Journal of Computer-Aided Molecular Design **18:** 375–388, 2004.

Homology modeling, force field design, and free energy simulation studies to optimize the activities of histone deacetylase inhibitors

Hwangseo Park* & Sangyoub Lee*
School of Chemistry and Molecular Engineering, and Center for Molecular Catalysis, Seoul National University, Seoul 151-747, Korea

Received 21 February 2004; accepted in revised form 29 July 2004

Key words: docking, free energy perturbation, histone deacetylase, homology modeling, lead optimization, molecular dynamics

Summary

As an effort to develop therapeutics for cancer treatments, a number of effective histone deacetylase inhibitors with structural diversity have been discovered. To gain insight into optimizing the activity of an identified lead compound, a computational protocol sequentially involving homology modeling, docking experiments, molecular dynamics simulation, and free energy perturbation calculations was applied for rationalizing the relative activities of known histone deacetylase inhibitors. With the newly developed force field parameters for the coordination environment of the catalytic zinc ion in hand, the computational strategy proved to be successful in predicting the rank orders for 12 derivatives of three hydroxamate-based inhibitor scaffolds with indole amide, pyrrole, and sulfonamide moieties. The results showed that the free energy of an inhibitor in aqueous solution should be an important factor in determining the binding free energy. Hence, in order to enhance the inhibitory activity by adding or substituting a chemical group, the increased stabilization in solution due to the structural changes must be overcome by a stronger enzyme-inhibitor interaction. It was also found that to optimize inhibitor potency, the hydrophobic head of an inhibitor should be elongated or enlarged so that it can interact with Pro29 and His28 that are components of the flexible loop at the top of the active site.

Introduction

Chemical transformations of histone proteins complexed with DNA in the nucleosome, such as acetylation of their lysine sidechains, are known to have influence on gene transcription [1]. In this regard, inappropriate recruitment of histone deacetylases (HDACs) has been shown to cause transcriptional silencing in cancer cell lines. The inhibition of HDACs may thus lead to transcriptional reactivation, cell-cycle arrest, and terminal differentiation of transformed cells, and appears to be a new promising approach to cancer chemotherapy [2, 3].

The human histone deacetylases are classified into two distinct groups, the HDACs and the SIRTs [4]. The HDACs are further divided into two major subclasses according to the sequence homology. While class I HDACs (HDAC1, HDAC2, HDAC3, HDAC8, and HDAC11) are homologous to the yeast histone deacetylase called Rpd3, class II HDACs (HDAC4, HDAC5, HDAC6, HDAC7, HDAC9, HDAC10) have homology to the yeast HDAC Hda1. Although the class II HDACs are roughly twice as large as class I HDACs, all HDACs possess a highly conserved catalytic domain of approximately 390 amino acids with their catalytic activities dependent on the presence of a divalent zinc ion. On the other hand, SIRTs are a family of NAD-dependent deacetylases showing a significant sequence divergence from HDACs [5].

Although no three-dimensional structure of HDAC has been reported thus far except for the NAD-dependent HDAC (Sir2) [6], the crystal structure of histone deacetylase-like protein (HDLP) in unliganded

*Correspondence may be addressed to either author. Telephone: +82-2-880-6661; Fax: +82-2-874-3704; E-mail: hwangseo@snu.ac.kr; sangyoub@snu.ac.kr

and inhibitor-bound forms provided much information pertinent to the catalytic mechanism and inhibitor binding [7]. It was proposed that the acetylated lysine sidechain would be deacetylated by an activated water molecule in the active site through the electrophilic catalysis of a divalent zinc cation with one histidine and two aspartate ligands, which is reminiscent of a catalytic mechanism of zinc proteases [8]. The structure of HDLP also revealed that large and well-defined loops are associated with a binding pocket that has a tube-like shape with a depth of ∼11 Å. The pocket whose wall consists of hydrophobic residues gets wider near the bottom at which the zinc ion and polar residues are positioned so as to properly accommodate the acetylated amine moiety of the substrate.

Since the discovery of the antifungal antibiotic trichostatin A (TSA) and a few hydrophobic cyclic peptides as potent inhibitors of HDAC [9–12], a number of structurally diverse HDAC inhibitors have been reported in the literature, most of which possess a hydroxamic acid moiety at one terminus of the molecular structure to chelate the active site zinc ion [13–20]. A common structural feature of these inhibitors is the existence of a hydrophobic head that is separated from the zinc-binding group through a hydrophobic tether that extends 5 or 6 carbons long. Several non-hydroxamate inhibitors have also been identified, including short-chain fatty acids [21, 22], MS-27-275 [23], trifluoromethyl ketones [24], phosphorus-based compounds [25], bisulfide bromotyrosine derivatives [26], and α-keto amides [27]. Still in progress are the synthetic studies to optimize inhibitory strength, stability, and bioavailability.

The present study is undertaken to gain insight into improving the potency of identified inhibitor scaffolds by means of free energy perturbation (FEP) calculations based on molecular dynamics (MD) simulation. As a starting point, we constructed a model for the catalytic domain of HDAC1 using the X-ray structure of HDLP as a template that shares 35.2% sequence identity with the human HDAC1. For such a high sequence identity, homology modeling generally yields a structure that is accurate enough for use in subsequent computational studies [28]. For example, it is known that homology modeling from a template with a sequence identity of only about 20% can generate a protein model that serves as a good starting point for structure-based drug design if ligand binding information could be considered explicitly [29]. We analyzed the binding modes of recently reported HDAC inhibitors from their docking simulation in the HDAC1 catalytic core modeled with HDLP-TSA complex. Shown in Figure 1 are the structures of 12 derivatives of three inhibitor scaffolds under investigation. For these inhibitors, the most stable HDAC1-inhibitor complex found in the docking experiment was taken as a starting structure of the MD-based FEP calculation. Prior to the free energy simulation, we developed force field parameters to describe the active site zinc cluster complexed with the hydroxamic acid moiety because they are unavailable in the standard force field database. Then, we tried to explain the structure-activity relationship (SAR) of the known inhibitors by calculating the relative binding free energies with the hope that the resulting structural and energetic features for a variety of enzyme-inhibitor complexes would provide information pertinent to the optimization of inhibitor potencies.

