Published in final edited form as:

J Mol Model. 2012 August ; 18(8): 3927–3939. doi:10.1007/s00894-011-1297-8.

Insights from Comprehensive Multiple Receptor Docking to HDAC8

Michael Brunsteiner[‡] and Pavel A. Petukhov^{*}

Department of Medicinal Chemistry and Pharmacognosy, College of Pharmacy, University of Illinois at Chicago, 833 South Wood Street, Chicago, IL 60612, USA

Abstract

A systematic investigation of available crystal structures of HDAC8 and of the influence of different receptor structures and docking protocols is presented. The study shows that the open conformation of HDAC8 may be preferred by ligands with flexible surface binding groups as such conformation allows the ligands to minimize their exposure to solvent upon binding. This observation allowed us to rationalize excellent potency of the pyrazole-based inhibitors compared to that of the isoxazole-based inhibitors.

Keywords

Histone deacetylase 8; docking; multiple receptor; binding site flexibility

Introduction

Histone Deacetylases (HDAC) comprise a group of proteins that were first described in the context of histone deacetylation, as part of the machinery for the control of gene transcription,¹ and since then were found to have various other regulatory functions.² The interest in HDACs and their inhibitors from the point of view of basic biology and therapeutic application is hard to underestimate. Most notably they are considered a promising cancer target,³ with two HDAC inhibitors –suberoyl anilide hydroxamic acid (SAHA, Vorinostat, **1a**, Table 1) and Romidepsin - being currently used in the treatment of cutaneous T-cell lymphoma, and a number of other compounds being in various phases of clinical development.⁴ In humans at least 11 classical isoforms (HDAC1-HDAC11) are expressed that are grouped into three classes (class I, II and IV), and each of the 11 isoforms appears to have a distinctive function and possibly different therapeutic potentials.⁴ For this reason isoform selective HDAC inhibitors are of particular interest, but only very few drug like compounds with a clear selectivity profile have been proposed to date.^{4, 5}

Molecular modeling and simulation have been used as tools in the research towards HDAC inhibitors with improved activity and/or selectivity. This work has included the generation of various types of QSAR models,^{6–11} or docking of inhibitors to protein models followed by a discussion of docking poses.^{9, 12–18} These studies provide interesting insights into structure activity relationships, and, in at least one case, an impressive accuracy in the prediction of ligand activities for an external test set.¹¹ However, few of the structure-based studies concentrated on the HDAC8 isoform and flexibility of its binding site.^{19–21} The

^{*}To whom correspondence should be addressed. PAP: Phone: 312-996-4174. Fax: 312-996-7107. pap4@uic.edu.

[‡]Present address for M.B.: Research Center Pharmaceutical Engineering, Inffeldgasse 21a/II, 8010 Graz, Austria Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12/1, 8010 Graz, Austria

modeling of HDAC8 and its inhibitors appear to be a challenging task due to a number of issues outlined below.

Interactions between inhibitors and the catalytic zinc ion, found in HDAC binding sites, are difficult to model accurately. A recent study showed that “out of the box” docking to metalloenzymes, including those containing Zn ions, can be successful when applied in virtual screening campaigns, yet its accuracy is limited and a rigorous parameterization of the metals is likely to be required for lead optimization.^{22, 23} As opposed to most other metal ions commonly found in proteins, zinc can have three different coordination numbers. An accurate description of the Zn-ligand interactions and evaluation of their contribution to the binding energy can probably not be achieved without resorting to high-level *ab initio* calculations. Attempts have been made to generate more accurate classical model potentials for zinc by including bonded energy terms between the zinc and atoms chelating with it.^{24, 25} However, such potentials require the *a priori* knowledge of the coordination number, the calculation of protein ligand interaction energies when the two are connected by bonds is not straightforward, and to our knowledge no thorough evaluation of their ability to improve the accuracy of calculated protein ligand interaction energies is available.

The part of a typical HDAC ligand³ binding to the Zn²⁺ ion at the bottom of the HDAC active site, the zinc binding group (ZBG), is connected to a more hydrophobic linker that extends through a tunnel towards the binding site entrance, towards the, so called, surface binding or CAP group (SBG) which is partially exposed to the solvent. It has been argued that variations of the SBG can potentially lead to isoform selective inhibitors,²⁶ as the residues residing at the binding site entrance show variations in different isoforms. In many cases the SBG of HDAC inhibitors is relatively large, flexible, and solvent exposed. For conventional docking/scoring approaches this makes the reliable determination of the SBGs conformation, and its contribution to the binding energy difficult. The situation is exacerbated by the fact that some of the residues on the surface of HDAC proteins close to the binding site are very flexible,²⁶ rendering traditional docking to a single rigid receptor structure unreliable. This flexibility includes extensive backbone movements due to two flexible loops. Such large variations cannot be accounted for by allowing for limited side chain flexibility as implemented in a number of docking programs.^{27, 28} One of these loops (residues PRO91-THR105, loop A) appears to be disordered in the apo-protein,²⁹ but in all existing crystal structures with a bound ligand, it has a similar conformation.^{26, 29–31} In particular the crucial Asp101 residue, that can form a hydrogen bond with a ligand, shows some variation in its orientation, but is generally in the same location throughout (Figure 1). This is not true for the second loop (residues SER30-LYS36, loop B) which, in co-crystal structures, was found in two distinctively different conformations, open (1VKG, 1T64) and closed (1T67, 1T69, 1W22, 2V5X, 2V5W),²⁶ resulting in the presence or absence of a second pocket next to the primary HDAC binding pocket (Figure 1) depending on the ligand. In a recent study Wiest et al²¹ explored flexibility of the HDAC8 binding site using molecular dynamics simulations and confirmed that, indeed, the binding site of HDAC8 is rather malleable and, if required, can be subjected to induced fit conformational changes to accommodate the ligands. In addition, the interpretation of the binding poses may be biased due to the fact that the conformation of the HDAC ligands co-crystallized with HDAC8 is influenced by the interactions between the ligand and neighboring copies of the protein and/or ligand in the crystal lattice as evidenced by analysis of the X-ray structures of HDAC8. When the protein is solvated under physiological conditions the conformation of the SBG may differ from that seen in crystal structures, making the interpretation even more complicated.

