filter analysis

The intrinsic activity of a drug is not the only factor that determines its effectiveness upon administration. A drug should be capable of permeating a wide variety of barriers before reaching its site of action and should undertake so without being metabolized. Therefore, evaluating the absorption, distribution, metabolism, excretion, and toxicity (ADMET) profile of drug candidates is of utmost importance in the early stages of drug discovery, thereby

improving the probability of clinical success later during the drug development process. Nowadays, in silico techniques are widely used for the fast prediction of safety, pharmacokinetic, bioavailability and toxicity properties of new chemical entities in human body preceding expensive experimental procedures [101]. Consequently, as a step of the developed screening workflow, QikProp software was used to assess many crucial pharmacokinetic parameters of 105 hits that passed the last step. This step enabled us to eliminate weak druggable candidates and further focus on the molecules with favorable pharmacokinetic properties. The special parameters considered for ligand selection in this step of filtering were as follows: Caco-2 and MDCK cell permeability in nm/sec (QPPCaco-2 and QPPMDCK >500), blood-brain barrier permeability (QPlogBB: -3 to 1.2), the percentage of human oral absorption (HOA > 80%), lipophilicity (QPlogP: -2 to 6.5), aqueous solubility (QPlogP: -6.5 to 0.5), and predicted IC50 values for blockade of hERG K+ Channels and so cardiotoxicity (QPlogHERG < -5). Data indicated that predicted pharmacokinetic properties of 77 out of 105 compounds concurred with permissible ranges described for the human being. The results for all 77 retrieved molecules were also comparable with the corresponding descriptors for standard drugs, qualifying them as drug-like molecules. Furthermore, the resultant compounds were fruitfully screened for their potential capability to behave as PAINS by FAFDrugs 4.0 web server. Based on this filtration criterion, 3 molecules among 77 compounds have sub-structural features that marked them as “frequent hitters” in high-throughput screenings. Eventually, this step reduced the number of screened hits, thereby highly enriching the library with 74 more promising virtual hits. The ADMET prediction results of these top-screened candidates were provided in the Supplementary Table S1.

3.5. Induced-fit docking (IFD)

In order to narrow down the number of the potential hits and guarantee the reliability of the screening protocol, 74 retrieved drug-like molecules were further analyzed using IFD calculations, a flexible docking technique, as previously described by us [102-104]. Considering a rigid receptor in standard virtual docking studies can lead to misleading results, since many proteins undergo side-chain or backbone movements, upon ligand binding. The most important feature of IFD is that both ligand and residues in the receptor’s active site and its vicinity are imparted flexibility. Thus, this docking protocol can significantly improve the prediction of binding mode of hit candidates into the HDAC1 active site. The best pose of the compounds obtained from initial virtual docking experiments was exploited for IFD studies targeting the key residues involved in HDAC1 active binding site. To validate the applied docking protocol, the binding modes of Mocetinostat, Entinostat, Tacedinaline, and Chidamide, four well-known HDAC1 inhibitors, were also explored via IFD procedure. The obtained IFD results for all compounds were sorted based on XP GlideScore values. With the purpose of screening compounds with docking score lower than those found for the reference compounds, the score values lower than -11.00 kcal/mol were considered as filtering criteria. This led to identify 37 top-ranked hit candidates as prominent HDAC1 inhibitors with docking score within the range from -13.91 to -11.02 kcal/mol, Glide energy from -69.06 to -44.56 kcal/mol, glide Emodel energy from -123.49 to -61.91 kcal/mol and IFD score from -1217.85 to -1209.36 kcal/mol

(Table S2). To confirm that, the best binding poses of each selected ligand were engaged in critical interactions within the HDAC1 active site, all predicted IFD poses were visually scrutinized. The different scores and energies obtained from IFD and the interaction details for 37 selected hits are presented in Tables S2 and S3, respectively.

The active site of zinc-dependent HDAC1 consisted of a long and narrow hydrophobic tube-like channel, leading to an internal cavity that contains the catalytic Zn^{2+} ion. Various hydrophobic residues such as phenylalanine, histidine, glycine, methionine, leucine and tyrosine, play significant roles in the creation of this active site. Therefore, these crucial residues along with the catalytic zinc ion were proposed to be involved in the potential binding interaction between inhibitors and HDAC1 active site [68, 69, 105-107]. The analysis of the IFD results indicated that best binding pose of 37 selected hits was properly settled in the same space as occupied by the approved HDAC1 inhibitors. In consonance with docking models of reference ligands, the predicted binding mode of these hits was predominantly driven by the extensive hydrophobic interactions with key residues Met_{30}, Leu_{139}, Phe_{150}, Cys_{151}, Phe_{205}, Leu_{271}, and Tyr_{303} (Table S3). Besides the potential metal-coordination bond with the catalytic Zn^{2+} ion, several H-bond interactions were detected with residues Asp_{99}, His_{140}, His_{141}, Gly_{149}, His_{178}, Tyr_{204}, Phe_{205}, Leu_{271}, and Tyr_{303}, that contribute to further stability of molecules in the binding site. The latter interactions were strengthened by the formation of π–π stacking contacts with residues His_{140}, His_{141}, His_{178}, Phe_{150}, Tyr_{204}, Phe_{205}, and Tyr_{303}. Moreover, the amino acids Arg_{34}, Phe_{150}, His_{178}, and Phe_{205} were occasionally observed to be engaged in additional cation-π interaction with some of the selected hits (Table S3).

3.6. MM-GB/SA rescoring of hit compounds and binding mode analysis

Although docking methods and the associated scoring functions exhibit good predictive power in offering the best ligand pose within the protein-binding site, they are not reliable enough to rank order compounds with respect to their binding affinities and thereby biological activity. This poor correlation can be due to severe approximations employed by docking scoring functions, which can substantially amplify inaccuracies in such calculations [86, 108]. It has recently been appeared that the incorporation of more physically relevant energy terms such as solvation energy and surface accessibility area with a molecular mechanical force field provides ligand binding energy calculations with a more acceptable accuracy [88, 89]. Thus, the best-docked pose of each ligand selected from the previous IFD studies was rescored using a subsequent MM-GBSA post-docking protocol. MM-GBSA allow us to perform relatively accurate predictions of ligand binding affinity for a specific target and so can be applied with confidence to prioritize active hits in a virtual library. Rescoring using MM-GBSA leads to minor changes in ligand conformations within the receptor site. These changes result from the minimization of the ligand in the receptor's environment and consequent stabilization of receptor-ligand complex [88, 89].

In this step of screening workflow, ranking of the ligands was conducted based on obtained free binding energy values (ΔG_{bind}). The ΔG_{bind} values lower than -40 kcal/mol were considered to retrieve the final set of compounds, leading to the identification of 12

top-ranked hits. This implies that these compounds were the most stable ligands within protein-binding site, thereby possessing the highest in silico binding affinity for HDAC1 active site. The chemical structures of these compounds are shown in Figure 3. The calculated ΔG_{bind} values of the selected hits along with their contributions to the total binding free energy from various energy components are provided in Table 1.

About the free binding energy values of known inhibitors reported in Table 1 (Mocetinostat, Entinostat, Chidamide, and Tacedinaline), the selected hits were predicted to have a higher binding affinity than known ligands toward HDAC1 active site (ΔG_{bind} ranged from -54.095 to -40.199 kcal/mol). Specifically, most of these ligands exerted more favorable van der Waals and lipophilic terms than known inhibitors. In addition, general inspection of the free energy components in this Table revealed that the van der Waals and the lipophilic interaction energies (ΔG_{vdW} and ΔG_{Lipo}) are the most important contributors to the ligands binding energy. This observation emphasizes critical importance of hydrophobic interactions in the stability of the igand–protein complexes, which is logical considering the highly hydrophobic nature of HDAC1 binding pocket. A close-up view of the best-docked pose of the six top-ranked hits inside the HDAC1 active site along with their 2D interaction diagram is depicted in Figure 4. The binding mode analyses of these compounds were also described in detail below. The details regarding atoms/residues of top six potential hits involved in key interactions within HDAC1 binding site as well as the predicted binding mode of the other final hits selected from this step are also provided in the supplementary Tables S4, S5 and Figure S1.

3.6.1. Binding mode of compound CID_38265326

The hit CID_38265326 was identified to occupy the HDAC1 binding site with the most favorable binding energy (ΔG_{bind}: -54.095 kcal/mol). Interestingly, this result was in good accordance with data predicted by IFD modeling studies. As shown in Table S2, the best XP Glide and IFD scores (-13.913 and -1217.85 kcal/mol, respectively) were also obtained for this compound amongst screened hits. This ligand passes the tube-like gorge of the HDAC1 enzyme while interacting by hydrophobic contacts with the key residues Leu_{271}, Phe_{205}, and Phe_{150}, as well as two hydrogen bonds with Asp99 (see Tables S4, S5 and Figure 4). In this way, the molecule was extended into the internal cavity and favorably participated in expected hydrophobic interactions with Val_{19}, Tyr_{24}, Met_{30}, Pro_{32}, Ile_{35}, Phe_{109}, Trp_{135}, Ala_{136}, Leu_{139}, Cys_{151}, and Tyr_{303}. This situation allowed the carbonyl oxygen of the benzamide group to get involved in a direct interaction with the Zn^{2+} 2.12 Å away, while the N-H of amide can form a hydrogen bond with Gly_{149} (Table S4). Moreover, phenyl and pyridine rings of the ligand were observed to engage residues phe_{150}, His_{140}, His_{141}, and Arg_{34} through π-π stacking and cation-π interactions, thereby improving further binding affinity of the ligand to HDAC1 active site.

3.6.2. Binding mode of compound CID_56064109

As shown in the MM-GBSA results in Table 1, the hit CID_56064109 displayed the binding affinity for the HDAC1 active site with the ΔG_{bind} value of -49.947 kcal/mol. The putative binding mode of this molecule within HDAC1 binding site (with an IFD score of -1216.121 kcal/mol and GlideScore of -12.107 kcal/mol) indicated that the ligand binding is primarily mediated by hydrophobic interactions (see Tables S2 and S3). In this compound, 2-(benzyloxy)pyridine moiety address the active site tunnel, providing appropriate hydrophobic interactions with residues Leu_{271}, Tyr_{204}, and Pro_{206} (see Table S5 and Figure 4). The pyridine portion was also involved in an additional π–π stacking interaction with Tyr_{204}. The rest of the molecule embedded in the hydrophobic internal cavity and stabilized by favorable hydrophobic contacts with Pro_{29}, Tyr_{303}, Phe_{205}, Met_{30}, Leu_{139}, and Cys_{151} (Table S5) as well as π– π stacking interactions with His_{178} and Phe_{205}. The adopted conformation of the ligand within HDAC1 active site, benefits from the establishment of additional hydrogen bonds with Gly_{149}, His_{178} (Table S4), in addition to the zinc ion coordination.

3.6.3. Binding mode of compound CID_8136932

The binding energy of -49.046 kcal/mol was predicted for the complex formation of this hit with HDAC1 binding site (Table 1). As shown in Figure 4, phenyl and cyclopropyl groups of this ligand were found to fit inside the lipophilic tube and main cavity of HDAC1 active site, so that they were able to establish intensive hydrophobic interactions with the neighboring residues Tyr_{24}, Met_{30}, Phe_{109}, Leu_{139}, Phe_{150}, Cys_{151}, Tyr_{204}, Phe_{205}, Leu_{271}, and Tyr_{303} (See Table S5). This accommodation enables phenyl rings of ligand to engage in π–π stacking interactions with residues Tyr_{204}, His_{178}, and Phe_{150}, while the nitro functional group participated in two cation-π interactions with Phe_{150} and Phe_{205}. In addition to interaction of the carbonyl oxygen of benzamide moiety with the catalytic zinc ion at the end of the pocket, this moiety was involved in additional hydrogen bond formation with residues Tyr_{303} and Gly_{149} (Table S4).