Computational methods

Homology modeling

The peptide sequence of human HDAC1 comprising 482 amino acid residues was retrieved from the SWISS-PROT protein sequence data bank (http://us.expasy.or/sprot/; accession number Q13547) [30]. Domain analysis of the whole amino acid sequence in the Pfam database (http://www.sanger.ac.uk/Software/Pfam/) showed that residues 9–321 should constitute the catalytic domain with the identification of HDLP from *Aquifex aeolicus* as its homologue [31]. Sequence alignment between HDLP and the catalytic domain of HDAC1 was then derived with the CLUSTAL W package [32] using the BLOSUM matrices for scoring the alignments. The parameters of GAP OPEN, GAP EXTENSION, and GAP DISTANCE were set equal to 10, 0.05, and 8, respectively. Opening and extensions gap penalty were thus changed systematically, and the obtained alignment was inspected for violation of structural integrity in the structurally conserved regions. Based on the best-scored sequence alignment, the 3D structure of HDAC1 was constructed using the MODELLER 6v2 program [33]. In this model building, we employed an optimization method involving conjugate gradients and molecular dynamics to minimize violations of the spatial restraints. With respect to the structure of gap regions, the coordinates were built from a randomized distorted structure that is located approximately

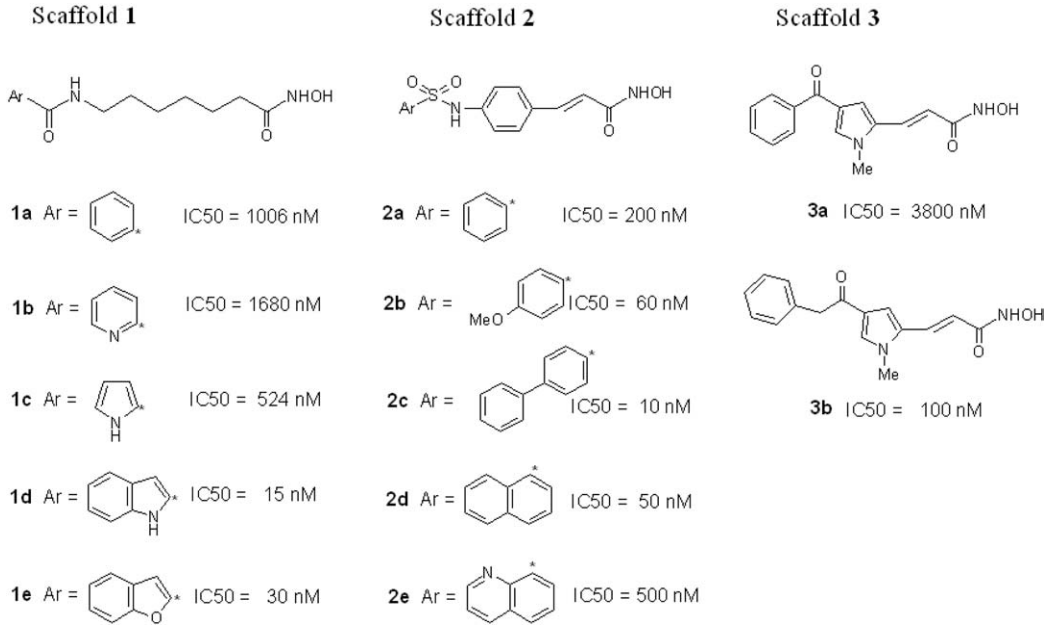

Figure 1. Molecular structures of 12 HDAC inhibitors of three scaffolds under investigation, together with their inhibitory data. An asterisk indicates the position at which the remaining structure of each inhibitor is connected.

between the two anchoring regions as implemented in MODELLER 6v2. To increase the accuracy of the calculated structure, loop modeling was also performed with the enumeration algorithm [34].

Docking experiments

The AutoDock 3.0.5 program [35] was used to obtain a starting structure for the MD simulations of the enzyme-inhibitor complexes. It combines a rapid energy evaluation through precalculated grids of affinity potentials with various search algorithms to find suitable binding positions for a ligand on a protein receptor. Although the protein structure has to be fixed, the program allows torsional flexibility of a ligand. The coordinates of the protein atoms were taken from the final model of homology modeling. Docking to the HDAC1 active site was then carried out using the new empirical free energy function and the Lamarckian genetic algorithm [35], applying a standard protocol. In this docking simulation, we used the standard AMBER force field for protein atoms and the newly developed parameters for the active-site zinc complex of HDAC1 to be consistent with those used in the subsequent free energy simulations. Of the conformations obtained from 200 independent docking runs, results differing by less than 1.5 Å in positional root-mean-square deviation (RMSD) were clustered together. Since zinc chelation is believed to be critical for HDAC1 inhibition, only the solutions in which the hydroxamic acid moiety is coordinated to the active site zinc ion with associated interatomic distances of less than 3.0 Å, were selected for further analyses. The suitability of this distance criterion may be supported by a recent docking study with AutoDock reported by Mai et al. in which the interatomic distance between the zinc ion and the carbonyl oxygen of hydroxamate inhibitors falls within 3.0 Å [16].

Force field design

The bonded approach [36, 37] was adopted to represent binding of the metal ion to the enzyme-inhibitor complex. Explicit bonds were placed between the zinc cation and its ligands, since other non-bonded counterparts have been known to be sensitive to the used electrostatic model, leading to an undesirable coordination geometry [38]. To derive the associated force-field parameters that are unavailable in the standard force field database, we followed the procedure suggested by Fox and Kollman [39] to be consistent with the AMBER force field [40]. The equilibrium bond lengths and angles for the active site zinc cluster were taken from the optimized structure of a model zinc complex representing the active site zinc cluster che-

lated with a hydroxamic acid group. The geometry optimizations were performed at the B3LYP/6-31G* level of theory with the Gaussian98 suite of programs. We extracted the force constant parameters involving the zinc ion from the earlier work of Hoops et al. [36] and Ryde [37]. Using the energy-minimized structure, RHF/6-31G* atomic partial charges were derived by using the RESP methodology [41], to be consistent with the standard AMBER force field. The Van der Waals parameters for the zinc ion were taken from the work of Hoops et al. [36], while those for the atoms in hydroxamic acid, carboxylate, and imidazole groups were assigned from the standard AMBER force field database. We derived the potential parameters for all HDAC inhibitors under investigation with the same procedure as used for the active-site zinc complex, which involves the geometry optimization and charge fitting with the RESP method. The newly obtained potential parameters for the active-site zinc complex and the inhibitors were also used in the precedent docking simulation.