Clearly, the problem of modeling the binding poses of HDAC inhibitors remains to be one of the key challenges in HDAC inhibitor structure-based drug design. To address it we

recently developed a Binding Ensemble Profiling with Photoaffinity Labeling (BEPProFL) approach that utilizes a diazide probe to map experimentally the multiple binding poses of the SBGs of HDAC ligands.³² We also designed and synthesized several highly potent cell permeable HDAC isoxazole- and pyrazole-based photoreactive probes.³³ Interestingly, we found that our probe **1b** (Table 1) was able to crosslink to the residues at the bottom of the second pocket in the open conformation of HDAC8 PDB:1T64 mentioned above. We also found that despite the presence of relatively bulky and lipophilic diazide moieties in the SBGs necessary for BEProFL, both the isoxazole- and the pyrazole-based probes exhibited excellent low double-digit nanomolar inhibitory activity against HDAC3 and HDAC8. Having IC₅₀ of 17 nM, the pyrazole-based probe **2a** (Table 1) is one of the most active HDAC8 inhibitors reported to date. On the basis of these data, we hypothesize that even in the case of the ligands (or photoreactive probes) that do not require substantial conformational changes, the ligands and their SBGs have an opportunity to bind to either the closed or the open conformations of the HDAC8 protein, allowing the SBGs of the ligands to minimize the solvent exposure and gain additional interactions with the protein. For docking to HDAC8 it will be a significant departure from the published protocol and may open new directions for design of inhibitors that may target the open instead/in addition to the closed conformation of HDAC8. To our knowledge, with the exception of our publications^{32, 33} and a study by Wiest et al²¹ in all previous docking studies using HDAC8 binding data only receptors based on a crystal structure with closed loop B were used.^{19, 34} Here we present and rationalize a multiple receptor docking and scoring study of a series of novel HDAC8 inhibitors to a set of the closed and open conformations of the HDAC8 protein.

Materials and Methods

Modeling

Coordinates of HDAC8 protein structures were downloaded from the protein data bank³⁵ (PDB). Visual inspection of the crystal structures (PDB IDs 1T64, 1T67, 1T69, 1VKG, 1W22, 2V5W, and 2V5X)^{26, 29–31} was performed using Chimera.³⁶ The reconstruction of the full unit cell was done starting from each PDB file using the python script `crystalcoords.py` in Chimera. The resulting coordinates, containing all copies of the asymmetric unit in the unit cell of each HDAC8 crystal structure, were analyzed with Chimera and in-house awk scripts to establish contacts between the different symmetry copies of the HDAC8 monomer. For re-docking of the native ligands, all protein copies closer than 10 Å to the ligand in the primary binding site were identified. Typically this included only one extra protein chain. The resulting two HDAC monomers were then saved, and used as a single protein for docking.

The five crystal structures (1T64, 1VKG, 1W22, 2V5X, 2V5W) that contained two non-equivalent protein chains in the asymmetric unit were split to yield two different receptor structures. Two structures (1T67 and 1T69) contained only one chain per asymmetric unit. Taken together this provided twelve different receptor structures. Conserved water molecules were identified by splitting the HDAC8 crystal structures into the monomers, aligning and super-positioning them using the Matchmaker tool in Chimera, followed by visual inspection. All water molecules that were found in less than three out of the 12 HDAC8 monomers were discarded resulting in only two conserved clusters of water molecules. The oxygen atoms of the two water clusters identified in this way were saved, and the center of mass of each cluster was used as the oxygen position in all receptor structures used for docking. The receptor set was further extended by adding to each structure either zero (W0), one (W1), or both (W2) of the water molecules. The resulting 36

receptor structures are labeled accordingly, e.g., 1T64B-W2 is the label for the receptor based on chain B of the structure with PDB:1T64 plus both water molecules W1 and W2.

If present, gaps in the sequence of each monomer were filled, and the resulting coordinates subsequently refined, using the homology modeling program *Modeller*.³⁷ Specifically, the structure 2V5X was used as a template for modeling the unresolved amino acids in monomers of HDAC8 with the missing residues. 2V5X was chosen because it has a reasonably good resolution ($R=2.25\text{\AA}$) and has all the residues resolved. The geometry and position of the newly modeled residues were further refined by a Monte Carlo simulation using a built-in in *Modeller* empirical force field. The resulting HDAC8 monomers were protonated using the program *reduce*.³⁸ This program analyses the hydrogen bond network in each monomer and adjusts (flips) GLN, ASN, and HIS side-chains accordingly. Then, the protonation states of the proteins were fixed and the water oxygen atoms were protonated using MOE.³⁹ The ligands were prepared using MOE,³⁹ and for each ligand the hydroxamic acid was chosen to be de-protonated.⁴⁰ For the re-docking experiments the same procedure was followed using either one monomer, or the entire unit cells.

The programs GOLD⁴¹ and Surflex-dock⁴² were used for docking/scoring, and default parameters and settings were used unless mentioned explicitly. For the definition of the binding site the native ligand in each monomer structure was used. Docking was performed with zero, one, or two conserved water molecules as part of the receptor structure. The atoms of the HA group were constrained to their experimental positions using a weight of 5 (GOLD) and 10 (Surflex-dock). In preliminary studies it was found that these values for the force constants give poses with comparable HA geometries.

Some of the compounds used here contain azide groups, and preliminary studies with Surflex-dock showed that the DREIDING force field⁴³ employed by this program was unable to reproduce the angle of nearly 180 degrees formed by the three azide nitrogen atoms.⁴⁴ Since Surflex-dock does not allow for editing the force field parameters used we decided to replace the azide $N=N=N$ by a sufficiently similar residue for the docking calculations. As, in each case, the magnitude of AM1BCC partial charges⁴⁵ calculated with molcharge⁴⁶ were lower than 0.3 for the two terminal nitrogen atoms of the azides we concluded that this residue is best modeled as a hydrophobic moiety. Given this, the approximate nature of a scoring function, and the similar van der Waals radii of the two residues we concluded that an $N=C=CH_2$ group, which does give the correct linear geometry,⁴⁴ can be used as an analog mimicking the azido group.

For docking with Surflex-dock the *pgeomx* flag was used to ensure a more exhaustive search, resulting in Gold and Surflex-dock using approximately the same amount of CPU-time per compound. For each ligand/receptor combination 20 poses were generated.

All ligands docked with Surflex-dock were processed by deleting the HA group and four of the six CH_2 units of the linker connecting ZBG and SBG. This task was automated using in-house awk scripts and programs based on the openbabel C-library, version 2.2.⁴⁷ For re-scoring of the resulting fragments with the HA group removed the same settings were used as in the original docking runs. This protocol is referred to as SC2, the protocol scoring the entire ligand as SC6. Additional calculations were performed using fragments including zero (SC0) or four (SC4) CH_2 units.

Multiple receptor docking was done by docking to each possible pair of receptor structures and using the better of two scores for each compound for the final ranking. The same docking procedure was used for the full ligands and the fragments. In addition, the scores calculated for the ligands and ligand fragments with each receptor structure were normalized. For this purpose a histogram was generated from all the calculated scores from

each receptor structure, a normal distribution curve was fitted to each histogram using the statistics package R,⁴⁸ and all individual scores were normalized by subtracting the mean, and dividing by the standard deviation.