3.6.4. Binding mode of compound CID_55802151

The best-bound conformation of this ligand within HDAC1 active site exhibited a ΔG_{bind} value of -48.745 kcal/mol (Table 1). The predicted binding mode for this hit (Figure 4) showed that HDAC1 active site properly encompassed aromatic groups of the ligand, forming the expected hydrophobic interactions (Table S5). Moreover, the pyridine moiety binding to the active site tunnel was engaged in a π-π stacking with the phenyl ring of Phe_{205} residue. The phenyl ring of benzamide moiety buried in the internal cavity, was also able to establish a triple π-π stacking with Phe_{150}, His_{141}, and His_{178}. This orientation enables the terminal amide side-chain of ligand to embed in a region close to the Zn-binding site of protein. Consequently, beyond the zinc ion coordination, carbonyl oxygen

and N-H of this benzamide group established two hydrogen bonds with residues His_{140} and Gly_{149}, respectively (Table S4). The carbonyl group of amide linker between pyridine and phenyl rings also form another hydrogen bond with His_{178}, which contribute to further stabilization of ligand binding inside HDAC1 active site (Table S4).

3.6.5. Binding mode of compound CID_133901641

The proposed binding mode of ligand CID_133901641 accounted for a ΔGbind value of -47.975 kcal/mol, filling the tunnel and main cavity of HDAC1 active site. In this context, terminal ether side-chain of ligand overlaid the mouth of the active site tunnel, such that stabilized by favorable hydrophobic contacts with Tyr_{204}, Leu_{271}, and Phe_{205}, as well as hydrogen bond interaction with His_{178} (see Tables S4, S5 and Figure 4). The rest of the ligand perfectly nestles in the tunnel and internal cavity of the active site, making favorable hydrophobic contacts with Met_{30}, Phe_{150}, Tyr_{303}, Cys_{151}, and Leu_{139} (Table S5). With this accommodation, carbonyl oxygen of amide linker between phenyl rings is able to interact with the catalytic zinc ion at the end of the pocket. In addition, residues His_{140}, His_{141}, and His_{178} along with Gly_{149} were observed to participate in π - π stacking and hydrogen bond interactions with ligand, respectively.

3.6.6. Binding mode of compound CID_18150975

This ligand was predicted to orient inside the HDAC1 binding site (ΔGbind: -47.873 kcal/mol) in such a way that the benzimidazole, phenyl and cyclopropyl moieties were properly clamped by lipophilic tube and main cavity of HDAC1 active site, providing expected hydrophobic interactions with the key residues Met_{30}, Phe_{109}, Phe_{150}, Leu_{139}, Cys_{151}, Phe_{205}, Leu_{271}, Tyr_{303}, Trp_{312} (Table S5). In this respect, phenyl ring favorably took part in additional π-π stacking interactions with residues Phe_{150} and His_{141} as well. In addition to the observed zinc ion coordination, the carbonyl oxygen and N-H of benzamide groups of the ligand exerted interactions with the residues Tyr_{303}, Gly_{149} and Gly_{138} through hydrogen bonding (Table S4).

3.7. Molecular dynamics simulation analysis

Molecular dynamics (MD) simulation has evolved into a fundamental in silico technique for capturing dynamic events of biological systems under specific conditions of physiological environment. A high-performance MD, especially when coupled with other computational tools, can provide detailed insight into conformational changes and internal interactions of protein-ligand complexes on the time scales by introducing atomic-level perturbations [109]. Thus, the best binary complexes of the 12 final screened hits and known ligands with HDAC1, in terms of the lowest binding free energy and the best orientation of the ligand in the active site, were further subjected to an extensive MD study for 30 ns. The

objective of MD simulations was to ensure the stability of the binding mode and the main intermolecular interactions of the ligands in the HDAC1 active site. In this regard, the resulting trajectories for all complexes were completely scrutinized through different standard simulation parameters including the root mean square deviations (RMSDs) for all backbone atoms and ligands, the root mean square fluctuations (RMSFs) of individual amino acid residues, intermolecular hydrogen bond formation and gyration radius.

Overall structural fluctuations and conformational stability of each complex were evaluated by analyzing the RMSD of protein backbone atoms versus simulation time. The results of the RMSD analysis for the selected hits and known inhibitors are reported in Figures S2 and S3, respectively. As seen in the Figure S2, after an initial period of RMSD fluctuations, all ligand-protein complexes, except for **CID_17479772** and **CID_29693448**, reached equilibrium status and then the conformational stasis was accomplished without much fluctuation in the rest of the simulation time. These results implied that these systems folded into a more stable state than the starting structure, consequently reinforcing the reliability of the docking outcomes.

By comparing the RMSD plots, it can be found that **CID_38265326**, **CID_56064109**, **CID_8136932**, **CID_55802151**, **CID_133901641**, and **CID_18150975** (indicated by asterisk in Figure 3) had the least fluctuations in RMSD amongst hits. Accordingly, it was concluded that the complex of these hits with HDAC1 active site are more stable than the other ligands under the same MD simulation conditions. As reported in Figure 5, the RMSD profiles of these six complexes were almost the same at the end of the simulation with average RMSD values of 1.89, 2.02, 2.03, 2.12, 1.95, and 1.93 Å, respectively. In particular, they were perfectly superimposed in the last 10 ns of the simulation time. For **CID_56041109**, **CID_8136932**, and **CID_55802151**, the system gradually reached an overall stability after about 15 ns, whereas the bound state of compounds **CID_38265326**, **CID_133901641**, and **CID_18150975** displayed a longer equilibration time (after 10 ns of MD simulation). Moreover, the RMSD values of these compounds were comparable with those found for known inhibitors, (Entinostat (RMSD: 1.71 Å), Tacedinaline (RMSD: 2.13 Å), Chidamide (RMSD: 1.84 Å), and Mocetinostat (RMSD: 1.82 Å)), fortified their stable binding with HDAC1 active site under similar simulation conditions.

RMSF refers to the variation of the atomic Cα coordinates of the protein from its average position throughout the MD simulation. This assessment is particularly useful for characterizing the flexibility of individual residues in the protein backbone. Thus, RMSF of backbone atoms for six aforementioned ligand-protein complexes with the best RMSD profile, was analyzed to explore the dynamic behavior of the essential residues involved in the interaction with a specified ligand. Figure 6 illustrates the superimposed RMSF graph of protein-ligand complexes. Although some residues in several complexes fluctuated much more than in other complexes, overall pattern of residue fluctuations in complexes except for CID_38265326 was found to be similar to that seen in Entinostat. The results of RMSF analysis indicated that the main fluctuations in all systems corresponded to residues that were far from the ligand binding site. As shown in Figure 6, in all systems a noticeable value of RMSF was observed in a restricted number of residues at the N-terminal and the C-terminal tails of protein. Furthermore, residues such as Glu_{86}, Ala_{223}, Ala_{239}, Ser_{290}, and Ala_{317}, with high fluctuation were located in the flexible loop regions. In contrast, the conformational changes of crucial residues in the HDAC1 active site, (lowest RMSF values for all complexes), verifying the capability of ligands to form stable

interactions within the protein. In particular, in **CID_38265326**, the key residues Leu_{139}, His_{140}, His_{141}, and Phe_{150} represented maximum RMSFs of 1.31, 1.40, 1.48, and 1.86 Å, respectively. The RMSF fluctuations of these residues in five other hits and Entinostat were observed to be higher than **CID_38265326**. In agreement with results of MM-GBSA analysis, these findings perfectly corroborated the favourable binding affinity of **CID_38265326** for the protein-binding site compared with other hits and so supported the reliability of MD simulation.

The radius of gyration (Rg) is an indicator of how the structural compactness of protein changes during a simulation, which in turn reflects system stability. Time-dependency plots of the radius of gyration for the simulated complexes of six hits and known inhibitors are shown in Figures S4 and S5, respectively. As shown in Figure S4, the Rg values of HDAC1 backbone atoms in the presence of **CID_38265326**, **CID_8136932**, **CID-55802151**, **CID_133901641**, and **CID_18150975**, were slightly reduced during the simulation time, implying a more compact protein structure. In contrast, insignificant expansion of the protein structure has been observed with **CID_56064109**. As a result, the target protein showed reasonably stable structure lacking any major expansion/contraction, after the binding of these ligands throughout the simulation period. Moreover, the comparative results revealed that hits were found as lesser fluctuated and thereby with more stable protein structure than known inhibitors (Figure S5).

The MD trajectories were also examined for an in-depth study of the underlying forces in relation to protein-ligand structure stability during simulation time. The pattern of molecular interaction of six best-selected ligands over the binding site revealed that main contacts were essentially preserved as underlined by the dynamic analysis of the simulation interaction diagrams (Figures 7, 8 and S7). In this regard, **CID_38265326** and **CID_8136932** displayed the highest number of contacts with the protein over the course of the trajectory, (0-20 and 0-18 contacts, respectively). These findings clearly support the lowest amount of RMSF being observed in these two hits compared with other ligands (Figure 6), consolidating their thermodynamic stability. We found a total number of 0-16 and 0-15 contacts for ligands **CID_56041109** and **CID_18150975** respectively, whereas the lowest number of contacts was found for ligands **CID_55802151** and **CID_133901641** (for both about 0-12 contacts) (Figure 7).

As depicted in Figure 8, the carbonyl oxygen of the benzamide moiety of all hits, tightly interacted with the catalytic zinc center, during the entire simulation time, suggesting the relevance of this interaction for substantial binding of ligands with HDAC1 active site. Moreover, hydrogen bonds were found as another important interaction involved in the stability of selected HDAC1-ligand complexes. In fact, protein rigidity is denoted by increased intra-molecular hydrogen bonds. Hydrogen bond monitoring revealed that same interactions observed in docking models of most hits were reproduced after MD simulations. For example, in line with the docking studies, the Gly_{149} and Asp_{99} were among hydrogen bond interacting residues with hit **CID_38265326**. In particular, the hydrogen bond formed by Gly_{149} was also observed in the simulation process of five other ligands and known inhibitors bound to HDAC1 active site (Figure 7). Notably, this residue is one of the critical residues within the ligand binding pocket of HDAC1. The results also revealed that **CID_38265326** formed additional hydrogen bonds with residues Arg_{34}, Arg_{270}, and Tyr_{303} during 37%, 20%, and 7% of the simulation time, respectively. In

addition, the crucial residue Tyr_{303} can occasionally be targeted by two other compounds **CID_8136932** and **CID_55802151** through hydrogen bonding.

The 30 ns MD simulations were also deciphered some additional polar contacts as water mediated H-bonds, which were beneficial to the stability of the molecules in the HDAC1 active site. As illustrated in Figure 8, residues Arg_{34}, Asp_{99}, Arg_{270}, Leu_{271}, Phe_{205}, and Tyr_{303} were found to be particularly relevant in forming water-bridging H-bonds with selected hits. Moreover, residue Arg_{34} was involved in further electrostatic interaction with compound **CID_38265326** (Figure 7).

As indicated in Figure 7, the hydrophobic interactions are the main contacts that manage the binding mode of six nominated ligands toward HDAC1 active site. Given the analyzed data, the selected hits depicted hydrophobic interactions such as those found in docking calculations. In this regard, the key lipophilic residues in HDAC1 binding site, Met_{30}, His_{178}, Tyr_{303}, Phe_{150}, Phe_{205}, and Leu_{271} were particularly targeted for establishment of various hydrophobic contacts with hits during the simulation. Moreover, further stabilization of most of these ligands in the HDAC1 active site was facilitated, although occasionally, by π- π stacking with His_{178}, His_{141}, Phe_{150}, Tyr_{303}, and Tyr_{204} as well as cation-π stacking interactions with Arg_{34}, Arg_{270}, Phe_{205}, Phe_{150}, and His_{178}.

Overall, the MD simulation results clearly confirmed the advantageous interactions of six screened compounds with reasonable thermodynamic stability in the HDAC1 active site, further consolidating their capability as plausible HDAC1 inhibitors.