In relation to the importance of the enzyme-inhibitor environment in zinc coordination, Vanommeslaeghe et al. have reported that more reliable force field parameters can be obtained with geometry optimization and population analysis on an extended model system at a modest level of theory [42]. Therefore, we carried out semiempirical PM3 and AM1 calculations with the MOPAC program as an effort to improve the potential parameters for the active site zinc complex of HDLP. However, it is found that the coordination geometry in the X-ray structure may not be maintained in geometry optimization, leading to the formation of an undesirable tetrahedral complex in the local energy minimum. We believe that the reason for this failure should be the lack of reliable semiempirical parameters for the zinc ion that can describe its pentacoordination.

Molecular dynamics and FEP calculation

MD simulations of the HDAC1 in complex with the inhibitors were carried out using the SANDER module of AMBER 7 [43] with the newly developed force field parameters in addition to those reported by Cornell et al. [40]. As a starting structure of the free energy simulation, we used the energy-minimized structure of the best-scored enzyme-inhibitor complex for which the coordinates of the inhibitor were extracted in the top-ranked conformational cluster. The top-ranked clusters exhibited the highest population

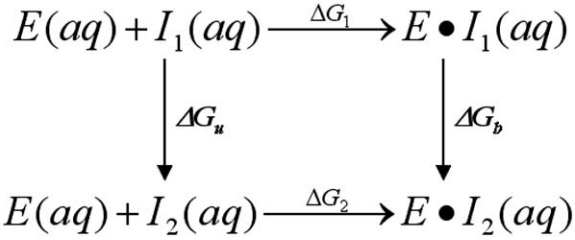

Figure 2. Thermodynamic cycle used in calculating free energy changes in aqueous solution (ΔG_u) and in HDAC1 active sites (ΔG_b), and the relative binding free energy ($\Delta\Delta G_{bind}$).

except for the inhibitors **1a**, **1b**, and **3b**. Even in these cases, the lowest energy cluster was considered only in further analysis because of a small difference in populations among the conformational clusters (see Supporting Information for details). The all-atom models for the enzyme-inhibitor complexes were neutralized by adding chloride ions and were immersed in rectangular boxes containing about 11,000 TIP3P water molecules. After 1000 cycles of energy minimization to remove any bad van der Waals contacts, we equilibrated the systems beginning with 20 ps equilibration dynamics of the solvent molecules at 300 K. The next step involved equilibration of the solute with a fixed configuration of the solvent molecules for 5 ps at 10, 50, 100, 150, 200, 250, and 300 K. Then, the equilibration dynamics of the entire system was performed at 300 K for 20 ps.

MD-FEP calculation was performed using a thermodynamic cycle-perturbation (TCP) approach [44–46], which has been widely used in the theoretical studies of various enzyme-inhibitor complexes [47–52]. As depicted in Figure 2, the relative change in binding free energy ($\Delta\Delta G_{\text{bind}}$) can be computed through nonphysical paths connecting the desired initial and final states. More specifically, this approach enables calculation of $\Delta\Delta G_{\text{bind}}$ between two structurally similar inhibitors by computationally simulating the 'mutation' of one to the other. Since the free energy is a state function and, therefore, the sum of all energies in the thermodynamic cycle in Figure 2 is zero, $\Delta\Delta G_{\text{bind}}$ can be expressed as follows:

$$\Delta\Delta G_{bind} = \Delta G_2 - \Delta G_1 = \Delta G_b - \Delta G_u \quad (1)$$

where ΔG_1 and ΔG_2 refer to the binding free energies of inhibitors I_1 and I_2, respectively, and ΔG_u and ΔG_b to the free energies associated with the mutation of I_1 into I_2 in aqueous solution and in the enzymatic active site.

```
HDLP     2  KKVKLIGTLDYGKYRYPKNHPLKIPRVSLLLRFLDAMNLIDEKELIKSRPATKEELLLFH  61
HDAC1    9  RKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKANAEEMTKYH  68
            :**      * *:* * : **:*  *: :   :*    *  : *: : : *  **:  :*

HDLP    62  TEDYINTIMEAERCQCVPKGAREKYNIGGYENPVSYAMFTGSSLATGSTVQAIEEFLKG-  120
HDAC1   69  SDDYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGLFEFCQLSTGGSVASAVKLNKQQ  128
            ::***: *      :      : :    * : **   :*    *:** :* :  :: *

HDLP   121  -NVAFNPAGGMHHAFKSRANGFCYINDPAVGIEYLRKKGFKRILYIDLDAHHCDGVQEAF  179
HDAC1  129  TDIAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKY-HQRVLYIDIDIHHGDGVEEAF  187
             ::* * ***:*** ** * ****:**  : *  * *   :*:****:* ** ***:***

HDLP   180  YDTDQVFVLSLHQSPEYAFPFEKGFLEEIGEGKGKGYNLNIPLPKGLNDNEFLFALEKSL  239
HDAC1  188  YTTDRVMTVSFHKYGEYFP--GTGDLRDIGAGKGKYYAVNYPLRDGIDDESYEAIFKPVM  245
            * **:*: :*:*:  **      * * :** **** * :* **  *::*: :   ::  :

HDLP   240  EIVKEVFEPEVYLLQLGTDPLLEDYLSKFNLSNVAFLKAFNIVREVFGEGVYLGGGGYHP  299
HDAC1  246  SKVMEMFQPSAVVLQCGSDSLSGDRLGCFNLTIKGHAKCVEFVKSFNLPMLMLGGGGYTI  305
               * *:*: *:** *:* *  * *  ***:     *  ::*:      : ******

HDLP   300  YALARAWTLIWCELSGREVPEKLN-NKAKELLK  331
HDAC1  306  RNVARCWTYETAVALDTEIPNELPYNDYFEYFG  338
              :** **         *:*::*  *   * :
```

Figure 3. Sequence alignment of the catalytic domain of HDAC1 with its homologue, HDLP. Asterisk (*) and colon (:) represent identity and similarity between the corresponding residues, respectively.