HDAC8 Inhibitors and assays

The compounds considered here – **1c**, **1d**,⁴⁹ **2a–d**, **3a,b**,³³ **1b,e**,³² **4a,b**⁴⁹ – represent several series of new HDAC inhibitors and probes (Table 1) that were synthesized by us as a part of our project to develop photoreactive probes for BEProFL.^{32, 33}

The inhibition of HDAC8 was measured as recommended by the supplier BIOMOL International using the fluorescent acetylated HDAC substrate *Fluor de Lys* (BIOMOL, KI178) and commercially available recombinant human HDAC8 (BIOMOL). The activity data are summarized in Table 1. The procedure is identical to that published previously.³²

Results and discussion

Analysis of co-crystal structures

An analysis of the HDAC8 crystal structures was performed to identify a set of proteins and water molecules appropriate for the docking. A total of 12 wild-type HDAC8 structures co-crystallized with a hydroxamic acid based inhibitor were available at the time this study was performed. A visual inspection of the aligned monomers of HDAC8 with resolved water molecules (Figure 2C) revealed the presence of one water molecule (W1) that forms a hydrogen bond with His 180 and an acceptor atom, typically a carboxamide oxygen atom, in the ligand. This interaction pattern was found in nine of the 11 monomers. Another water molecule (W2) forms H-bonds with W1, Phe208, and, in four cases, also with a bound ligand. The RMSD of the water oxygen atoms in the W1 cluster is 0.49 Å, for W2 this value is 0.37 Å. Since inclusion of conserved water molecules may improve accuracy of the docking⁵⁰ both W1 and W2 were considered during docking.

To further investigate the influence of crystal packing and ligands' structure on the poses of SBGs and conformation of the protein, the coordinates of the full unit cell were regenerated from each of the PDB files, and interactions between the ligands and the protein residues in its vicinity were analyzed by visual inspection. In all cases it was found that the ligands' SBG not only interacts with the HDAC8 monomer it is bound to, but also with residues and/or ligands of neighboring copies of the protein in the crystal lattice (Figure 2A). It is unclear whether the pose and the resulting protein-ligand interactions of a given compounds SBG would also be found under physiological conditions, i.e. in solution, or whether this pose is, at least in part, determined by interactions between the SBG and symmetry copies of the primary protein.

The generation and visual inspection of the full unit cells of HDAC8 crystal structures also highlighted another phenomenon that can possibly complicate the interpretation of co-crystal structures in the context of the drug design: in nearly all cases the smallest distance between protein residues in loop B and any atom found in one of the symmetry copies of the protein in the crystal lattice is well below 4Å. In all structures that have a closed loop B (Figure 1) at least one hydrogen bond is formed between loop B and an atom in another HDAC8 monomer (see Table S1 in Supplemental material). Interestingly the only structure in which no atom in loop B is in contact with other proteins in the crystal lattice is 1T64, which is one of the two structures with an open loop B. The two chains, A and B, in the asymmetric unit of 1T64 (loop B open) were aligned on the corresponding chains in the full unit cell of 1T69 (loop B closed) and are shown in Figure 2B. We find that the two copies of loop B of the 1T64 chains overlap, the closest distance between two non-H atoms being 1.2 Å, while for the PDB:1T69 structure this distance is 6.6 Å. The data indicate that the open

and closed conformations observed in the X-ray structures may be a combined result of the conformational changes caused by a specific ligand bound and interactions between the symmetry copies. Thus, neither of the conformations can be excluded or prioritized only on the basis of the structure of the ligands.

Re-docking of native ligands

The analysis of the crystal structures of HDAC8 provided above shows that the second copy of the HDAC8 protein may affect the conformation of the HDAC8 protein and the bound ligands, and therefore we decided to determine how this would affect the results of the re-docking experiments.

HDAC inhibitors were re-docked to their native receptors, comparing the docking accuracy achieved with two different docking protocols/scoring functions. All the available at the time of the writing this paper co-crystal structures of hydroxamic acid based HDAC inhibitors were used for this purpose. If more than one chain was present in the asymmetric unit the one with the lower average B-factor was used. Given the results of the previous section and to account for the effect of interactions between the ligands SBG and symmetry copies of the HDAC8 monomers, as well as the water molecules W1 and W2 found in the majority of cases, receptor structures were set up in three different ways: i) one HDAC8 monomer (chainA); ii) as (i) plus W1 (chainA+W1); iii) as (i) plus W1 and W2 (chainA+W1+W2); iv) one HDAC monomer plus all symmetry copies of protein and bound ligand that can possibly come into contact with the docked ligands SBG (x-tal); v) as (iv) plus the conserved water molecule W1 (x-tal+W1); vi) as (iv) plus the conserved water molecules W1 and W2 (x-tal+W1+W2). In preliminary docking runs with GOLD and Surflex-Dock it was found that in some cases the best-scored docking pose corresponded to a structure with the ligands ZBG far from the protein's Zn^{2+} ion (not shown). Therefore constraints were applied in all docking runs to ensure that the ZBG coordinates are close to their experimental values.

The results of the docking for five ligands with different protocols are shown in Table 2. Analysis of the poses obtained for 2V5X shows that the docking software tends to position the ligand such that its two indol substituents switch positions compared to those in the crystal structure. The ligand is somewhat symmetrical and can occupy almost the same space in the binding site in both the cases, making prediction of the docking poses a rather challenging task. Given the exceptionally large and flexible SBG found in 2V5X, even values 3.70 Å and above can still be considered acceptable. To minimize bias due to the unusually high RMSD values obtained for 2V5X we excluded 2V5X from calculation of the average RMSD values given in Table 2 and discussion below.

The RMSD values obtained with Surflex-dock are smaller if the receptor structure is constructed in the most accurate way, i.e. it includes symmetry copies of the HDAC8 protein directly adjacent to the binding site. A similar trend is observed for re-docking with GOLD though it is somewhat less pronounced. Re-dockings to the primary monomer structure of HDAC8 without a symmetry copy and the conserved water molecules resulted in the average RMSD 2.65 Å smaller for GOLD compared to Surflex (Table 2). Since the SBG of the ligands co-crystallized with HDAC8 are in the vicinity of a symmetry copy of another HDAC8 it is unclear why GOLD was able to generate poses with smaller RMSD and Surflex not. In case of re-docking to a monomer of HDAC8, inclusion of water molecules W1 and W2 leads to a decrease of an average RMSD for Surflex by ca 2 Å, whereas for GOLD the average RMSD increases by ca 1 Å. In case of "x-tal" receptors inclusion of water molecules affects the accuracy only marginally. We note that the best RMSD values obtained for 1VKG with particularly Surflex are small, confirming that the modeled structure of loop A in 1VKG is unlikely to introduce artifacts.

In one case (1VKG) Surflex dock gives a clearly better result than GOLD, but in all other cases the two docking programs and scoring functions used here show a comparably good performance in reproducing experimental poses in their native “x-tal”, “x-tal+W1”, and “x-tal+W1+W2” receptors - the entire unit cell with all symmetry copies of the primary monomer.