4. Conclusion

In this work, it was attempted to develop an exhaustive virtual screening protocol involving a combination of conceptually different in silico methods (structure and ligand-based) for the efficient identification of novel benzamide-based analogs as HDAC1 inhibitors. To achieve this goal, computational evaluation of a library of 736,160 compounds, attained by similarity search of four renowned HDAC1 inhibitors, was conducted in several hierarchical steps. First, the five-point pharmacophore hypothesis (ADDRR), supplemented with a highly predictive 3D-QSAR model, recently developed by us [53], was used as a preliminary filtering tool, for selecting portions of the starting library that were compliant with chemical and geometrical aspects responsible for HDAC1 inhibitory activity. Subsequently, 28,252 identified hits with the highest predicted inhibitory activity (pIC50 > 0.7) were subjected to various filtering processes to elicit an enriched set of promising candidates as HDAC1 inhibitors. The exploited filtering criteria in the order of steps included 1) the estimated XP GlideScore values lower than −7.0 kcal/mol obtained from the structure-based virtual screening workflow; 2) having adequate pharmacokinetic and ADMET properties within the acceptable range described for a drug-like molecule as predicted by using QikProp software and FAFDrugs webserver; 3) the top IFD score and the lowest GlidScore value along with considering the best interaction pattern with the key residues in the HDAC1 active site obtained from flexible docking studies (IFD calculations); 4) the best estimated binding free energy values calculated using Prime/MM-GBSA simulation.

The discussed filtering process afforded a final selection of 12 top-ranked drug-like hits. These compounds fulfilled necessary ADMET properties calculated for human purpose and passed the false positive evaluation during PAINS analysis. Furthermore, the results of IFD docking simulations coupled with MM-GBSA rescoring method were indicative of higher binding affinities of selected hits to HDAC1 active site compared with currently known active inhibitors. As a final step, the best-docked pose of the top-screened hits in complex with the HDAC1 active site was employed for a series of 30 ns simulations to disclose the binding strength, the overall stability of the interaction profiles and novel interactions not observed during docking studies. Regarding the comparative analysis of RMSD, RMSF, Rg parameters, and pattern of intermolecular interactions during MD simulation time, among the final hits, six compounds **CID_38265326**, **CID_56064109**, **CID_8136932**, **CID_55802151**, **CID_133901641** and **CID_18150975** presented the best stability profiles and binding mode in the HDAC1 active site. The IFD and MD results cooperatively exposed the importance of hydrophobic contacts and optimum hydrogen bonds with crucial residues of HDAC1 binding pocket, such as Gly_{149}, Phe_{150}, His_{178}, Tyr_{204}, Phe_{205}, Arg_{270}, Leu_{271}, and Tyr_{303}. Consequently, top-six presented hits can be considered as encouraging templates for the further development of new benzamide chemotypes as HDAC1 inhibitors. We expect that the findings obtained here, provide an outline for future experimental explorations to stimulate successful design of improved analogs.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that can influence the work reported in this paper.

Acknowledgments

The authors thank the Bioinformatics Research Center in Isfahan University of Medical Sciences (Iran) for financing and supporting this project (Grant No. 298025 to H.S.), and the Università di Pisa under the “PRA – Progetti di Ricerca di Ateneo” (Institutional Research Grants) (Project no PRA_2020_58 “Agenti innovativi e nanosistemi per target molecolari nell'ambito dell'oncologia di precisione” to S.B.).

References:

[1] S.J. Conway, P.M. Woster, W.J. Greenlee, G. Georg, S. Wang, Epigenetics: Novel Therapeutics Targeting Epigenetics, J Med Chem, 59 (2016) 1247-1248.

[2] K.J. Falkenberg, R.W. Johnstone, Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders, Nat Rev Drug Discov, 13 (2014) 673-691.

[3] M. Paris, M. Porcelloni, M. Binaschi, D. Fattori, Histone deacetylase inhibitors: from bench to clinic, J Med Chem, 51 (2008) 1505-1529.

[4] A.J. de Ruijter, A.H. van Gennip, H.N. Caron, S. Kemp, A.B. van Kuilenburg, Histone deacetylases (HDACs): characterization of the classical HDAC family, Biochem J, 370 (2003) 737-749.

[5] T. Kouzarides, Chromatin modifications and their function, Cell, 128 (2007) 693-705.

[6] L. Ellis, P.W. Atadja, R.W. Johnstone, Epigenetics in cancer: targeting chromatin modifications, Mol Cancer Ther, 8 (2009) 1409-1420.

[7] I.V. Gregoretti, Y.M. Lee, H.V. Goodson, Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis, J Mol Biol, 338 (2004) 17-31.

[8] T. Finkel, C.X. Deng, R. Mostoslavsky, Recent progress in the biology and physiology of sirtuins, Nature, 460 (2009) 587-591.

[9] W. Weichert, HDAC expression and clinical prognosis in human malignancies, Cancer Lett, 280 (2009) 168-176.

[10] J.E. Bolden, M.J. Peart, R.W. Johnstone, Anticancer activities of histone deacetylase inhibitors, Nat Rev Drug Discov, 5 (2006) 769-784.

[11] S. Minucci, P.G. Pelicci, Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer, Nat Rev Cancer, 6 (2006) 38-51.

[12] M. Duvic, R. Talpur, X. Ni, C. Zhang, P. Hazarika, C. Kelly, J.H. Chiao, J.F. Reilly, J.L. Ricker, V.M. Richon, S.R. Frankel, Phase 2 trial of oral vorinostat (suberoylanilide hydroxamic acid, SAHA) for refractory cutaneous T-cell lymphoma (CTCL), Blood, 109 (2007) 31-39.

[13] H.Z. Lee, V.E. Kwitkowski, P.L. Del Valle, M.S. Ricci, H. Saber, B.A. Habtemariam, J. Bullock, E. Bloomquist, Y. Li Shen, X.H. Chen, J. Brown, N. Mehrotra, S. Dorff, R. Charlab, R.C. Kane, E. Kaminskas, R. Justice, A.T. Farrell, R. Pazdur, FDA Approval: Belinostat for the Treatment of Patients with Relapsed or Refractory Peripheral T-cell Lymphoma, Clin Cancer Res, 21 (2015) 2666-2670.

[14] K.P. Garnock-Jones, Panobinostat: first global approval, Drugs, 75 (2015) 695-704.

[15] Z. Qiao, S. Ren, W. Li, X. Wang, M. He, Y. Guo, L. Sun, Y. He, Y. Ge, Q. Yu, Chidamide, a novel histone deacetylase inhibitor, synergistically enhances gemcitabine cytotoxicity in pancreatic cancer cells, Biochem Biophys Res Commun, 434 (2013) 95-101.

[16] K.M. VanderMolen, W. McCulloch, C.J. Pearce, N.H. Oberlies, Romidepsin (Istodax, NSC 630176, FR901228, FK228, depsipeptide): a natural product recently approved for cutaneous T-cell lymphoma, J Antibiot (Tokyo), 64 (2011) 525-531.

[17] G. Garcia-Manero, S. Assouline, J. Cortes, Z. Estrov, H. Kantarjian, H. Yang, W.M. Newsome, W.H. Miller, Jr., C. Rousseau, A. Kalita, C. Bonfils, M. Dubay, T.A. Patterson, Z. Li, J.M. Besterman, G. Reid, E. Laille, R.E. Martell, M. Minden, Phase 1 study of the oral isotype specific histone deacetylase inhibitor MGCD0103 in leukemia, Blood, 112 (2008) 981-989.

[18] J. Knipstein, L. Gore, Entinostat for treatment of solid tumors and hematologic malignancies, Expert Opin Investig Drugs, 20 (2011) 1455-1467.

[19] L.K. Gediya, A. Belosay, A. Khandelwal, P. Purushottamachar, V.C. Njar, Improved synthesis of histone deacetylase inhibitors (HDIs) (MS-275 and CI-994) and inhibitory effects of HDIs alone or in combination with RAMBAs or retinoids on growth of human LNCaP prostate cancer cells and tumor xenografts, Bioorg Med Chem, 16 (2008) 3352-3360.

[20] G. Finazzi, A.M. Vannucchi, V. Martinelli, M. Ruggeri, F. Nobile, G. Specchia, E.M. Pogliani, O.M. Olimpieri, G. Fioritoni, C. Musolino, D. Cilloni, P. Sivera, G. Barosi, M.C. Finazzi, S. Di Tollo, T. Demuth, T. Barbui, A. Rambaldi, A phase II study of Givinostat in combination with hydroxycarbamide in patients with polycythaemia vera unresponsive to hydroxycarbamide monotherapy, Br J Haematol, 161 (2013) 688-694.

[21] A.M. Evens, S. Balasubramanian, J.M. Vose, W. Harb, L.I. Gordon, R. Langdon, J. Sprague, M. Sirisawad, C. Mani, J. Yue, Y. Luan, S. Horton, T. Graef, N.L. Bartlett, A Phase I/II Multicenter, Open-Label Study of the Oral Histone Deacetylase Inhibitor Abexinostat in Relapsed/Refractory Lymphoma, Clin Cancer Res, 22 (2016) 1059-1066.

[22] T.A. Miller, D.J. Witter, S. Belvedere, Histone deacetylase inhibitors, J Med Chem, 46 (2003) 5097-5116.

[23] P. Bertrand, Inside HDAC with HDAC inhibitors, Eur J Med Chem, 45 (2010) 2095-2116.

[24] M. Mottamal, S. Zheng, T.L. Huang, G. Wang, Histone deacetylase inhibitors in clinical studies as templates for new anticancer agents, Molecules, 20 (2015) 3898-3941.

[25] S. Shen, A.P. Kozikowski, Why Hydroxamates May Not Be the Best Histone Deacetylase Inhibitors--What Some May Have Forgotten or Would Rather Forget?, ChemMedChem, 11 (2016) 15-21.

[26] N. Khan, M. Jeffers, S. Kumar, C. Hackett, F. Boldog, N. Khramtsov, X. Qian, E. Mills, S.C. Berghs, N. Carey, P.W. Finn, L.S. Collins, A. Tumber, J.W. Ritchie, P.B. Jensen, H.S. Lichenstein, M. Sehested, Determination of the class and isoform selectivity of small-molecule histone deacetylase inhibitors, Biochem J, 409 (2008) 581-589.

[27] O. Bruserud, C. Stapnes, E. Ersvaer, B.T. Gjertsen, A. Ryningen, Histone deacetylase inhibitors in cancer treatment: a review of the clinical toxicity and the modulation of gene expression in cancer cell, Curr Pharm Biotechnol, 8 (2007) 388-400.

[28] S. Subramanian, S.E. Bates, J.J. Wright, I. Espinoza-Delgado, R.L. Piekarz, Clinical Toxicities of Histone Deacetylase Inhibitors, Pharmaceuticals (Basel), 3 (2010) 2751-2767.

[29] J. Roche, P. Bertrand, Inside HDACs with more selective HDAC inhibitors, Eur J Med Chem, 121 (2016) 451-483.

[30] S.N. Ononye, M. van Heyst, E.M. Falcone, A.C. Anderson, D.L. Wright, Toward isozyme-selective inhibitors of histone deacetylase as therapeutic agents for the treatment of cancer, Pharm Pat Anal, 1 (2012) 207-221.

[31] T. Beckers, C. Burkhardt, H. Wieland, P. Gimmnich, T. Ciossek, T. Maier, K. Sanders, Distinct pharmacological properties of second generation HDAC inhibitors with the benzamide or hydroxamate head group, Int J Cancer, 121 (2007) 1138-1148.

[32] E. Hu, E. Dul, C.M. Sung, Z. Chen, R. Kirkpatrick, G.F. Zhang, K. Johanson, R. Liu, A. Lago, G. Hofmann, R. Macarron, M. de los Frailes, P. Perez, J. Krawiec, J. Winkler, M. Jaye, Identification of novel isoform-selective inhibitors within class I histone deacetylases, J Pharmacol Exp Ther, 307 (2003) 720-728.

[33] J.H. Choi, H.J. Kwon, B.I. Yoon, J.H. Kim, S.U. Han, H.J. Joo, D.Y. Kim, Expression profile of histone deacetylase 1 in gastric cancer tissues, Jpn J Cancer Res, 92 (2001) 1300-1304.