To determine $\Delta\Delta G_{\text{bind}}$ between I_1 and I_2, therefore, the free energy changes of mutating I_1 into I_2 should be calculated both in aqueous solution and in the HDAC1 active site. The free energy change for converting I_1 into I_2 is computed by perturbing the Hamiltonian of I_1 (initial state) into I_2 (final state). This transformation is accomplished through a parametrization of terms comprising the interaction potentials of the system with a change of state variable (λ) that maps onto reactant ($\lambda = 0$) and product ($\lambda = 1$) states. The total free energy change for the mutation from the initial to the final state is then computed by summing incremental free energy changes over several windows visited by λ as it changes from 0 to 1, which is summarized in the following equation:

$$\begin{aligned}\Delta G &= G_1 - G_0 = \sum_i \left(G_{\lambda(i+1)} - G_{\lambda(i)}\right) \\ &= -RT \sum_i \ln \left\langle \exp - \left[\frac{(V_{\lambda(i+1)} - V_{\lambda(i)})}{RT}\right] \right\rangle_{\lambda(i)}\end{aligned} \tag{2}$$

where G_0 and G_1 are the free energies of states 0 and 1, respectively, $V_{\lambda(i)}$ represents the potential energy function for the representative state $\lambda(i)$, and $<>_{\lambda(i)}$ designates the ensemble average of the enclosed quantity for the representative state $\lambda(i)$.

Following the aforementioned equilibration dynamics, we performed 315 ps of perturbation for a selected pair of inhibitors. Each perturbation consisted of 21 windows with 5000 steps of equilibration and 10000 steps of data collection. In this MD-based FEP calculation, a periodic boundary condition was employed in the NPT ensemble at 300 K using Berendsen temperature coupling [53] and constant pressure (1 atm) with isotropic molecule-based scaling. The SHAKE algorithm [54], with a tolerance of 10^{-6}, was applied to fix all bond lengths involving hydrogen atom. We used a time step of 1 fs and a nonbonded-interaction cutoff radius of 12 Å. A doublewide sampling procedure was performed for all of the structural transformations, and the reported results are based on the averages from the backward and forward simulations. As a rough estimation of the statistical error, we used half the difference between the absolute values of forward and reverse free energy changes, and the error in $\Delta\Delta G_{\text{bind}}$ is calculated as the square root of the sum of the squares of the individual errors in ΔG_u and ΔG_b.

Results and discussion

HDAC1 homology modeling

Figure 3 displays the sequence alignment of HDAC1 with HDLP, which was used as the template for constructing the homology model. Of the ten models cal-

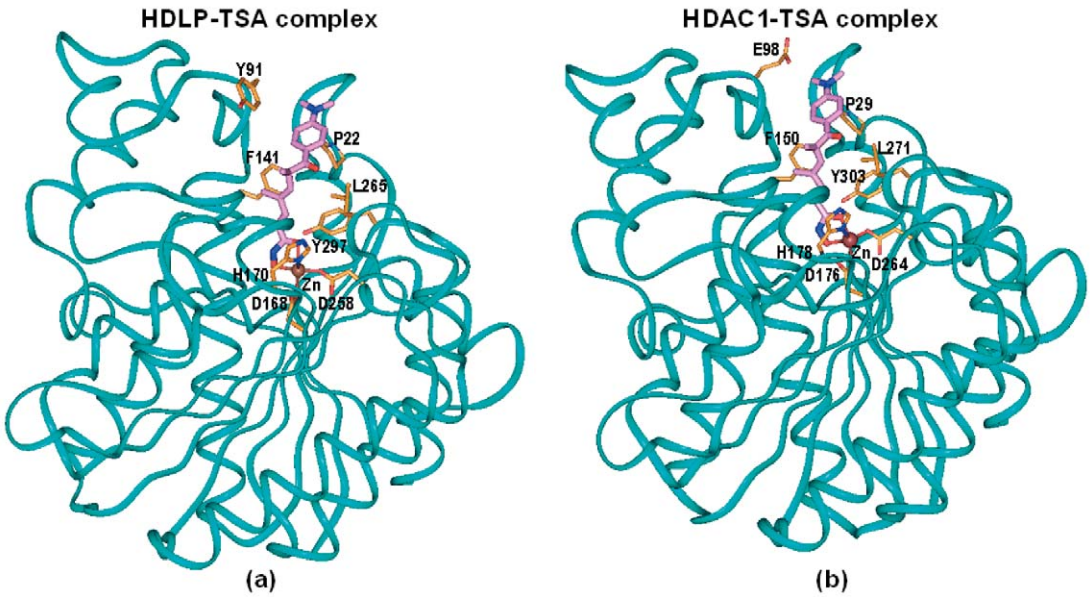

Figure 4. Comparative view of (a) the X-ray structure of HDLP in complex with TSA and (b) the final model in the homology modeling of HDAC1 complexed with TSA. Carbon atoms of the inhibitor and the enzymatic groups are shown in pink and orange, respectively.

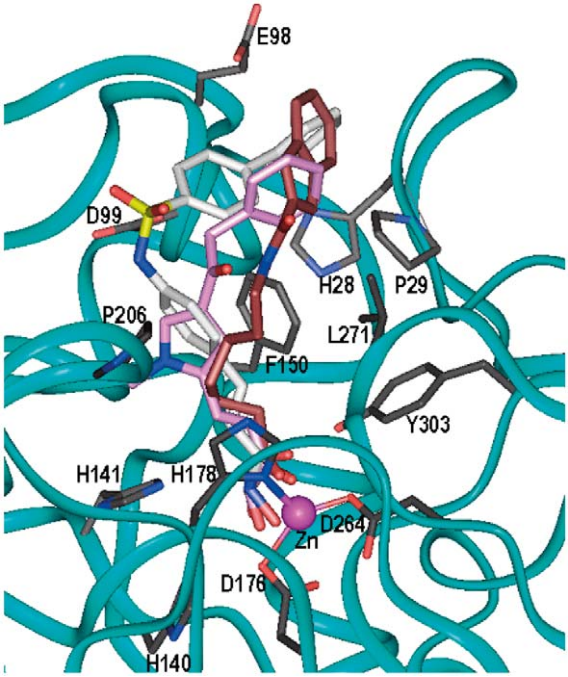

Figure 5. Superimposition of docked conformations of **1d** (brown), **2c** (pink), and **3d** (gray).