Since generally re-docking to several open and closed HDAC8 crystal structures containing the second copy of the HDAC8 protein resulted in improved RMSD values, neither of the crystal structure should be excluded based on the re-docking accuracy and we elected to use a multiple receptor docking protocol. Inclusion of multiple receptor structures in docking to account for large variations in protein conformation has become a commonly applied strategy.^{51–55} Despite the better accuracy observed for the “x-tal”-based receptors we could not use them for docking ligands other than the native ones, and, therefore, all the following studies were performed only with the monomers of HDAC8.

Docking to single receptors

Receptor structures based on 12 different HDAC8 monomers were prepared as described in Methods, and zero, one, or two water molecules were added to each structure, resulting in 36 receptor structures for docking. The compounds in Table 1 were docked to each of the receptor structures using two different docking programs, Surflex-dock and GOLD. Constraints on the hydroxamic acid group and other parameters were used as described in Methods. The pK_i values for all compounds were compared to the calculated docking scores, resulting in 36 correlations (Pearson correlations) for, each, the GOLD (not shown) and the Surflex-dock (Figure 3 Left) results. The Pearson correlations were used to differentiate between the positive and negative correlation. The correlations show a clear positive trend. Both the average and the best (average $r=0.17$, best $r=0.43$) correlations obtained from the GOLD results are somewhat poorer than those obtained with Surflex-dock (average $r=0.23$, best $r=0.67$). Taking into account that Surflex-dock was also successful in re-docking native HDAC ligands and contains a desolvation term we decided to perform the calculations discussed below only based on the Surflex-dock results. In the following we employ two techniques - a variation of fragment based scoring and multiple receptor based docking - to evaluate their effect on the results of docking/scoring calculations.

Fragments

To avoid difficulties with the parameterization Zn^{2+} ion in the HDAC active site we decided to explore how truncation of the ZBG would affect the docking. Specifically, the docking poses generated as described in the previous section were processed by removing the hydroxamic acid moiety and a part of the linker from each docked ligand and the resulting fragments were re-scored. Here we make the assumption that the contribution of the Zn^{2+} -ZBG interactions is similar for all ligands, and that these contributions cancel out to a good degree. The best correlation obtained from fragment based scores (SC2) is $r=0.74$ for receptor structure 1VKGB-W0. For the full ligands (SC6) the best correlation was weaker, with $r=0.67$ for receptor structure 2V5WB-W0. In both cases SC2 and SC6 the best correlations were substantially higher than the average correlations of $r\sim 0.2$ (Table 4). We also performed equivalent calculations for ligands that were truncated to leave fragments with four (SC4) or zero (SC0) instead of two (SC2) CH_2 units of the linker. The correlations obtained with SC0 were somewhat weaker, and those with SC4 were comparable to the correlations from the SC2 protocol. The top two receptor structures, in terms of the resulting correlations, were identical for SC2 and SC4 (2V5XA-W2 and 1VKGB-W0). The correlations obtained with the three protocols SC0, SC2, and SC4, for all receptor structures, are similar, and show good correlations of $r>0.8$ between each other, while the same kind of correlations between SC6 and the three other protocols is smaller than $r=0.7$ in all cases.

We found that the truncated form of this type of HDAC8 inhibitors can be used instead of the full ligands, leading to a slight improvement in correlation between the docking scores and the experimental affinities. The relative scores were also not very sensitive to the precise choice of the truncation. This is particularly important as the ionization state of the ligands and residues^{56, 57} and the identity of the metal ion in the HDAC active site⁵⁸ have been a matter of debate. In the following we only discuss the results obtained with the SC6 and the SC2 protocols.

Multiple receptors

The docking scores obtained with each compound/receptor combination were chosen from all possible pairs of receptor structures, using the best out of two scores for each compound. This was done independently for the results from both the SC2 and SC6 protocol using the normalized scores. A summary of the results is given in Figure 3B and Table 4. The best correlations achieved with two receptors were 0.72 for SC6 and 0.80 for SC2. With SC6 two out of 630 combinations give correlations better than $r=0.70$. For SC2 this applies to 11 out of 630 combinations. On average the resulting 630 new correlations for each of the two scoring protocols considered did not improve substantially. However, in a majority of cases the correlations are above zero, and, compared to the single receptor-SC6 protocol, we find more cases in which the correlations are particularly good with $r>0.7$. The best correlation of $r=0.80$ is obtained with the SC2 protocol by combining open 1VKGB-0W and closed 2V5XA-2W receptor structures (Table 4 and Figure 4A).

We also compared the results obtained by choosing from all possible combinations of three receptor structures and choosing from all the 36 receptors. The latter gave a correlation close to zero ($r=0.02$), while the best correlation obtained with a combination of three receptors is $r=0.83$, only a marginal improvement over the best correlations with two receptors. Therefore, and in order to keep results more interpretable, we decided to only consider results from pairs of receptors.

Given the large number of combinations of receptors we consider here, it was important to understand to what extent the good correlations are fortuitous. Thus, in the next sections, we evaluated a test set of HDAC8 inhibitors generated by a different laboratory and performed an analysis of the trends observed with the open and closed conformations of HDAC8 for the training and test sets.

Test

Our dataset of HDAC8 inhibitors has the largest published range of HDAC8 activities among SAHA-like HDAC8 inhibitors. Although SAHA, and SAHA derivatives, belong to the most widely discussed HDAC inhibitors, the available affinity data generated using the same biological protocol in the same laboratory for the compounds based on this scaffold appear to be rather limited. We decided to use a subset of compounds published by Chen et al.⁵⁹ The data in this paper include IC_{50} values of 17 hydroxamic acid based HDAC8 inhibitors with flexible linkers as used in the training set. Since the activities of these compounds only cover one order of magnitude we decided to only use a subset of six compounds, those with the three lowest, and those with the three highest IC_{50} values, resulting in two groups of compounds that are separated by about one log-unit in their activities. The docking scores were then analyzed using a confusion matrix of true and false positives (TP, FP), and true and false negatives (TN, FN) respectively. The ability of the scores to categorize the ligands into weaker or stronger binders was calculated as accuracy $a=(TP+TN)/(TP+TN+FP+FN)$. A value of $a=0.5$ implying random, and a value of $a=1.0$ implying perfect categorization. Using the protocol that gave the best results for our training set, i.e., the best of two fragment based scores from the receptor structures 1VKGB-W0 and

2V5XA-W2 we obtain an accuracy of $a=0.67$. A clearer, and statistically more significant, picture emerges if one considers the entire range of results obtained with all possible receptor combinations. Of all the calculated accuracies 93% lie above, and only 7% at, or below, $a=0.5$. If the docking protocol used here was unable to correctly categorize compounds with weaker or stronger activities these numbers would reside around 50%.