[34] K. Halkidou, L. Gaughan, S. Cook, H.Y. Leung, D.E. Neal, C.N. Robson, Upregulation and nuclear recruitment of HDAC1 in hormone refractory prostate cancer, Prostate, 59 (2004) 177-189.

[35] K. Miyake, T. Yoshizumi, S. Imura, K. Sugimoto, E. Batmunkh, H. Kanemura, Y. Morine, M. Shimada, Expression of hypoxia-inducible factor-1alpha, histone deacetylase 1, and metastasis-associated protein 1 in pancreatic carcinoma: correlation with poor prognosis with possible regulation, Pancreas, 36 (2008) e1-9.

[36] M. Thangaraju, K.N. Carswell, P.D. Prasad, V. Ganapathy, Colon cancer cells maintain low levels of pyruvate to avoid cell death caused by inhibition of HDAC1/HDAC3, Biochem J, 417 (2009) 379-389.

[37] Z. Zhang, H. Yamashita, T. Toyama, H. Sugiura, Y. Ando, K. Mita, M. Hamaguchi, Y. Hara, S. Kobayashi, H. Iwase, Quantitation of HDAC1 mRNA expression in invasive carcinoma of the breast*, Breast Cancer Res Treat, 94 (2005) 11-16.

[38] J. Luo, F. Su, D. Chen, A. Shiloh, W. Gu, Deacetylation of p53 modulates its effect on cell growth and apoptosis, Nature, 408 (2000) 377-381.

[39] S.J. Macalino, V. Gosu, S. Hong, S. Choi, Role of computer-aided drug design in modern drug discovery, Arch Pharm Res, 38 (2015) 1686-1701.

[40] B.K. Shoichet, Virtual screening of chemical libraries, Nature, 432 (2004) 862-865.

[41] P. Prathipati, A. Dixit, A. Saxena, Computer-Aided Drug Design: Integration of Structure-Based and Ligand-Based Approaches in Drug Design, Curr Comput Aided Drug Des, 3 (2007) 133-148.

[42] H. Sirous, G. Chemi, G. Campiani, S. Brogi, An integrated in silico screening strategy for identifying promising disruptors of p53-MDM2 interaction, Comput Biol Chem, 83 (2019) 107105.

[43] S. Brogi, M. Kladi, C. Vagias, P. Papazafiri, V. Roussis, A. Tafi, Pharmacophore modeling for qualitative prediction of antiestrogenic activity, J Chem Inf Model, 49 (2009) 2489-2497.

[44] S. Brogi, P. Papazafiri, V. Roussis, A. Tafi, 3D-QSAR using pharmacophore-based alignment and virtual screening for discovery of novel MCF-7 cell line inhibitors, Eur J Med Chem, 67 (2013) 344-351.

[45] L. Zaccagnini, S. Brogi, M. Brindisi, S. Gemma, G. Chemi, G. Legname, G. Campiani, S. Butini, Identification of novel fluorescent probes preventing PrP(Sc) replication in prion diseases, Eur J Med Chem, 127 (2017) 859-873.

[46] G. Melagraki, A. Afantitis, H. Sarimveis, P.A. Koutentis, G. Kollias, O. Igglessi-Markopoulou, Predictive QSAR workflow for the in silico identification and screening of novel HDAC inhibitors, Mol Divers, 13 (2009) 301-311.

[47] A.I. Uba, K. Yelekci, Identification of potential isoform-selective histone deacetylase inhibitors for cancer therapy: a combined approach of structure-based virtual screening, ADMET prediction and molecular dynamics simulation assay, J Biomol Struct Dyn, 36 (2018) 3231-3245.

[48] A. Ibrahim Uba, K. Yelekci, Homology modeling of human histone deacetylase 10 and design of potential selective inhibitors, J Biomol Struct Dyn, 37 (2019) 3627-3636.

[49] L. Zhang, M. Li, J. Feng, H. Fang, W. Xu, Discovery of a novel histone deacetylase 8 inhibitor by virtual screening, Med Chem Res, 21 (2010) 152-156.

[50] S.B. Nair, M.K. Teli, H. Pradeep, G.K. Rajanikant, Computational identification of novel histone deacetylase inhibitors by docking based QSAR, Comput Biol Med, 42 (2012) 697-705.

[51] S. Vadivelan, B.N. Sinha, G. Rambabu, K. Boppana, S.A. Jagarlapudi, Pharmacophore modeling and virtual screening studies to design some potential histone deacetylase inhibitors as new leads, J Mol Graph Model, 26 (2008) 935-946.

[52] S. Krishna, A.D. Lakra, N. Shukla, S. Khan, D.P. Mishra, S. Ahmed, M.I. Siddiqi, Identification of potential histone deacetylase1 (HDAC1) inhibitors using multistep virtual screening approach including SVM model, pharmacophore modeling, molecular docking and biological evaluation, J Biomol Struct Dyn, 38 (2020) 3280-3295.

[53] H. Sirous, G. Campiani, S. Brogi, V. Calderone, G. Chemi, Computer-Driven Development of an in Silico Tool for Finding Selective Histone Deacetylase 1 Inhibitors, Molecules, 25 (2020).

[54] Maestro, version 10.1, Schrödinger, LLC, New York, NY, 2015.

[55] S. Kim, P.A. Thiessen, E.E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B.A. Shoemaker, J. Wang, B. Yu, J. Zhang, S.H. Bryant, PubChem Substance and Compound databases, Nucleic Acids Res, 44 (2016) D1202-1213.

[56] MacroModel, version 10.7, Schrödinger, LLC, New York, NY, 2015.

[57] LigPrep, version 3.3, Schrödinger, LLC, New York, NY, 2015.

[58] H. Sirous, G. Chemi, S. Gemma, S. Butini, Z. Debyser, F. Christ, L. Saghaie, S. Brogi, A. Fassihi, G. Campiani, M. Brindisi, Identification of Novel 3-Hydroxy-pyran-4-One Derivatives as Potent HIV-1 Integrase Inhibitors Using in silico Structure-Based Combinatorial Library Design Approach, Front Chem, 7 (2019) 574.

[59] H. Sirous, A. Fassihi, S. Brogi, G. Campiani, F. Christ, Z. Debyser, S. Gemma, S. Butini, G. Chemi, A. Grillo, R. Zabihollahi, M.R. Aghasadeghi, L. Saghaie, H.R. Memarian, Synthesis, Molecular Modelling and Biological Studies of 3-hydroxypyrane- 4-one and 3-

hydroxy-pyridine-4-one Derivatives as HIV-1 Integrase Inhibitors, Med Chem, 15 (2019) 755-770.

[60] W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids, J Am Chem Soc, 118 (1996) 11225-11236.

[61] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, W.L. Jorgensen, Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides†, J Phys Chem B, 105 (2001) 6474-6487.

[62] W.C. Still, A. Tempczyk, R.C. Hawley, T. Hendrickson, Semianalytical treatment of solvation for molecular mechanics and dynamics, J Am Chem Soc, 112 (2002) 6127-6129.

[63] Phase, version 4.2, Schrödinger, LLC, New York, NY, 2015.

[64] S.L. Dixon, A.M. Smondyrev, E.H. Knoll, S.N. Rao, D.E. Shaw, R.A. Friesner, PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results, J Comput Aided Mol Des, 20 (2006) 647-671.

[65] C.J. Millard, P.J. Watson, I. Celardo, Y. Gordiyenko, S.M. Cowley, C.V. Robinson, L. Fairall, J.W. Schwabe, Class I HDACs share a common mechanism of regulation by inositol phosphates, Mol Cell, 51 (2013) 57-67.

[66] Protein Preparation Wizard 2015, Epik version 2.4, Schrödinger, LLC, New York, NY, 2015; Impact version 5.9, Schrödinger, LLC, New York, NY, 2015; Prime version 3.2, Schrödinger, LLC, New York, NY, 2015.

[67] G.M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments, J Comput Aided Mol Des, 27 (2013) 221-234.

[68] M. Brindisi, C. Cavella, S. Brogi, A. Nebbioso, J. Senger, S. Maramai, A. Ciotta, C. Iside, S. Butini, S. Lamponi, E. Novellino, L. Altucci, M. Jung, G. Campiani, S. Gemma, Phenylpyrrole-based HDAC inhibitors: synthesis, molecular modeling and biological studies, Future Med Chem, 8 (2016) 1573-1587.

[69] M. Brindisi, J. Senger, C. Cavella, A. Grillo, G. Chemi, S. Gemma, D.M. Cucinella, S. Lamponi, F. Sarno, C. Iside, A. Nebbioso, E. Novellino, T.B. Shaik, C. Romier, D. Herp, M. Jung, S. Butini, G. Campiani, L. Altucci, S. Brogi, Novel spiroindoline HDAC inhibitors: Synthesis, molecular modelling and biological studies, Eur J Med Chem, 157 (2018) 127-138.

[70] Glide, version 6.6. Schrödinger, LLC, New York, NY, 2015.

[71] Virtual Screening Workflow 2015-3, Glide version 6.6, LigPrep version 3.3, QikProp version 4.3, Schrödinger, LLC, New York, NY, 2015.

[72] R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, P.S. Shenkin, Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, J Med Chem, 47 (2004) 1739-1749.

[73] R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, P.C. Sanschagrin, D.T. Mainz, Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, J Med Chem, 49 (2006) 6177-6196.

[74] The PyMOL Molecular Graphics System, 2013, PyMOL Version 1.6-alpha, Schrodinger, LLC, New York, NY.

[75] P.A. Smith, M.J. Sorich, L.S. Low, R.A. McKinnon, J.O. Miners, Towards integrated ADME prediction: past, present and future directions for modelling metabolism by UDP-glucuronosyltransferases, J Mol Graph Model, 22 (2004) 507-517.

[76] QikProp, version 4.3, Schrödinger, LLC, New York, NY, 2015.

[77] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings 1PII of original article: S0169-409X(96)00423-1. The article was originally published in Advanced Drug Delivery Reviews 23 (1997) 3–25. 1, Adv Drug Deliv Rev, 46 (2001) 3-26.

[78] FAFDrugs, http://www.fafdrugs4.mti.univ-paris-diderot.fr 4.0 (Accessed 15 october 2020).

[79] D. Lagorce, J. Maupetit, J. Baell, O. Sperandio, P. Tuffery, M.A. Miteva, H. Galons, B.O. Villoutreix, The FAF-Drugs2 server: a multistep engine to prepare electronic chemical compound collections, Bioinformatics, 27 (2011) 2018-2020.

[80] D. Lagorce, O. Sperandio, H. Galons, M.A. Miteva, B.O. Villoutreix, FAF-Drugs2: free ADME/tox filtering tool to assist drug discovery and chemical biology projects, BMC Bioinformatics, 9 (2008) 396.

[81] Induced Fit Docking protocol 2015-1, Glide version 6.4, Prime version 3.7, Schrödinger, LLC, New York, NY, 2015.

[82] W. Sherman, T. Day, M.P. Jacobson, R.A. Friesner, R. Farid, Novel procedure for modelling ligand/receptor induced fit effects, J Med Chem, 49 (2006) 534-553.

[83] W. Sherman, H.S. Beard, R. Farid, Use of an induced fit receptor structure in virtual screening, Chem Biol Drug Des, 67 (2006) 83-84.

[84] M.P. Jacobson, D.L. Pincus, C.S. Rapp, T.J. Day, B. Honig, D.E. Shaw, R.A. Friesner, A hierarchical approach to all-atom protein loop prediction, Proteins, 55 (2004) 351-367.

[85] M.P. Jacobson, R.A. Friesner, Z. Xiang, B. Honig, On the role of the crystal environment in determining protein side-chain conformations, J Mol Biol, 320 (2002) 597-608.

[86] D.A. Pearlman, P.S. Charifson, Are free energy calculations useful in practice? A comparison with rapid scoring functions for the p38 MAP kinase protein system, J Med Chem, 44 (2001) 3417-3423.

[87] Prime, version 3.9, Schrödinger, LLC, New York, NY, 2015.