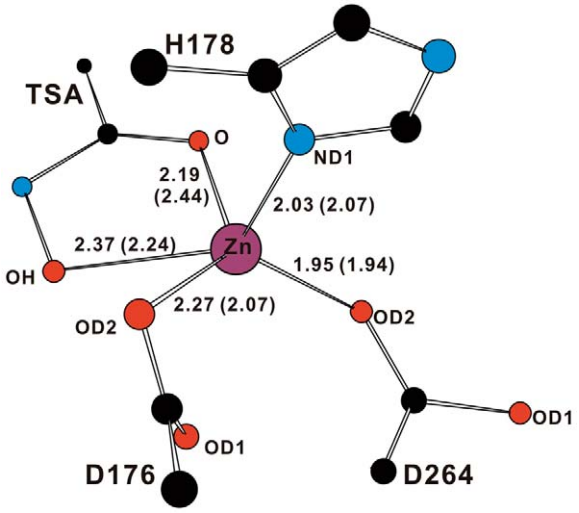

Figure 6. B3LYP/6-31G* optimized structures of the zinc cluster modeling the active site of HDAC1 in complex with a hydroxamic acid inhibitor. Numbers in and out of parentheses are, respectively, the interatomic distances (in Å) between the zinc ion and its ligand atoms in the X-ray structure of HDLP-TSA complex and those in the optimized structure of the model system. All hydrogen atoms are omitted for clarity.

culated for the same target (HDAC1) and the template (HDLP), the one with the lowest value of the MODELLER objective function was selected as the best model for HDAC1. The geometry of this final model was then evaluated with PROCHECK [55], revealing that the backbone Φ and Ψ dihedral angles of 84.7%, 14.0%, and 1.3% of the residues are located within most favorable, additionally allowed, and generously allowed regions, respectively, of the Ramachandran plot. This good stereochemical quality is not surprising for the high sequence identity (35%) and similarity (56%) between the template and the target, which are shown in Figure 3.

Since the model building of an enzyme-inhibitor complex is believed to be more relevant for the purpose of our study of the enzyme inhibition rather than that of apoenzyme, homology modeling of HDAC1 was performed in the presence of the potent inhibitor TSA. Compared in Figure 4 are the X-ray structures of HDLP-TSA and the final model of HDAC1-TSA complexes. The template and the target possess very similar folding structure and are superimposable over the main-chain atoms. As in the crystal structure of HDLP, a zinc ion resides at the bottom of the catalytic domain of HDAC1, revealing a distorted square pyramidal coordination with five ligands: Asp176, His178, Asp264, and the hydroxamic acid group of the inhibitor. Tyr303 in the active site of HDAC1 that corresponds to Tyr297 in HDLP, donates a hydrogen bond to the inhibitor carbonyl oxygen. The residues of HDAC1 (Pro29, Phe150, and Leu271) responsible for hydrophobic interactions with inhibitor groups are also identical to those of HDLP (Pro22, Phe141, and Leu265). However, the residues Gln26, Gly27, Met30, Glu98, Asp99, and Cys100, which are components of the active site loops of HDAC1, differ from the corresponding Lys19, Asn20, Leu23, Tyr91, Glu92, and Asn93 in HDLP, respectively, suggesting that a structural refinement of the active site loops would be required for detailed studies of the catalytic mechanism and enzyme inhibition.

Docking experiments

All inhibitors that need to be mutated for free energy simulations were docked onto the modeled active site of HDAC1 from which the inhibitor TSA had been removed. The number of conformational clusters involving the chelation of the zinc ion by the hydroxamate moiety ranges from 5 to 15 depending on the number of rotatable bonds. No single cluster exhibits a dominant population except for the inhibitors of Scaffold 2 for which 45–68% of docking solutions are included in the best-scored one (see Supporting Information for details). Consistent with the docking analysis reported by Mai et al. [16], either carbonyl or hydroxyl oxygen of the inhibitors under investigation is situated at a distance of less than 3.0 Å from the zinc ion in all AutoDock solutions in the best-scored conformational cluster. Figure 5 shows the best-scored AutoDock conformations of compounds **1d, 2c**, and **3b**, each of which is the most potent inhibitor of the respective scaffold under investigation. The results from the docking simulation are self-consistent in that the functional groups of similar chemical character are placed in similar ways with comparable interactions with the protein groups. As revealed by superposition of their docked structures, for example, the hydroxamic acid group and the hydrophobic head of the inhibitors are directed to the catalytic zinc ion and the active site loop structure, respectively. The carbonyl and sulfonyl oxygens of the inhibitors are exposed to the exterior of protein, residing in different lateral positions in the active site gorge. These groups are expected to interact with bulk solvent in aqueous solution.

Force field design for active site zinc cluster in complex with a hydroxamic acid

We used a general approach to systematically derive the potential parameters for metalloproteins to extend the AMBER force field to model the zinc ion in HDAC1. The method involves geometry optimization of the model system for active site zinc complex chelated by a hydroxamic acid inhibitor. In this simplified model for an enzyme-inhibitor complex, methyl imidazole and two acetate ions are used to represent the active site histidine and two aspartate sidechains, respectively. Figure 6 displays the structure of a local energy minimum whose input structure was taken from the X-ray structure of HDLP-TSA complex. As in the crystal structure of HDLP-TSA complex and in the homology model of HDAC1-TSA complex, the zinc ion has a distorted square pyramidal coordination with the imidazole group occupying the axial position. As indicated in Figure 6, the interatomic distances involving the zinc ion compare reasonably well with those in the crystal structure of HDLP-TSA complex except for those associated with zinc chelation by the inhibitor hydroxamate group. This can be attributed to the lack of macromolecular structural environment

around the zinc complex including the polar residues such as His140, His141, and Tyr303 that stabilize the hydroxamate group at the bottom of the active site.

It is also noted that in the optimized structure, the RESP atomic charge of the zinc ion decreases from +2 e to +0.8 e. On the other hand, the atomic charges of ligand atoms become less negative by 0.1 e − 0.3 e when compared to those in the standard AMBER force-field database and in hydroxamic acid inhibitors *in vacuo*. These reflect charge redistribution between the zinc ion and its ligands upon formation of metal complex. We used these newly obtained atomic charges in the free energy simulation because it had been well appreciated that the 'Zn^{2+}' model with standard force field for ligands is inadequate for maintaining the coordination geometry of an active site zinc complex during the simulation [36–38].