Trends

To reduce the number of cases taken into account when using pairs of receptors and to determine the difference in trends observed with whether open and closed structures of HDAC8, we split the 630 combinations of structures into three groups: one containing only pairs of structures with open loop-B (O-O), one with a combination of one open and one closed structure (O-C), and a third one with pairs of two closed structures (C-C). The data were analyzed by calculating, for each group, the percentage of receptor pairs that give a correlation above a certain threshold for the compounds in the training set. With this threshold being -1 this percentage is simply the total amount of receptor combinations in each group divided by the number of all combinations: about 10, 44, and 46% for O-O, C-C, and O-C, respectively. If the threshold is raised the sum of the three percentages stays 100%, but the three numbers reflect the relative propensity of each group for giving increasingly good correlations. As the threshold gets higher (Figure 5), there is a clear trend towards better correlations for O-C and O-O cases, whereas the C-C combinations of receptor structures fall behind, and at correlations with $r \geq 0.7$ 100% of all cases are either from the O-C or the O-O groups.

A similar analysis was performed with the test set compounds. Figure 6 shows the accuracies achieved with all possible combinations of receptor structures for the test set in a 2D diagram. Each point represents a particular combination of receptors, and is colored according to the resulting accuracy, from zero (white) to one (black). In this diagram the four left columns and the four bottom rows represent open structures, while the remaining columns and rows represent closed structures. The diagram clearly demonstrates a trend for open structures 1VKGB, 1T64 alone and in combination with closed 1T67, 1T69, 1W22, 2V5X, 2V5W to provide better accuracies than for combinations of closed 1T67, 1T69, 1W22, 2V5X, 2V5W structures alone. The average accuracy obtained with O-O combinations is 0.96, for O-C it is 0.83, and for C-C 0.71.

The receptor combination 1VKGB-0W and 2V5XA-2W found to give the best correlation for the training set does give an accuracy of above $a=0.5$ also for the test set, but there are other combinations of receptors that give better results for the test set. We believe that the reason for this is that the test set only contains compounds with large and flexible SBGs, while the training set includes a number of smaller compounds, such as SAHA. Interestingly, we see that if we perform multiple receptor scoring for the test set compounds with combinations of two different open receptors only (Figure 6), then 60 out of the possible 66 combinations gives an accuracy of $a=1.0$, meaning that all six compounds are correctly assigned to one of the groups of weaker or stronger activities.

More interesting than correlations from particular docking results is the general trend we observe for correlations. They are noticeably improved if receptors based on HDAC8 structures with an open loop B are included into a set of multiple receptors. It may suggest that the SBGs of at least some of the ligands may not necessarily be solvent exposed but instead may be hidden in the open conformations of the HDAC8 protein. Moreover, availability of both the closed and the open conformations would be consistent with the photolabeling results obtained by us with probe **1b** that was found to photocrosslink to both the residues on the surface and to the residues at the bottom of the second binding site.³²

The possibility of a dynamical equilibrium between the open, closed, and other conformations of the protein was recently highlighted by Wiest et al.²¹

These results suggest that depending on the geometry and the size of the SBG the ligands may preferentially bind to an open structure, which gives them an opportunity to bury their, perhaps, bulky and largely hydrophobic SBG, or to a closed conformation of the protein if their SBG are small or cannot be accommodated by the open conformation. Indeed, we find that, with the best protocol used here, inhibitors **1c** and **1d**, two of the three compounds with the smallest SBGs, do give a better score when docked to the closed 2V5XA-W2 conformation, compared to the scores obtained with the open 1VKGB-W0 conformation. Also, nine out of the ten larger compounds give a better score when docked to the open structure. The best scoring poses of compounds **2a**, **2b**, **2d**, and **4b**, four out of the five compounds with pK_i values above 7, have their SBG buried in the second binding pocket. This would also be consistent with the SAR we recently presented for the pyrazole- and isoxazole-based photoreactive ligands/probes³³ **2a-d** and **3a,b**, respectively. The SBG of neither **3a** nor **3b** are small or flexible enough to fit in the second pocket, and, therefore, they occupy one of the shallow grooves on the protein surface. Unlike the rigid SBG of the isoxazole-based ligands in series **3**, the SBG of the pyrazole-based ligands is smaller and more flexible. Accordingly, the majority of the pyrazole-based ligands binds to the open conformation of HDAC8 (Figure 7, only **2a** is shown for clarity) where they have an opportunity to hide their SBG in the second binding site.

Summary

Availability of the second binding site in the open conformation of HDAC8 for binding of SAHA-like ligands was explored using a multiple-receptor fragment-based docking. Analysis of the crystallographic structures of the HDAC8 protein indicated that the second copy of the HDAC8 protein present in the crystal lattice may affect the conformation of both the residues in the binding site and its vicinity and the binding pose of the ligand, rendering inclusion of the second copy of the HDAC8 protein during re-docking important. Re-docking of the native ligands confirmed that both Surflex and GOLD performed comparably well. The RMSD values for poses generated by both Surflex and GOLD were found to be particularly small when a monomer of HDAC8 was used in combination with the symmetry copy of HDAC8. Inclusion of the conserved water molecules during re-docking affected the accuracy of the docking for some of the x-ray structures but not the others. The average RMSD obtained for re-docking of the native ligands to the HDAC8 monomers without the corresponding symmetry copies was somewhat smaller for GOLD than for Surflex.

Neither of the open or closed conformations of the HDAC8 protein was found to be superior and produced comparable RMSD in the re-docking experiments, indicating that neither of the open or the closed conformations can be excluded from consideration. To facilitate handling of the HDAC8-Zn²⁺-ligand complex by the docking forcefield the ligands were truncated at various positions of the flexible aliphatic linker. The fragment-based docking scores were shown to be relatively insensitive to the position of the truncation. A total of 630 receptor structures containing six open and closed structures of HDAC8 and up to two conserved water molecules were used to for the docking of 13 compounds in the training set and six compounds in the test set. We found that there is a clear trend towards better correlations between pK_i against HDAC8 and the docking scores for O-C and O-O cases. At correlations with $r \geq 0.7$ 100% of all cases are either from the O-C or the O-O groups. A similar trend is observed for a test set compounds docking to open structures 1VKG, 1T64 alone and in combination with closed 1T67, 1T69, 1W22, 2V5X, 2V5W resulted in overall more accurate differentiation between weaker and stronger binders than docking to closed 1T67, 1T69, 1W22, 2V5X, 2V5W structures alone. This trend is also consistent with the

SAR of a series of potent HDAC inhibitors recently published by us. Unlike the SBGs of the isoxazole-based ligands, the SBG of the pyrazole-based ligands is flexible enough to occupy the second binding pocket found in the open conformations of HDAC8. Further exploration of the novel ligand scaffolds that can utilize the second binding site of HDAC8 is underway in our laboratory.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was in part funded by the National Cancer Institute/NIH Grant R01 CA131970 and Alzheimer's Drug Discovery Foundation grant 20101103. We also thank Ajay Jane for providing a free academic version of Surflex-dock. Molecular modeling was in part conducted using free academic licenses for the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH Grant P41 RR-01081) and OpenEye Scientific Software, Santa Fe, NM.