[88] N. Huang, C. Kalyanaraman, J.J. Irwin, M.P. Jacobson, Physics-based scoring of protein-ligand complexes: enrichment of known inhibitors in large-scale virtual screening, J Chem Inf Model, 46 (2006) 243-253.

[89] B. Kuhn, P.A. Kollman, Binding of a diverse set of ligands to avidin and streptavidin: an accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models, J Med Chem, 43 (2000) 3786-3791.

[90] P.D. Lyne, M.L. Lamb, J.C. Saeh, Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring, J Med Chem, 49 (2006) 4805-4808.

[91] S.A. Adcock, J.A. McCammon, Molecular dynamics: survey of methods for simulating the activity of proteins, Chem Rev, 106 (2006) 1589-1615.

[92] J.D. Durrant, J.A. McCammon, Molecular dynamics simulations and drug discovery, BMC Biol, 9 (2011) 71.

[93] S. Brogi, H. Sirous, V. Calderone, G. Chemi, Amyloid beta fibril disruption by oleuropein aglycone: long-time molecular dynamics simulation to gain insight into the mechanism of action of this polyphenol from extra virgin olive oil, Food Funct, 11 (2020) 8122-8132.

[94] Desmond Molecular Dynamics System, version 4.1, D. E. Shaw Research, New York, NY, 2015. Maestro-Desmond Interoperability Tools, version 4.1, Schrödinger, New York, NY, 2015.

[95] K.J. Bowers, D.E. Chow, H. Xu, R.O. Dror, M.P. Eastwood, B.A. Gregersen, J.L. Klepeis, I. Kolossvary, M.A. Moraes, F.D. Sacerdoti, J.K. Salmon, Y. Shan, D.E. Shaw, Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters, SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, (2006) 43-43.

[96] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J Chem Phys, 79 (1983) 926-935.

[97] D.D. Humphreys, R.A. Friesner, B.J. Berne, A Multiple-Time-Step Molecular Dynamics Algorithm for Macromolecules, J Phys Chem, 98 (1994) 6885-6892.

[98] W.G. Hoover, Canonical dynamics: Equilibrium phase-space distributions, Phys Rev A Gen Phys, 31 (1985) 1695-1697.

[99] G.J. Martyna, D.J. Tobias, M.L. Klein, Constant pressure molecular dynamics algorithms, J Chem Phys, 101 (1994) 4177-4189.

[100] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, J Chem Phys, 103 (1995) 8577-8593.

[101] N.A. Hosea, H.M. Jones, Predicting pharmacokinetic profiles using in silico derived parameters, Mol Pharm, 10 (2013) 1207-1215.

[102] M. Brindisi, S. Brogi, S. Maramai, A. Grillo, G. Borrelli, S. Butini, E. Novellino, M. Allarà, A. Ligresti, G. Campiani, V. Di Marzo, S. Gemma, Harnessing the pyrroloquinoxaline scaffold for FAAH and MAGL interaction: definition of the structural determinants for enzyme inhibition, RSC Adv, 6 (2016) 64651-64664.

[103] M. Brindisi, G. Borrelli, S. Brogi, A. Grillo, S. Maramai, M. Paolino, M. Benedusi, A. Pecorelli, G. Valacchi, L. Di Cesare Mannelli, C. Ghelardini, M. Allara, A. Ligresti, P. Minetti, G. Campiani, V. di Marzo, S. Butini, S. Gemma, Development of Potent Inhibitors of Fatty Acid Amide Hydrolase Useful for the Treatment of Neuropathic Pain, ChemMedChem, 13 (2018) 2090-2103.

[104] A. Grillo, G. Chemi, S. Brogi, M. Brindisi, N. Relitti, F. Fezza, D. Fazio, L. Castelletti, E. Perdona, A. Wong, S. Lamponi, A. Pecorelli, M. Benedusi, M. Fantacci, M. Valoti, G. Valacchi, F. Micheli, E. Novellino, G. Campiani, S. Butini, M. Maccarrone, S. Gemma, Development of novel multipotent compounds modulating endocannabinoid and dopaminergic systems, Eur J Med Chem, 183 (2019) 111674.

[105] Y. Sixto-Lopez, M. Bello, J. Correa-Basurto, Insights into structural features of HDAC1 and its selectivity inhibition elucidated by Molecular dynamic simulation and Molecular Docking, J Biomol Struct Dyn, 37 (2019) 584-610.

[106] E. Pontiki, D. Hadjipavlou-Litina, Histone deacetylase inhibitors (HDACIs). Structure—activity relationships: history and new QSAR perspectives, Med Res Rev, 32 (2012) 1-165.

[107] T. Abdizadeh, M.R. Kalani, K. Abnous, Z. Tayarani-Najaran, B.Z. Khashyarmanesh, R. Abdizadeh, R. Ghodsi, F. Hadizadeh, Design, synthesis and biological evaluation of novel coumarin-based benzamides as potent histone deacetylase inhibitors and anticancer agents, Eur J Med Chem, 132 (2017) 42-62.

[108] R.D. Taylor, P.J. Jewsbury, J.W. Essex, A review of protein-small molecule docking methods, J Comput Aided Mol Des, 16 (2002) 151-166.

[109] A. Hospital, J.R. Goni, M. Orozco, J.L. Gelpi, Molecular dynamics simulations: advances and applications, Adv Appl Bioinform Chem, 8 (2015) 37-47.

Highlights

- An exhaustive virtual screening protocol was settled to screen more than 736,000 molecules as HDAC1 inhibitors.
- An in-house 3D-QSAR model was used for discriminating hits with the highest predicted HDAC1 inhibitory activity.
- In silico structure-based protocols includes IFD, MM-GBSA calculations and MD simulations.
- In silico ADMET and PAINs analysis, were performed to select an enriched set of the best drug-like molecules.
- Six top-ranked hits with the best in silico pharmacodynamics and pharmacokinetic profiles

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that can influence the work reported in this paper.

Figure 1. Chemical structures of FDA-approved HDACIs and drugs in clinical trials

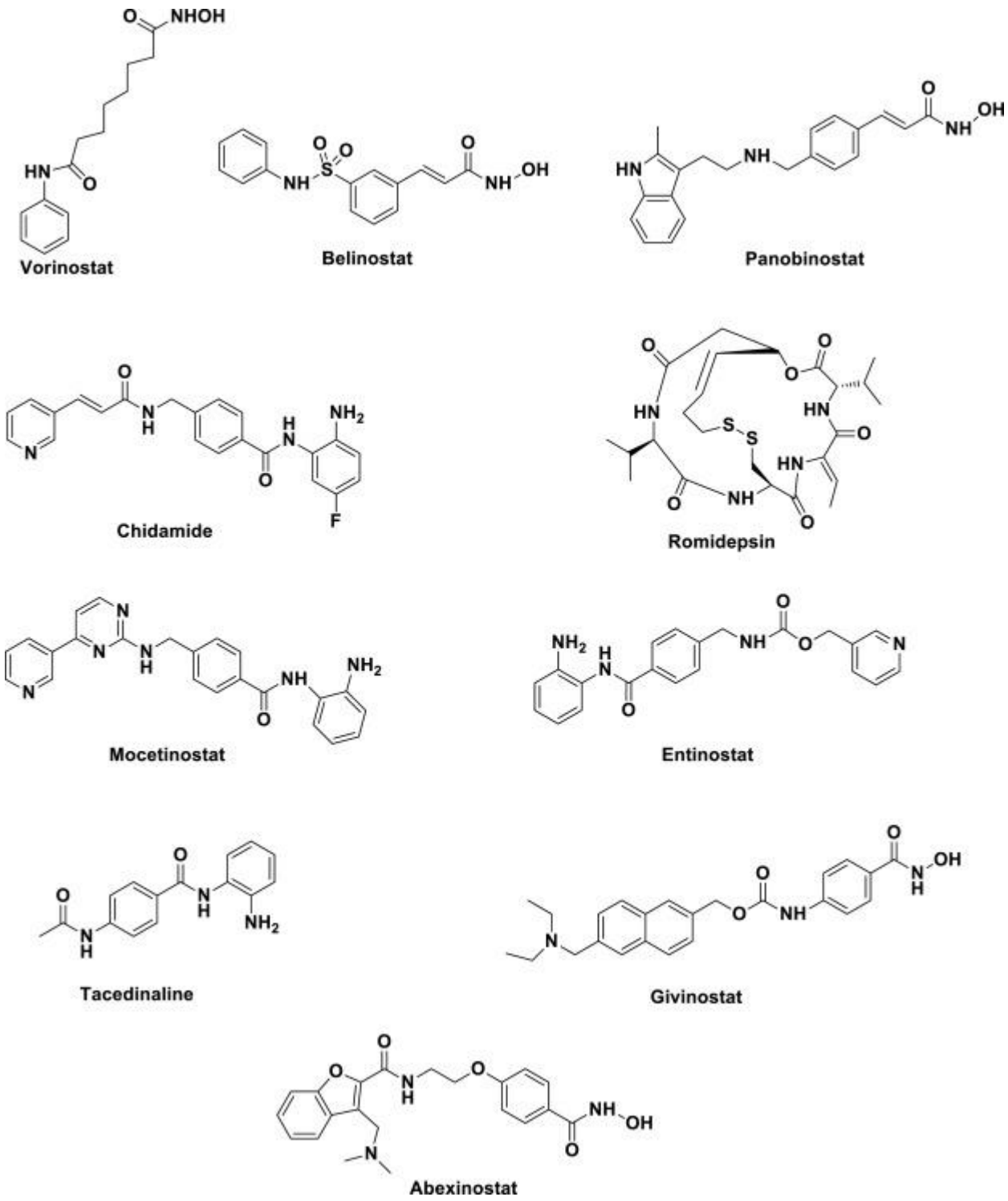


Figure 2. Schematic representation of the virtual screening process employed in this study

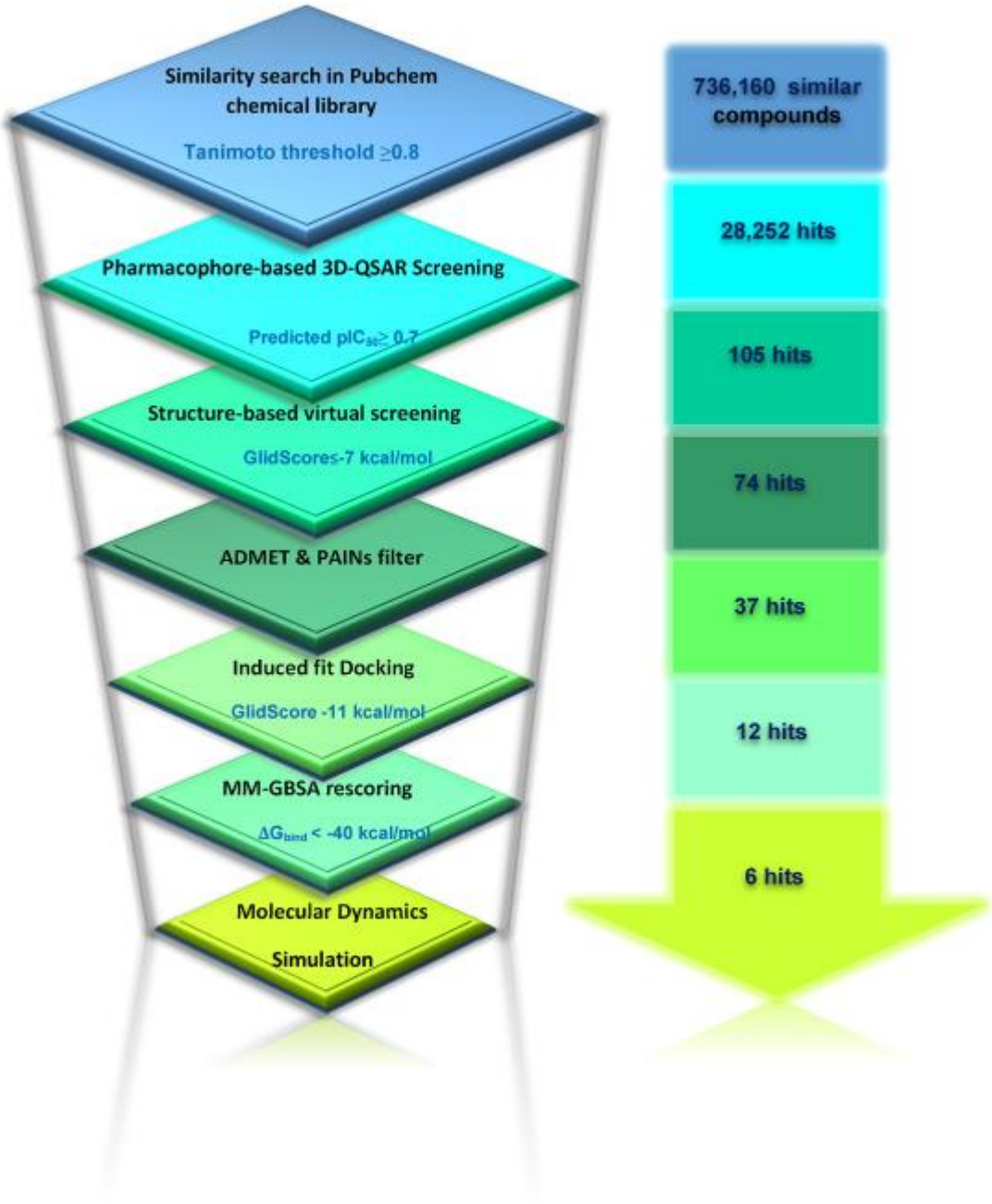


Figure 3. Chemical structures of the best hits nominated out by MM-GBSA studies as possible HDAC1 inhibitors