MD-FEP calculation

As a check on the reliability of the simulation, we check to see if the protein structure in the HDAC1-inhibitor complexes stays close to that of the starting structure obtained from homology modeling under the simulation conditions. For this purpose, we calculate the time dependence of the root-mean-square deviation from the initial structure ($RMSD_{init}$) for all backbone heavy atoms of HDAC1 in the free energy simulations. As can be seen in Figure 7, the backbone $RMSD_{init}$ values for the enzyme-inhibitor complexes remain within 2.0 Å with a convergent behavior with respect to simulation time, indicating the stability of the protein conformation during the entire course of simulations. We thus expect that the calculated free energy differences would be due to the structural perturbations of the ligand and the resulting changes in the enzyme-inhibitor interaction instead of some unwanted large conformational changes in the protein.

Figure 8 displays the time evolutions of $RMSD_{init}$ values for all heavy atoms of the HDAC inhibitors under investigation in the course of free energy simulations. The $RMSD_{init}$ values of all ligands are found to fall within 1.5 Å, and remain around 1.0 Å at the end of simulation in most cases. This indicates that during the structural transformations, the inhibitors have been maintained stable in the active site with a positional shift less than the conformational change of the protein, further supporting the reasonableness of the simulation conditions described in the previous section.

Summarized in Table 1 are the calculated relative binding free energies with respect to a single reference compound, in comparison with the corresponding ratio of experimental IC_{50} values of the latter to that of the former. In accordance with the previous experimental results of Dai et al. [14], our FEP calculations predict that potencies of the five inhibitors belonging to Scaffold 1 would increase in the following order: **1b < 1a < 1c < 1e < 1d**. However, it appears that the large difference (6.5 kcal/mol) in $\Delta\Delta G_{bind}$ between **1d** and **1e** should be overestimated in the calculations as compared to only a 3-fold difference in IC_{50} values. Judging from the comparison of ΔG_u and ΔG_b, the decrease of inhibitory strength in going from **1a** to **1b** may be attributed to the greater stabilization in aqueous solution than in the HDAC1 active site with the change of the hydrophobic head from phenyl to pyridinyl groups. This result indicates that to increase the inhibitor potency, the structural modification should be made in such a way that the resulting increase of enzyme-inhibitor interaction overbalances that of stabilization energy in bulk solvent, as in the case of the substitution of a pyrrole moiety in **1c** for the phenyl ring in **1a**. For the same reason, the extension of the hydrophobic head from pyrrole in **1c** to indole in **1d** leads to the increase of inhibitory strength. However, the replacement of -NH- in **1d** with -O- in **1e** lowers the inhibitor potency, due mainly to weakening of the enzyme-inhibitor interaction. Taken together, it is likely that the activity optimization of Scaffold 1 should involve the introduction of a hydrogen bond donor group and the enlargement of the aromatic group in the hydrophobic head. Indeed, a detailed trajectory analysis for the HDAC1-**1d** complex indicates that as shown in Figure 9, the hydrogen bond between the inhibitor indole moiety and the side-chain of Asp99 may be a significant binding force that stabilized **1d** in the HDAC1 active site.

The size of the hydrophobic head is also found to be important in raising the activity of Scaffold 2 [15]. As shown in Table 1, both the previous experimental and the present computational results indicate that among the substitutions under consideration, the biphenyl group in **2c** would be the best choice for the hydrophobic head. As can be inferred from the positive and negative signs of ΔG_u and ΔG_b in the **2a** → **2c** perturbation, respectively, the resulting increase in the inhibitor potency should be due to the combined effects of destabilization of the inhibitor in bulk solvent and strengthening of the enzyme-inhibitor interaction. The perturbation calculation for the **2a** →

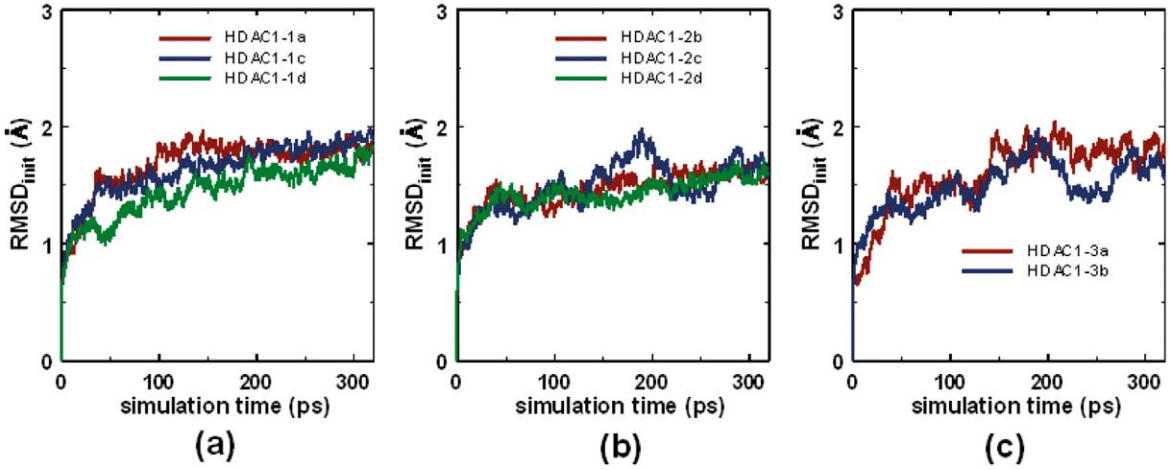

Figure 7. Time evolutions of the RMS deviation from initial structures of backbone heavy atoms in HDAC-inhibitor complexes for (a) Scaffold 1, (b) Scaffold 2, and (c) Scaffold 3.