References

1. Rundlett SE, Carmen AA, Kobayashi R, Bavykin S, Turner BM, Grunstein M. HDA1 and RPD3 are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription. *PNAS*. 1996; 93(25):14503–14508. [PubMed: 8962081]
2. Kouzarides T. Acetylation: a regulatory modification to rival phosphorylation? *EMBO J*. 2000; 19(6):1176–1179. [PubMed: 10716917]
3. Villar-Garea A, Esteller M. Histone deacetylase inhibitors: understanding a new wave of anticancer agents. *Int J Cancer*. 2004; 112(2):171–178. [PubMed: 15352027]
4. Balasubramanian S, Verner E, Buggy JJ. Isoform-specific histone deacetylase inhibitors: The next step? *Cancer Letters*. 2009; 280(2):211–221. [PubMed: 19289255]
5. Khan N, Jeffers M, Kumar S, Hackett C, Boldog F, Khramtsov N, Qian X, Mills E, Berghs SC, Carey N, Finn PW, Collins LS, Tumber A, Ritchie JW, Jensen PB, Lichenstein HS, Sehested M. Determination of the class and isoform selectivity of small-molecule histone deacetylase inhibitors. *Biochem J*. 2008; 409(2):581–589. [PubMed: 17868033]
6. Chen YD, Jiang YJ, Zhou JW, Yu QS, You QD. Identification of ligand features essential for HDACs inhibitors by pharmacophore modeling. *J Mol Graphics Modell*. 2008; 26(7):1160–1168.
7. Guo Y, Xiao J, Guo Z, Chu F, Cheng Y, Wu S. Exploration of a binding mode of indole amide analogues as potent histone deacetylase inhibitors and 3D-QSAR analyses. *Bioorg Med Chem*. 2005; 13(18):5424–5434. [PubMed: 15963726]
8. Ragno R, Simeoni S, Rotili D, Caroli A, Botta G, Brosch G, Massa S, Mai A. Class II-selective histone deacetylase inhibitors. Part 2: alignment-independent GRIND 3-D QSAR, homology and docking studies. *Eur J Med Chem*. 2008; 43(3):621–632. [PubMed: 17698257]
9. Ragno R, Simeoni S, Valente S, Massa S, Mai A. 3-D QSAR studies on histone deacetylase inhibitors. A GOLPE/GRID approach on different series of compounds. *J Chem Info Model*. 2006; 46(3):1420–1430.
10. Zhu Y, Li HF, Lu S, Zheng YX, Wu Z, Tang WF, Zhou X, Lu T. Investigation on the isoform selectivity of histone deacetylase inhibitors using chemical feature based pharmacophore and docking approaches. *Eur J Med Chem*. 2010; 45(5):1777–1791. [PubMed: 20153566]
11. Tang H, Wang XS, Huang XP, Roth BL, Butler KV, Kozikowski AP, Jung M, Tropsha A. Novel inhibitors of human histone deacetylase (HDAC) identified by QSAR modeling of known inhibitors, virtual screening, and experimental validation. *J Chem Info Model*. 2009; 49(2):461–476.
12. Di Micco S, Terracciano S, Bruno I, Rodriguez M, Riccio R, Taddei M, Bifulco G. Molecular modeling studies toward the structural optimization of new cyclopeptide-based HDAC inhibitors modeled on the natural product FR235222. *Bioorg Med Chem*. 2008; 16(18):8635–8642. [PubMed: 18715788]

13. Grolla AA, Podesta V, Chini MG, Di Micco S, Vallario A, Genazzani AA, Canonico PL, Bifulco G, Tron GC, Sorba G, Pirali T. Synthesis, biological evaluation, and molecular docking of Ugi products containing a zinc-chelating moiety as novel inhibitors of histone deacetylases. *J Med Chem.* 2009; 52(9):2776–2785. [PubMed: 19344175]
14. Huang WJ, Chen CC, Chao SW, Lee SS, Hsu FL, Lu YL, Hung MF, Chang CI. Synthesis of N-Hydroxycinnamides Capped with a Naturally Occurring Moiety as Inhibitors of Histone Deacetylase. *ChemMedChem.* 2010; 5(4):598–607. [PubMed: 20209563]
15. Lu Q, Wang DS, Chen CS, Hu YD. Structure-based optimization of phenylbutyrate-derived histone deacetylase inhibitors. *J Med Chem.* 2005; 48(17):5530–5535. [PubMed: 16107152]
16. Mai A, Valente S, Nebbioso A, Simeoni S, Ragno R, Massa S, Brosch G, De Bellis F, Manzo F, Altucci L. New pyrrole-based histone deacetylase inhibitors: binding mode, enzyme- and cell-based investigations. *Int J Biochem Cell Biol.* 2009; 41(1):235–247. [PubMed: 18834955]
17. Pirali T, Faccio V, Mossetti R, Grolla AA, Di Micco S, Bifulco G, Genazzani AA, Tron GC. Synthesis, molecular docking and biological evaluation as HDAC inhibitors of cyclopeptide mimetics by a tandem three-component reaction and intramolecular [3+2] cycloaddition. *Mol Divers.* 2010; 14(1):109–121. [PubMed: 19475493]
18. Wang DF, Wiest O, Helquist P, Lan-Hargest HY, Wiech NL. On the function of the 14 Å long internal cavity of histone deacetylase-like protein: implications for the design of histone deacetylase inhibitors. *J Med Chem.* 2004; 47(13):3409–3417. [PubMed: 15189037]
19. Ortore G, Di CF, Martinelli A. Docking of Hydroxamic Acids into HDAC1 and HDAC8: A Rationalization of Activity Trends and Selectivities. *J Chem Info Model.* 2009; 49(12):2774–2785.
20. Wang DF, Helquist P, Wiech NL, Wiest O. Toward selective histone deacetylase inhibitor design: homology modeling, docking studies, and molecular dynamics simulations of human class I histone deacetylases. *J Med Chem.* 2005; 48(22):6936–6947. [PubMed: 16250652]
21. Estiu G, West N, Mazitschek R, Greenberg E, Bradner JE, Wiest O. On the inhibition of histone deacetylase 8. *Bioorg Med Chem.* 2010; 18(11):4103–4110. [PubMed: 20472442]
22. Raha K, Merz KM Jr. A quantum mechanics-based scoring function: study of zinc ion-mediated ligand binding. *J Am Chem Soc.* 2004; 126(4):1020–1021. [PubMed: 14746460]
23. Irwin JJ, Raushel FM, Shoichet BK. Virtual screening against metalloenzymes for inhibitors and substrates. *Biochemistry.* 2005; 44(37):12316–12328. [PubMed: 16156645]
24. Park H, Lee S. Homology modeling, force field design, and free energy simulation studies to optimize the activities of histone deacetylase inhibitors. *J Comput Aided Mol Des.* 2004; 18(6):375–388. [PubMed: 15662999]
25. Ryde U. Molecular dynamics simulations of alcohol dehydrogenase with a four- or five-coordinate catalytic zinc ion. *Proteins.* 1995; 21(1):40–56. [PubMed: 7716168]
26. Somoza JR, Skene RJ, Katz BA, Mol C, Ho JD, Jennings AJ, Luong C, Arvai A, Buggy JJ, Chi E, Tang J, Sang BC, Verner E, Wynands R, Leahy EM, Dougan DR, Snell G, Navre M, Knuth MW, Swanson RV, McRee DE, Tari LW. Structural snapshots of human HDAC8 provide insights into the class I histone deacetylases. *Structure.* 2004; 12(7):1325–1334. [PubMed: 15242608]
27. Sherman W, Day T, Jacobson MP, Friesner RA, Farid R. Novel procedure for modeling ligand/receptor induced fit effects. *J Med Chem.* 2006; 49(2):534–553. [PubMed: 16420040]
28. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved protein-ligand docking using GOLD. *Proteins: Struct, Funct and Gen.* 2003; 52(4):609–623.
29. Dowling DP, Gantt SL, Gattis SG, Fierke CA, Christianson DW. Structural studies of human histone deacetylase 8 and its site-specific variants complexed with substrate and inhibitors. *Biochemistry.* 2008; 47(51):13554–13563. [PubMed: 19053282]
30. Vannini A, Volpari C, Filocamo G, Casavola EC, Brunetti M, Renzoni D, Chakravarty P, Paolini C, De Francesco R, Gallinari P, Steinkühler C, Di Marco S. Crystal structure of a eukaryotic zinc-dependent histone deacetylase, human HDAC8, complexed with a hydroxamic acid inhibitor. *PNAS.* 2004; 101(42):15064–15069. [PubMed: 15477595]
31. Vannini A, Volpari C, Gallinari P, Jones P, Mattu M, Carfi A, De Francesco R, Steinkühler C, Di Marco S. Substrate binding to histone deacetylases as shown by the crystal structure of the HDAC8-substrate complex. *EMBO Rep.* 2007; 8(9):879–884. [PubMed: 17721440]