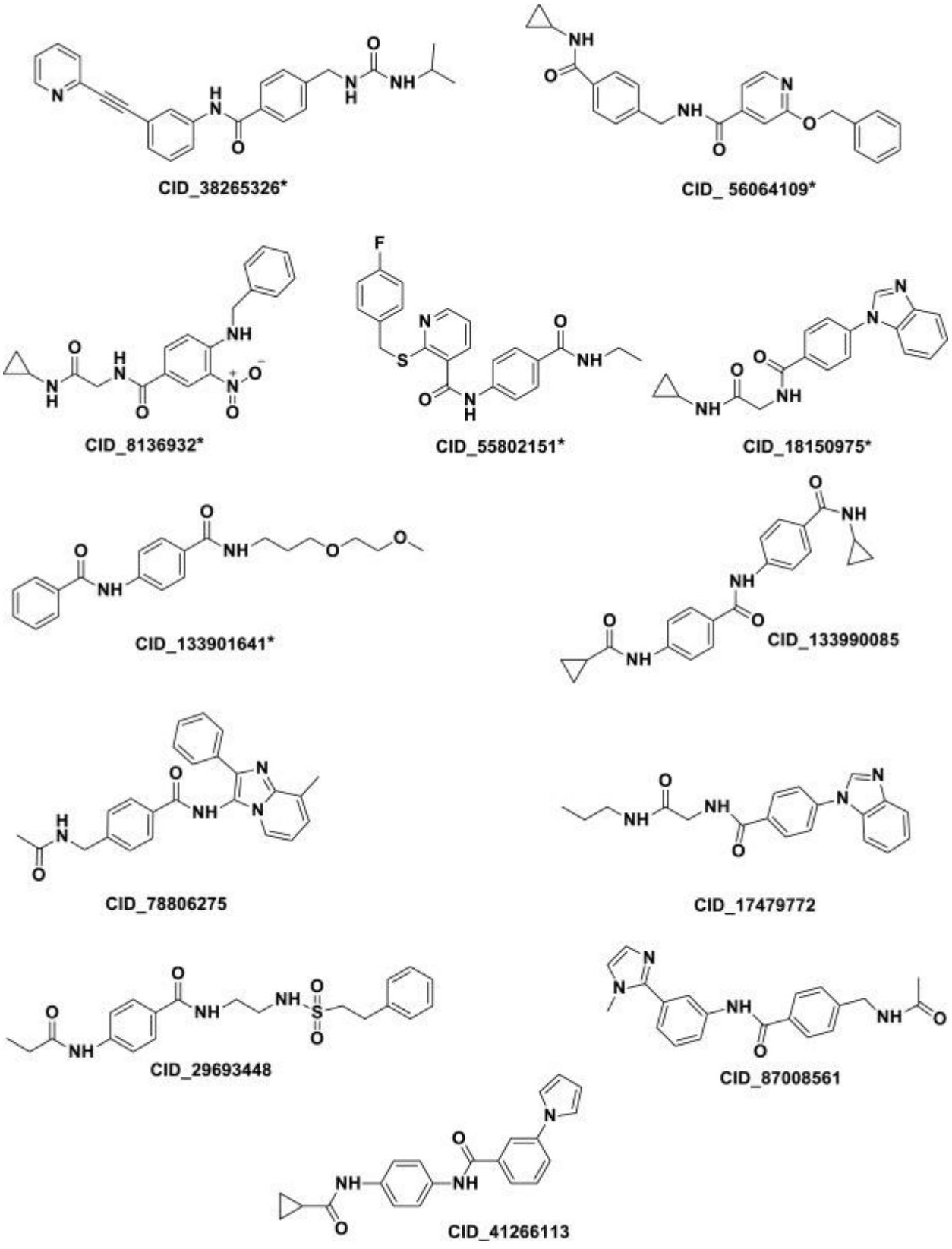
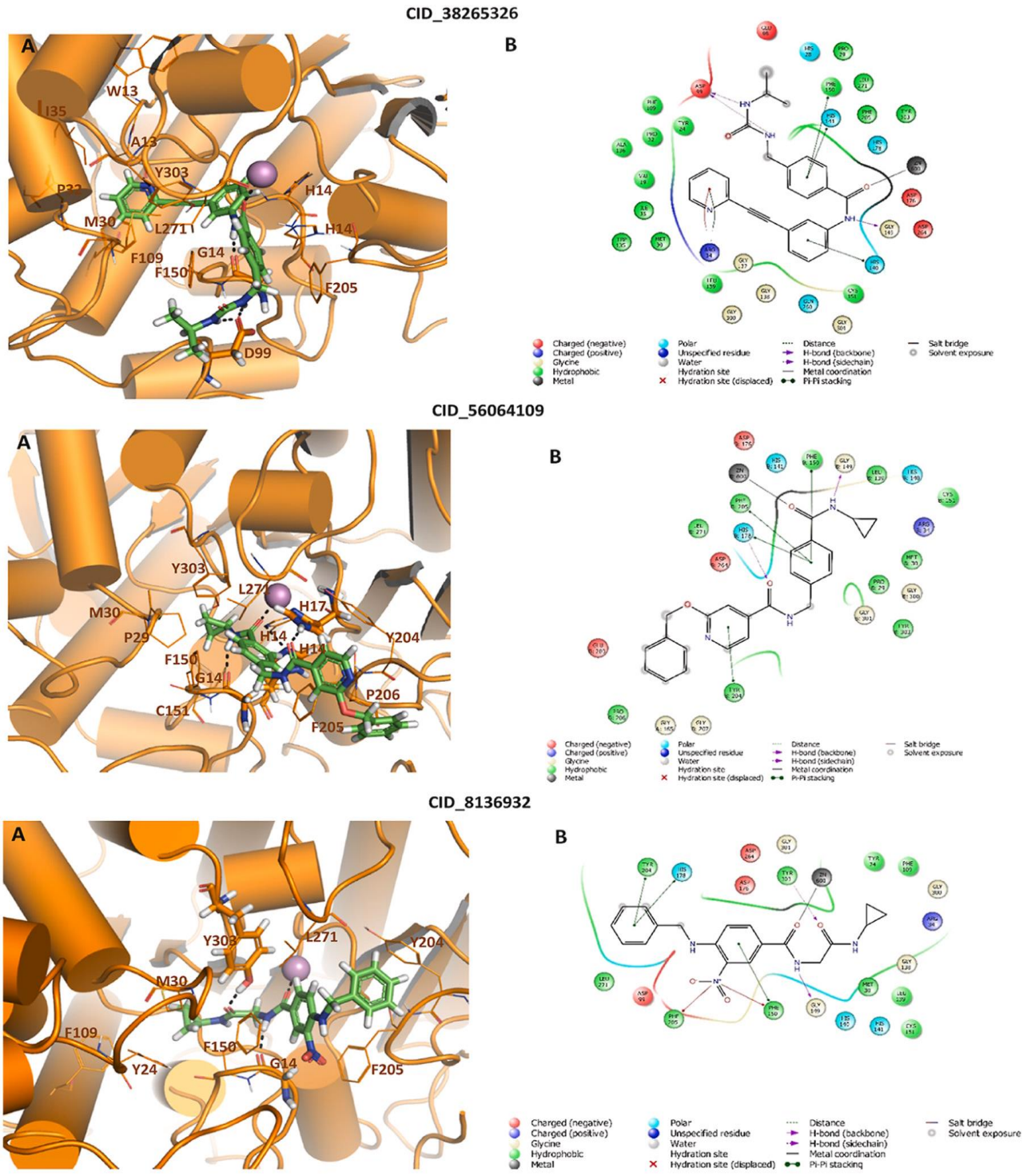
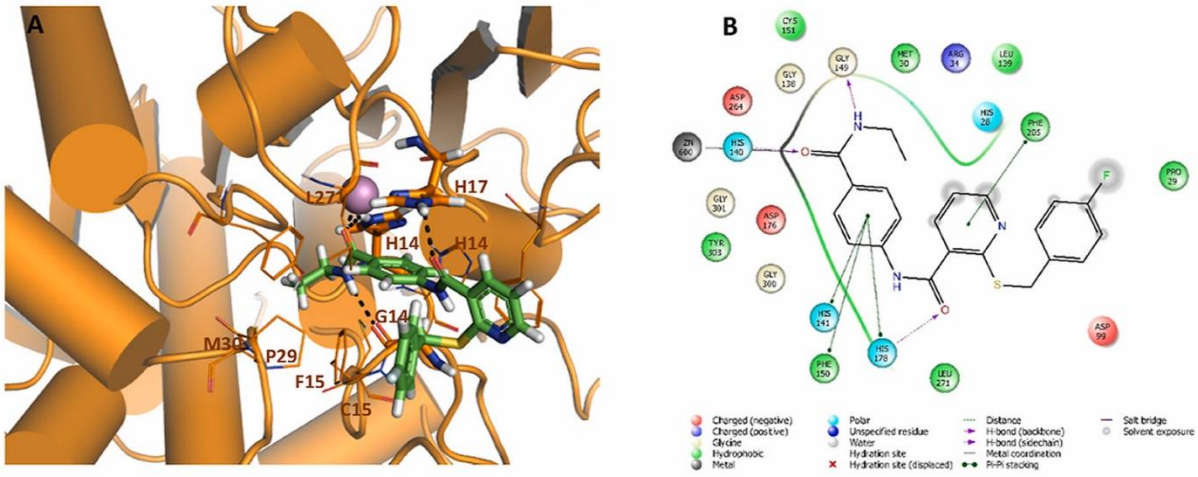


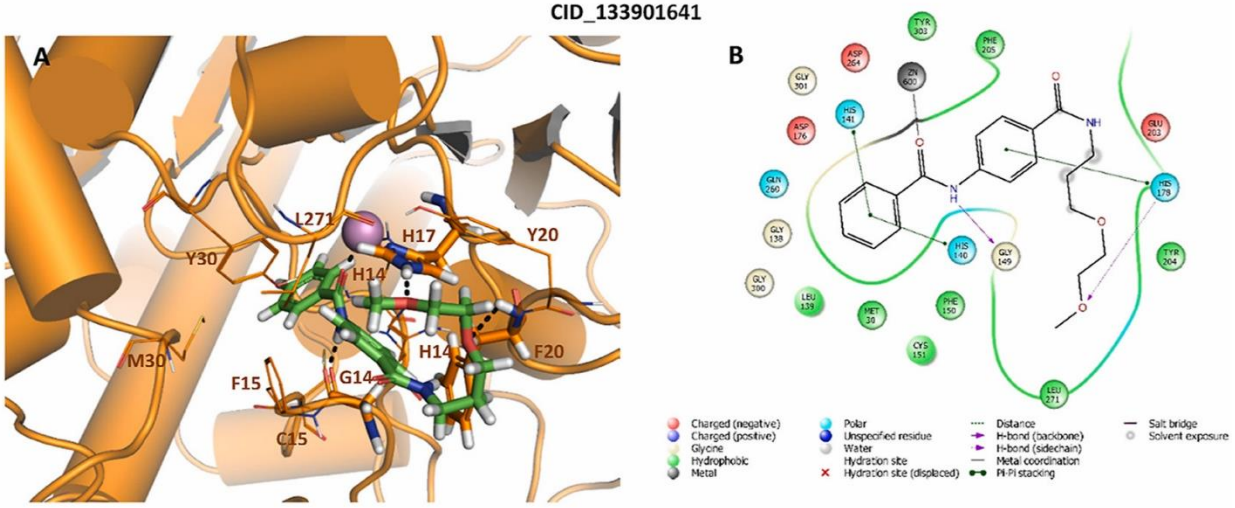
Figure 4. A) The putative binding mode of the six top-ranked compounds as found by MM-GBSA method into the active site of HDAC1. In each case, the protein and Zn^{2+} ion are represented as orange cartoon and violet spheres, respectively. The main hydrogen bond interacting residues and ligand in the binding site are shown as stick models and colored by elements. H-bonds and metal-coordination bonds are marked as black dotted lines. The figures were prepared using the PyMOL. (B) The representation of the 2D interaction diagram of the five top-screened hits with the key amino acid residues in the HDAC1 active site. The figures were generated using the ligand-interaction diagram available in Maestro modelling environment suite.