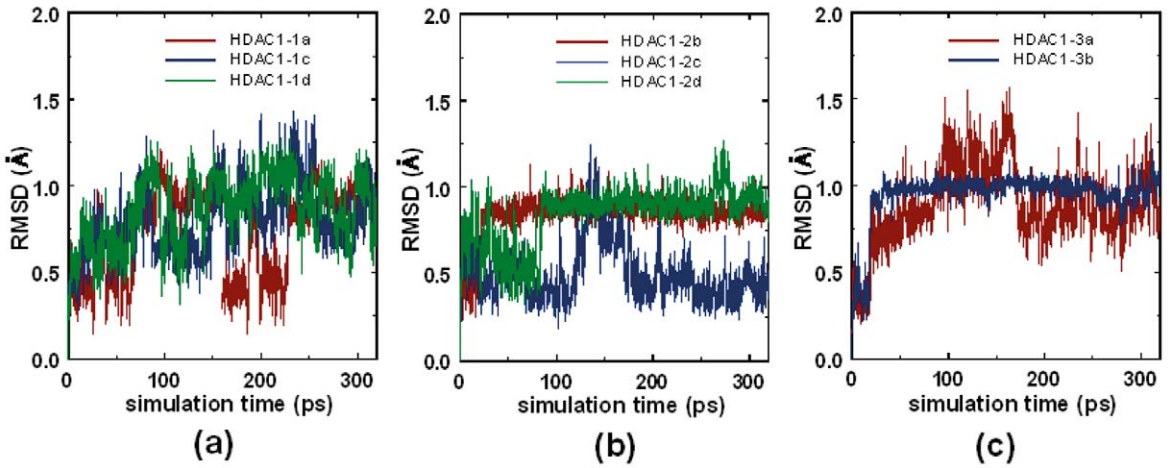

Figure 8. Time evolutions of the RMS deviation from initial structures for all heavy atoms of the inhibitors belonging to (a) Scaffold 1, (b) Scaffold 2, and (c) Scaffold 3.

2b mutation shows that although the stabilization of an inhibitor in the HDAC1 active site can be maximized by substitution of a methoxy group at the para position of **2a**, simultaneously increased stabilization in aqueous solution enhances the inhibitory strength of **2b** only by a factor of 3 as compared to that of **2a**. Both ΔG_u and ΔG_b become positive with the enlargement of the hydrophobic head from phenyl in **2a** to naphthyl in **2d**. In this mutation, however, the inhibitory strength increases because the destabilization in aqueous solution is greater than that in the HDAC1 active site. The present FEP calculation predicts that **2d** and **2e** are almost equipotent, which is inconsistent with the experimental finding that the former is more potent than the latter by a factor of 10. For the present, unfortunately, we are unable to find a reason for such an inconsistency without ambiguity.

To find the structural relevance to understand the size effects of the hydrophobic head on the potency of inhibitors belonging to Scaffold 2, we examine the characteristics of its interaction with the HDAC1 active site. Compared in Figure 10 are the representative MD trajectory snapshots of HDAC1-**2c** and HDAC1-**2d** complexes. We see that the biphenyl moiety of **2c** interacts with His28 and Pro29 that are components of a flexible loop at the top of the active site, and with the sidechain of Phe150 (Figure 10a). On the other hand, the naphthyl group of **2d** stays away from the loop residues and points toward the sidechains of Phe150 and Leu271 to establish a Van der Waals

Table 1. Calculated free energy changes in aqueous solution (ΔG_u) and in HDAC1 active site (ΔG_b), and the binding free energy change ($\Delta\Delta G_{bind}$) for the inhibitors with respect to a single reference compound, in comparison with the experimental IC_{50} value of the final state to that of the initial state. All energy values are given in kcal/mol.

Mutation	ΔG_u	ΔG_b	$\Delta\Delta G_{bind}$	IC_{50} ratio
Scaffold 1				
1a → **1b**	− 8.9 ± 0.0	− 8.4 ± 0.0	0.5 + 0.0	1.67
1a → **1c**	− 1.2 ± 1.5	− 3.5 ± 0.3	−2.3 +1.5	0.52
1a → **1d**	−64.8 ± 1.5	−74.6 + 0.4	−9.8 ± 1.6	0.01
1a → **1e**	−41.4 ± 1.6	−44.7 ± 0.7	−3.3 ± 1.7	0.03
Scaffold 2				
2a → **2b**	− 5.7 ± 0.2	− 6.6 ± 0.2	−0.9 ± 0.3	0.30
2a → **2c**	2.7 ± 0.5	− 1.2 ± 0.5	−3.9 ± 0.7	0.05
2a → **2d**	3.2 ± 0.1	1.2 ± 0.2	−2.0 ± 0.2	0.25
2a → **2e**	− 1.2 ± 0.1	− 3.5 ± 0.1	−2.3 ± 0.1	2.50
Scaffold 3				
3a → **3b**	− 0.1 ± 0.8	− 2.1 ± 0.2	−2.1 ± 0.8	0.03

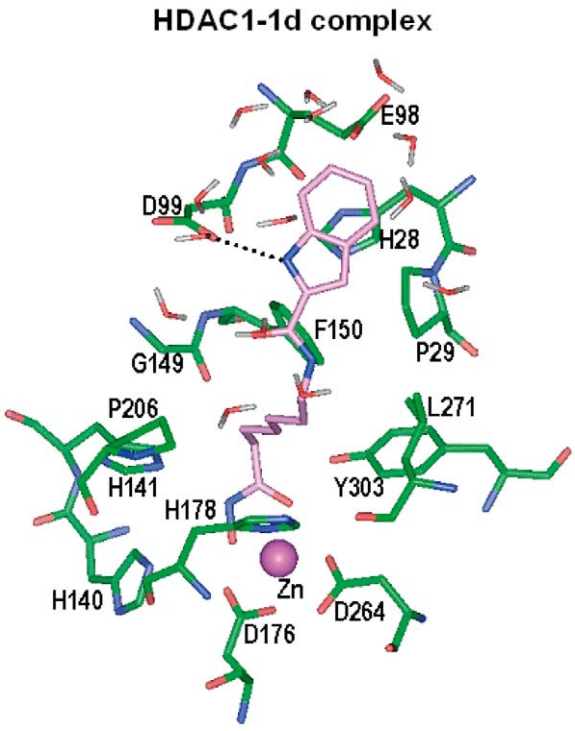

Figure 9. Representative MD trajectory snapshot of HDAC1-**1d** complex including solvent molecules near the inhibitor. The dotted line indicates the hydrogen bond that is established between the inhibitor indole moiety and the sidechain of Asp99.

contact (Figure 10b). Since the crystallographic temperature factors in the X-ray structure of free HDLP show a major peak at the loop consisting of residues 14–22 [7], the corresponding flexible loop of HDAC1 including His28 and Pro29 is likely to play a significant role in substrate binding and product release through a high-amplitude motion as confirmed in other proteins [56–58]. Therefore, a tight binding to such a flexible loop structure may be an important component of enzyme inhibition. Thus, the superiority of the inhibitor **2c** among those of Scaffold 2 can be attributed to the possession of a suitable hydrophobic group (biphenyl moiety) with appropriate length to interact with the loop residues.