32. He B, Velaparthi S, Pieffet G, Pennington C, Mahesh A, Holzle DL, Brunsteiner M, van Breemen R, Blond SY, Petukhov PA. Binding ensemble profiling with photoaffinity labeling (BEProFL) approach: mapping the binding poses of HDAC8 inhibitors. *J Med Chem.* 2009; 52(22):7003–7013. [PubMed: 19886628]
33. Neelarapu R, Holzle DL, Velaparthi S, Bai H, Brunsteiner M, Blond SY, Petukhov PA. Design, Synthesis, Docking, and Biological Evaluation of Novel Diazide-Containing Isoxazole- and Pyrazole-Based Histone Deacetylase Probes. *J Med Chem.* 2011; 54(13):4350–4364. [PubMed: 21548582]
34. Bora-Tatar G, Dayangac-Erden D, Demir AS, Dalkara S, Yelekci K, Erdem-Yurter H. Molecular modifications on carboxylic acid derivatives as potent histone deacetylase inhibitors: Activity and docking studies. *Bioorg Med Chem.* 2009; 17(14):5219–5228. [PubMed: 19520580]
35. Kirchmair J, Markt P, Distinto S, Schuster D, Spitzer GM, Liedl KR, Langer T, Wolber G. The Protein Data Bank (PDB), Its Related Services and Software Tools as Key Components for In Silico Guided Drug Discovery. *J Med Chem.* 2008; 51(22):7021–7040. [PubMed: 18975926]
36. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem.* 2004; 25(13):1605–1612. [PubMed: 15264254]
37. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A. Comparative protein structure modeling using MODELLER. *Curr Protoc Prot Sci.* 2007; Chapter 2(Unit 2 9)
38. Word JM, Lovell SC, Richardson JS, Richardson DC. Asparagine and glutamine: Using hydrogen atom contacts in the choice of side-chain amide orientation. *J Mol Biol.* 1999; 285(4):1735–1747. [PubMed: 9917408]
39. Molecular Operating Environment. <http://www.chemcomp.com>
40. Vanommeslaeghe K, Loverix S, Geerlings P, Tourwe D. DFT-based ranking of zinc-binding groups in histone deacetylase inhibitors. *Bioorg Med Chem.* 2005; 13(21):6070–6082.
41. Jones G, Willett P, Glen RC. Molecular Recognition of Receptor-Sites Using a Genetic Algorithm with a Description of Desolvation. *J Mol Biol.* 1995; 245(1):43–53. [PubMed: 7823319]
42. Jain AN. Surflex-Dock 2.1: Robust performance from ligand energetic modeling, ring flexibility, and knowledge-based search. *J Comput Aided Mol Des.* 2007; 21(5):281–306. [PubMed: 17387436]
43. Mayo SL, Olafson BD, Goddard WA. Dreiding - a Generic Force-Field for Molecular Simulations. *J Phys Chem.* 1990; 94(26):8897–8909.
44. Pieffet G, Petukhov PA. Parameterization of aromatic azido groups: application as photoaffinity probes in molecular dynamics studies. *J Mol Model.* 2009; 15(11):1291–1297. [PubMed: 19288146]
45. Jakalian A, Jack DB, Bayly CI. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem.* 2002; 23(16):1623–1641. [PubMed: 12395429]
46. molcharge. Openeye Scientific Software; Santa Fe: 2010.
47. Guha R, Howard MT, Hutchison GR, Murray-Rust P, Rzepa H, Steinbeck C, Wegner J, Willighagen EL. The Blue Obelisk-interoperability in chemical informatics. *J Chem Info Model.* 2006; 46(3):991–998.
48. R: A Language and Environment for Statistical Computing. Vienna: 2009.
49. Vaidya AS, Abdelkarim H, Neelarapu R, Veleparthi S, Brunsteiner M, Bai H, Blond SY, Petukhov Pavel A. 2011 unpublished.
50. Huang N, Shoichet BK. Exploiting ordered waters in molecular docking. *J Med Chem.* 2008; 51(16):4862–4865. [PubMed: 18680357]
51. Barril X, Fradera X. Incorporating Protein Flexibility into Docking and Structure-based Drug Design. *Expert Opinion in Drug Discovery.* 2006; 1(4):1–14.
52. Rueda M, Bottegoni G, Abagyan R. Recipes for the selection of experimental protein conformations for virtual screening. *J Chem Info Model.* 2010; 50(1):186–193.