CID_55802151



CID_133901641



CID_18150975

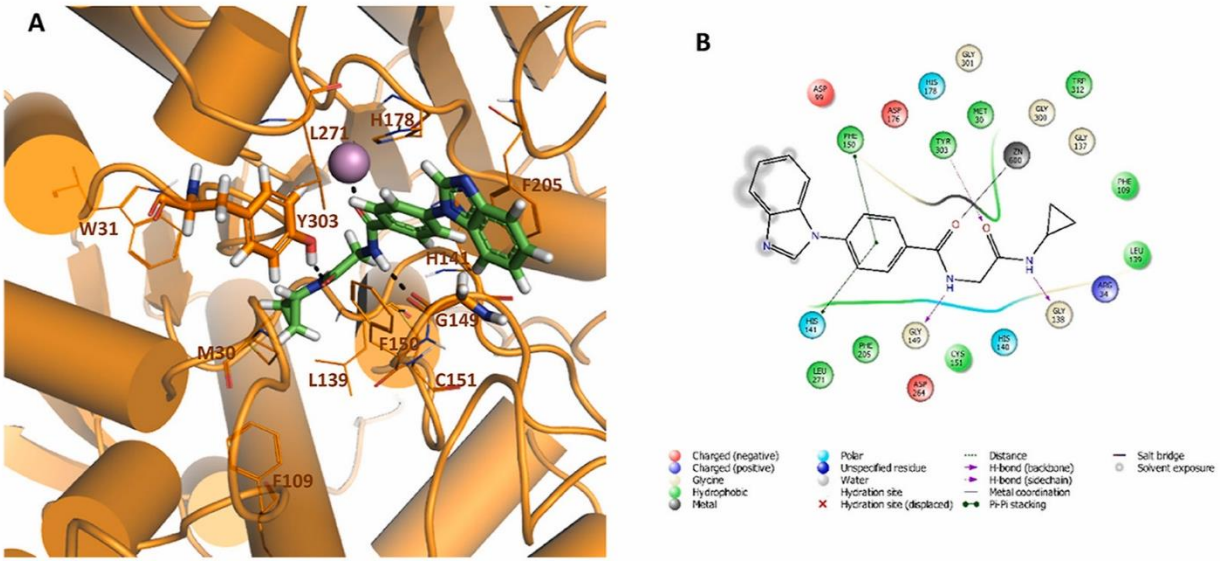


Figure 5. Superimposed RMSD (upper bound) graph of the backbone atoms of HDAC1 in complex with the best hits; **CID_38265326** (blue), **CID_56064109** (orange), **CID_8136932** (pink), **CID_55802151** (gray), **CID_133901641** (yellow), and **CID_18150975** (green), as well as known inhibitor Entinostat (purple). RMSD were calculated between the final conformation and the starting conformation through the 30 ns of the MD simulation.

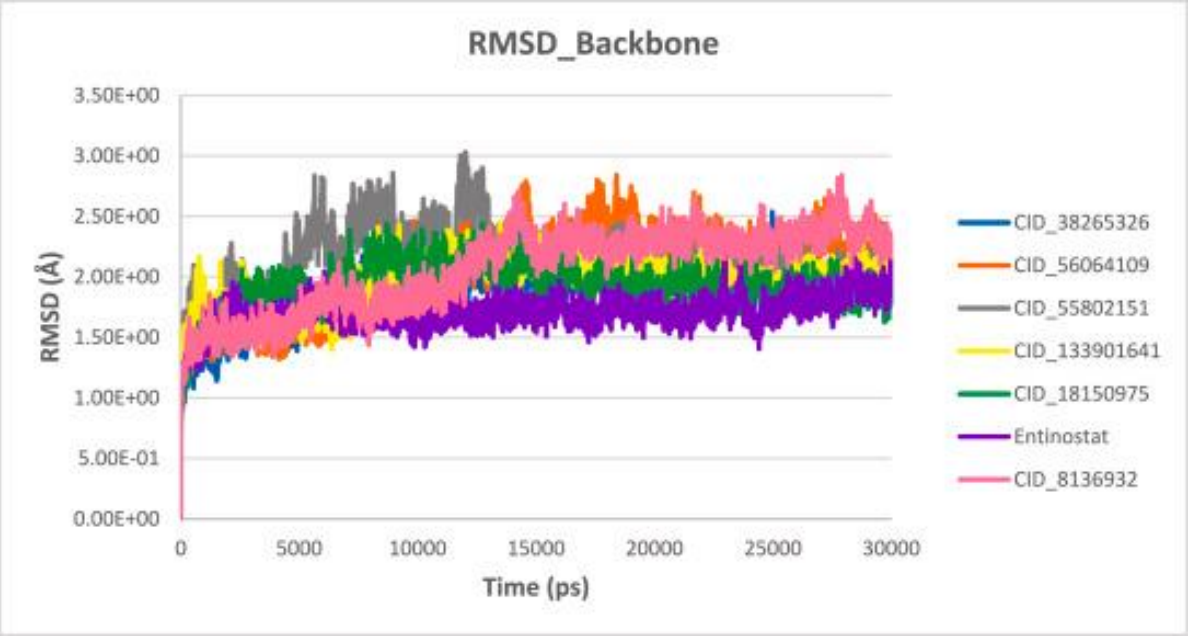


Figure 6. Superimposed RMSF graph of MDM2 in complex with the best hits; **CID_38265326** (blue), **CID_56064109** (orange), **CID_8136932** (pink), **CID_55802151** (gray), **CID_133901641** (yellow), and **CID_18150975** (green), as well as known inhibitor Entinostat (purple), obtained from 30 ns MD simulation.

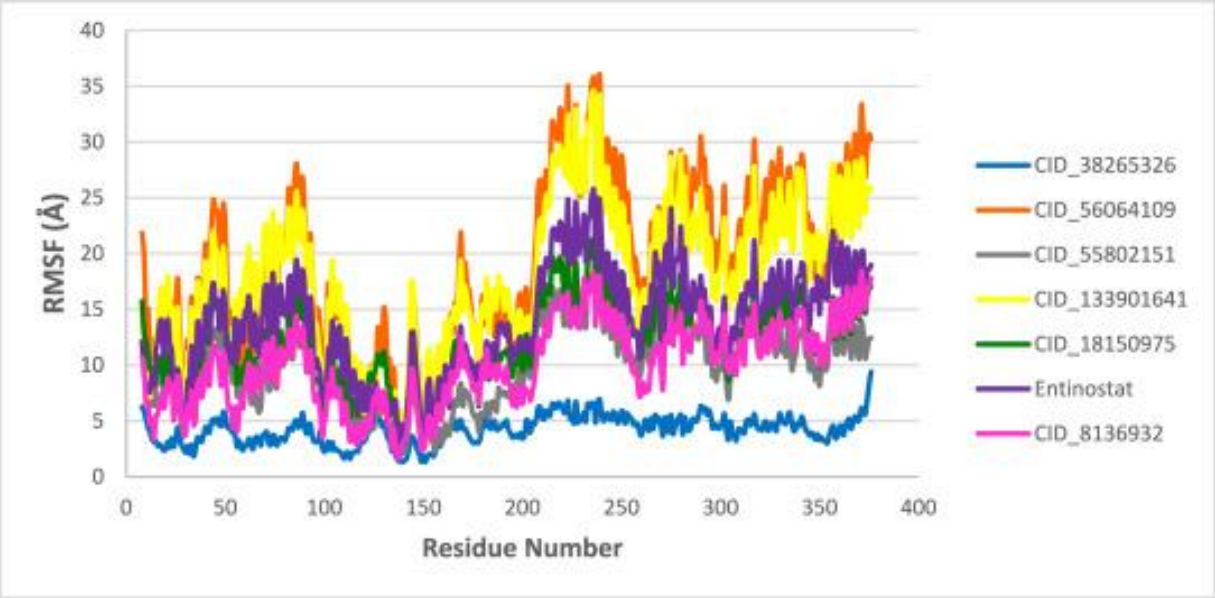
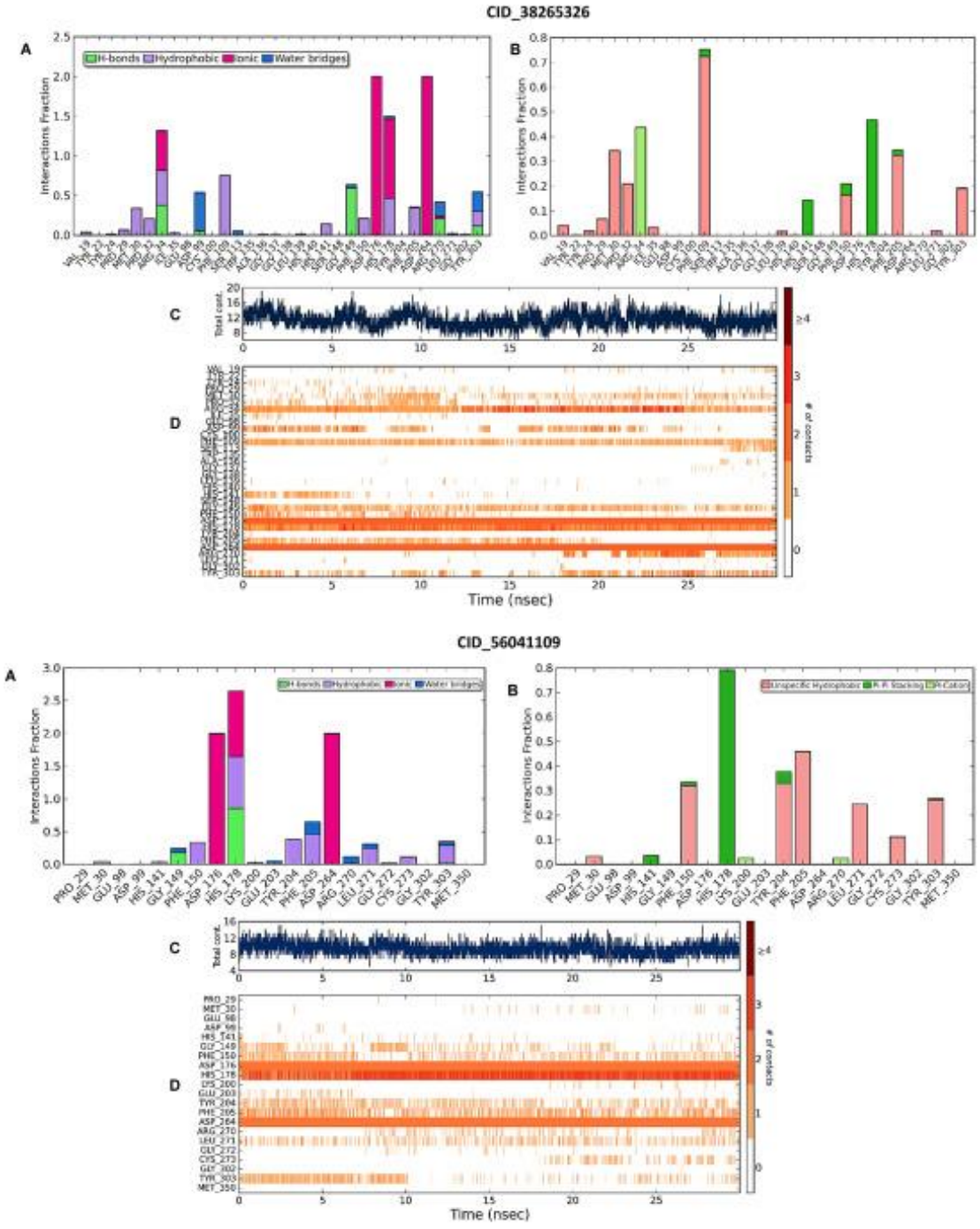


Figure 7. Protein interactions with the three top-ranked hits; **CID_38265326**, **CID_56064109** and **CID_8136932** monitored throughout the simulation. These interactions can be clustered by type and summarized, as shown in the plots. (A) Categorization of protein-ligand interactions into four types: H-bonds, hydrophobic, ionic, and water bridges. (B) Hydrophobic contacts of the ligand into the HDAC1 binding site over the course of the trajectory. (C) The representation of the total number of specific contacts that the protein makes with the ligands over the course of the trajectory (D) a timeline representation of the interactions and contacts of panel A (H -bonds, hydrophobic, ionic, water bridges). Some residues make more than one specific contact with the ligand, which is represented by a darker shade of orange, according to the scale to the right of the plot.



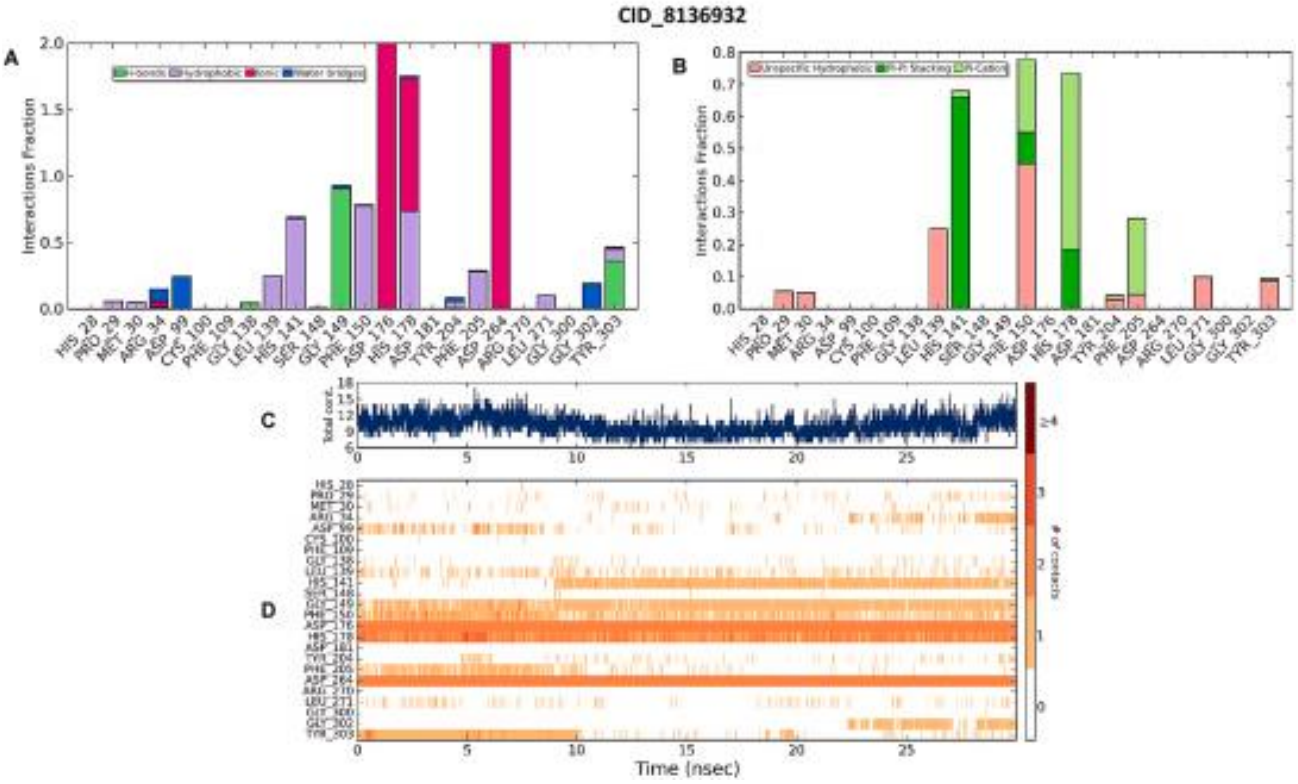

CID_8136932
A
B
C
D
Interactions Fraction
Total cont.
# of contacts
Time (nsec)

Figure 8. 2D schematic of the detailed atomic interactions of the best hits, **CID_38265326**, **CID_56064109**, **CID_8136932**, **CID_55802151**, **CID_133901641**, and **CID_18150975**, with the protein residues monitored throughout the simulation. In each case, interactions that occur more than 5.0% of the simulation time in the selected trajectory are shown.

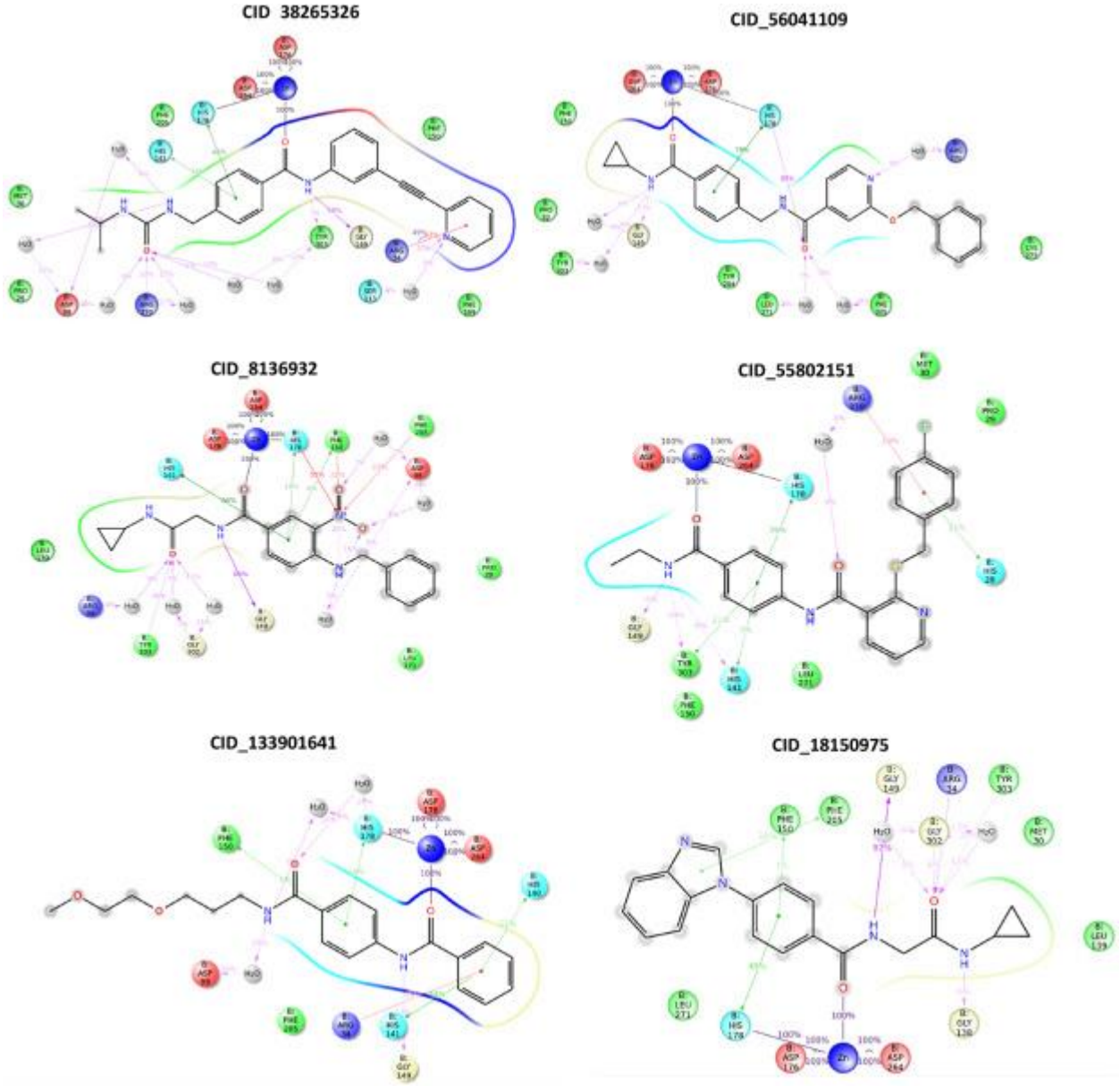


Table 1. Binding energy results of the compounds screened by Prime/MM-GBSA calculations.

entry	ΔG_{bind}[a] (Kcal/mol)	ΔG_{Coulom}[b]	$\Delta G_{Covalent}$[c]	ΔG_{Hbond}[d]	ΔG_{Lipo}[e]	ΔG_{SolvGB}[f]	ΔG_{vdW}[g]
CID_38265326	-54.095	-6.612	2.544	-1.232	-45.962	9.449	-47.401
CID_56064109	-49.947	-5.892	11.670	-0.523	-38.906	32.564	-48.004
CID_8136932	-49.046	37.220	13.041	-5.000	-37.852	- 9.097	-58.336
CID_55802151	-48.745	15.105	3.749	-0.691	-37.878	17.406	-51.870
CID_133901641	-47.975	20.358	4.452	-1.124	-32.390	10.085	-53.583
CID_18150975	-47.873	-2.626	8.857	-0.146	-31.002	21.174	-37.724
CID_133990085	-46.099	28.226	7.512	-1.225	-34.867	11.343	-64.611
CID_78806275	-46.004	3.153	18.225	-1.089	-31.107	25.501	-35.351
CID_17479772	-45.685	-5.038	10.475	-0.277	-31.233	20.921	-39.158
CID_29693448	-42.477	75.584	2.973	-5.414	-32.005	-36.140	-60.010
CID_87008561	-40.710	-5.532	9.679	-0.979	-36.212	35.713	-44.158
CID_41266113	-40.199	-1.235	12.262	-1.939	-33.647	32.907	-51.820
Mocetinostat	-35.993	21.296	9.713	-1.000	-33.283	18.191	-24.324
Entinostat	-38.087	13.033	16.310	-0.739	-31.001	20.025	-39.976
Chidamide	-37.145	-31.014	11.130	-1.186	-24.912	37.151	-27.360
Tacedinaline	-34.598	10.017	7.029	-1.199	-30.346	30.952	-23.705

[a] Binding free energy

[b] Coulomb energy contribution to the binding free energy

[c] Covalent energy contribution to the binding free energy

[d] hydrogen bonding Contribution to the binding free energy

[e] Lipophillic binding Contribution to the binding free energy

[f] the generalized born electrostatic solvation energy Contribution to the binding free energy

[g] van der Waals energy contribution to the binding free energy.