53. Yoon S, Welsh WJ. Identification of a minimal subset of receptor conformations for improved multiple conformation docking and two-step scoring. *J Chem Inf Comput Sci*. 2004; 44(1):88–96. [PubMed: 14741014]
54. Zhong H, Tran LM, Stang JL. Induced-fit docking studies of the active and inactive states of protein tyrosine kinases. *J Mol Graphics Modell*. 2009; 28(4):336–346.
55. Barril X, Morley SD. Unveiling the full potential of flexible receptor docking using multiple crystallographic structures. *J Med Chem*. 2005; 48(13):4432–4443. [PubMed: 15974595]
56. Wu R, Lu Z, Cao Z, Zhang Y. Zinc chelation with hydroxamate in histone deacetylases modulated by water access to the linker binding channel. *J Am Chem Soc*. 2011; 133(16):6110–6113. [PubMed: 21456530]
57. Wang D, Helquist P, Wiest O. Zinc binding in HDAC inhibitors: a DFT study. *J Org Chem*. 2007; 72(14):5446–5449. [PubMed: 17579460]
58. Gantt SL, Gattis SG, Fierke CA. Catalytic activity and inhibition of human histone deacetylase 8 is dependent on the identity of the active site metal ion. *Biochemistry*. 2006; 45(19):6170–6178. [PubMed: 16681389]
59. Chen Y, Lopez-Sanchez M, Savoy DN, Billadeau DD, Dow GS, Kozikowski AP. A series of potent and selective, triazolylphenyl-based histone deacetylases inhibitors with activity against pancreatic cancer cells and *Plasmodium falciparum*. *J Med Chem*. 2008; 51(12):3437–3448. [PubMed: 18494463]

Figure 1.

Twelve aligned and superimposed HDAC8 monomers. The backbones corresponding to Loop B are highlighted in red (closed) or blue (open). Also shown are the TSA ligands in PDB:1T64; one TSA occupies the primary pocket, interacting with the zinc (cyan), the other occupies the secondary pocket which is available only if loop B is open. Asp 101, representing the tip of loop A is shown in green.

Figure 2.

Interactions between HDAC8 protein/ligand and symmetry copies in co-crystal structures, and visualization of conserved water molecules. A: Example for interactions between a ligands SBG (blue) bound to one monomer (green) with a symmetry copy in the crystal lattice (yellow); B: Section of the unit cell content of PDB:1T64 (green) aligned on PDB:1T69 (yellow). The segments of the sequences shown represent the tip of loop-B, which is open in 1T64 and closed in 1T69. The closest distance between any two atoms in the two symmetry copies of the monomer are indicated for both structures. C: Aligned HDAC8 co-crystal structures (ligands yellow, protein white) revealing the presence of two conserved water molecules (red spheres).

Figure 3.

(A) Histograms of the correlations between experimental affinities and scores from docking to single receptors and SC6 scoring the entire ligand. (B) The same as A for multiple receptor docking protocol with a combination of all pair of HDAC8 receptors and SC2 scoring protocol.

Figure 4. Correlation between measured and predicted activities for the training set. Results are from multiple receptor docking using the best out of two scores from docking to two receptor structures: 1VKGB-0W (open spheres) and 2V5XA-2W (triangles) with normalized fragment-based scores.

Figure 5.

A comparison of the percentages of all correlations between measured and predicted activities for the training set above a threshold (given in the x-axis) as obtained with a combination of two open (O-O), two closed (C-C), or one open and one closed (O-C) receptor structures.

S-C2

Figure 6.

Accuracies obtained for the test set compounds with all combinations of two receptor structures used in multiple receptor scoring. Each row and column is labeled according to the PDB code of the crystal structure the receptor was based upon. In each single column and row three entries represent the different water occupancies, W0, W2 and W2, in this order from left to right, and from top to bottom.

Figure 7.

Ligand **2a** docked to the binding site of HDAC8 (PDB: 1VKG, chain B). The space available to the ligands is rendered by surface colored according to lipophilicity, green – lipophilic, purple – hydrophilic.

Table 1

HDAC8 inhibitors used in this study.

Comp. No.	K _i [μM]	R (SBG)
1a	0.395	
1b	6.594	
1c	0.616	
1d	0.453	
1e	2.431	
2a	0.015	
2b	0.068	

<p>SBG</p> <p>linker</p> <p>ZBG</p>		
Comp. No.	Ki[μ M]	R (SBG)
2c	0.074	
2d	0.025	
3a	0.585	
3b	0.635	
4a	0.209	
4b	0.095	

Table 2

Re-docking of native ligands to HDAC8 co-crystal structures. Results obtained with two docking programs/scoring functions are given as RMSD values between experimental and docked ligands. The receptor structures used for docking are: one monomer (chainA), one monomer plus one or two conserved water molecules in the binding site (chainA+W1, chainA+W1+W2), the entire unit cell with all symmetry copies of the primary monomer (x-tal), and the entire unit cell plus one or two conserved water molecules in the binding site (x-tal+W1, x-tal+W1+W2).

PDB	Surflex-dock				Gold/Goldscore			
	chainA	chainA+W1	ChainA+W1+W2	x-tal	x-tal+W1	x-tal+W1+W2	x-tal	x-tal+W1
1T64	1.18	1.12	1.80	0.63	2.07	2.29	0.43	0.43
1T67	5.95	1.78	2.18	1.57	1.31	1.43	2.01	0.96
1T69	4.74	4.60	4.40	1.86	1.88	4.65	1.85	1.76
1VKG	6.06	6.36	3.44	1.50	0.96	2.05	3.54	3.41
1W22	7.32	7.19	2.97	1.97	1.30	0.98	1.53	1.48
2V5X	5.52	8.66	9.63	5.93	4.41	7.72	5.09	5.06
Average RMSD ^a	5.05	4.21	2.96	1.51	1.50	2.28	1.87	1.61

^a – RMSD values for 2V5X were excluded from averaging

Table 4

Average and best Pearson correlations between experimental and calculated activities obtained by using SC2 and SC6 docking scores from single receptors (top), multiple receptors (center), or normalized scores from multiple receptors (bottom). Also included is/are the receptor structure(s) giving the best correlation.

protocol	avg	best	recept 1	recept 2
single receptor				
SC6	0.23	0.67	2V5WB-W0	-
SC2	0.21	0.74	1VKGB-W0	-
two receptors				
SC6	0.22	0.67	1VKGB-W0	2V5WB-W0
SC2	0.20	0.78	1VKGB-W0	2V5XA-W1
two receptors, normalized scores				
SC6	0.20	0.72	2V5WB-W0	1T69A-W2
SC2	0.18	0.80	1VKGB-W0	2V5XA-W2