We now turn to Scaffold 3 to further address the effect of the length of the hydrophobic head on binding free energies. As shown in Table 1, one-carbon elongation from **3a** and **3b** leads to a decrease in binding free energy by about 2.1 kcal/mol, which is consistent with the experimental data [16]. The representative MD trajectory snapshots for the two inhibitors are shown in Figure 11 to compare their binding modes in the HDAC1 active site. A common structural feature is an establishment of a hydrophobic contact between the inhibitor pyrrole moiety and Gly149, Phe150, and Pro206. It is characteristic of HDAC1-**3a** complex that the hydrophobic head is exposed to bulk solvent without involving a significant interaction with enzymatic groups because of the water molecules solvating the active site loop residues including His28,

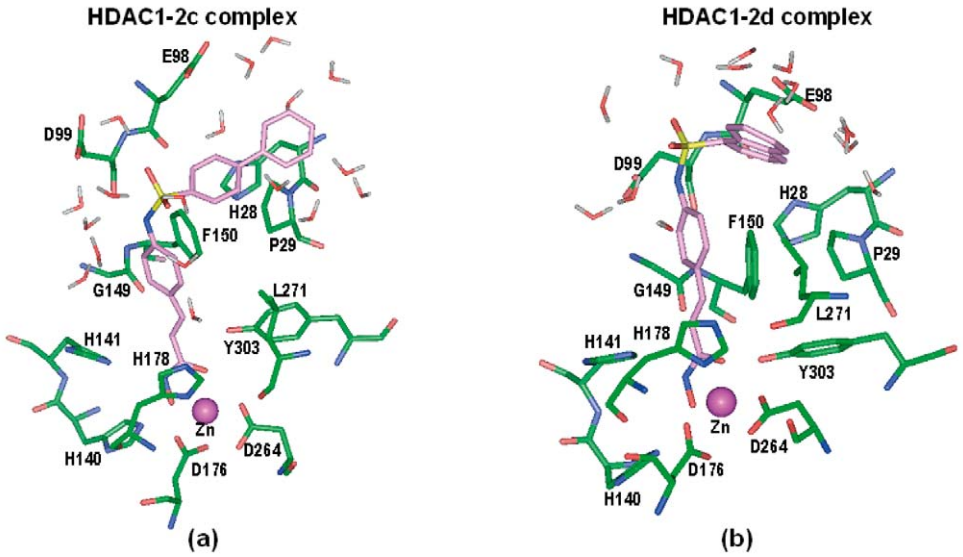

Figure 10. Representative MD trajectory snapshots of (a) HDAC1-**2c** and (b) HDAC1-**2d** complexes including solvent molecules near the inhibitors.

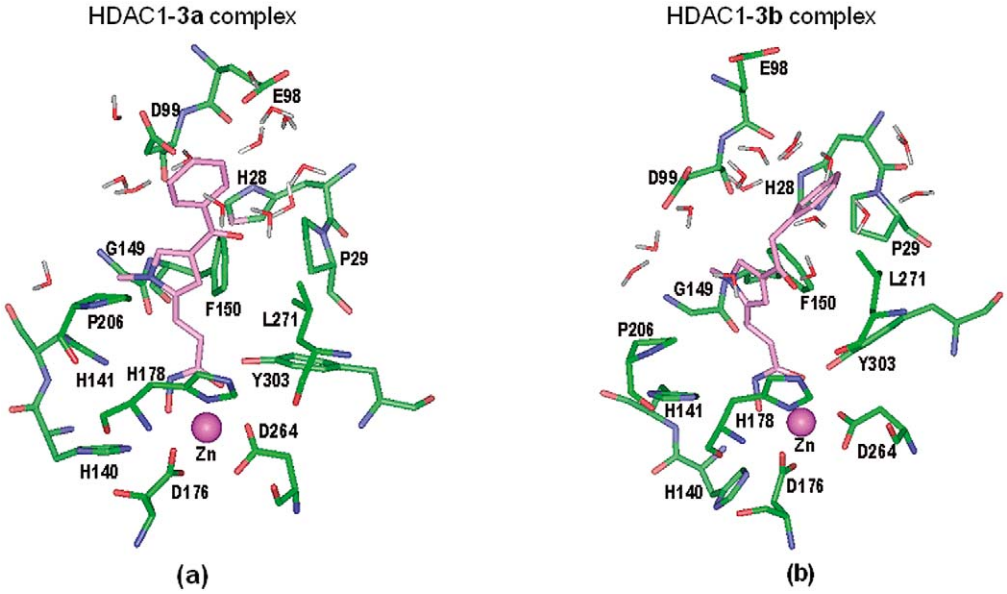

Figure 11. Representative MD trajectory snapshots of (a) HDAC1-**3a** and (b) HDAC1-**3b** complexes including solvent molecules near the inhibitors.

Pro29, Glu98, and Asp99. This hydration shell is most likely to prevent the phenyl moiety of the inhibitor from approaching the loop residues, playing a role of limiting the inhibitory strength of **3a**. On the other hand, as shown in the structure of HDAC1-**3b** complex (Figure 11b), the elongation of the hydrophobic head from phenyl to benzyl groups facilitates binding of the inhibitor to Pro29. Although a solvation shell is also found around Glu98 and Asp99, this is unlikely to significantly affect the inhibitor potency because neither of the inhibitors possesses a hydrogen-bonding group at the hydrophobic head.

Conclusions

We have applied a computational protocol sequentially involving homology modeling, docking experiments,

molecular dynamics simulation, and free energy perturbation calculations to find structural features of HDAC inhibitors that can lead to increase their potencies. The folding found in the final model of HDAC1 is very similar to that of its homolog HDLP with the same coordination environment around the catalytic zinc ion. With the newly developed potential parameters to describe the zinc coordination in the enzyme-inhibitor complex, FEP calculations could successfully explain the SAR of three inhibitor scaffolds. The free energy of an inhibitor in aqueous solution proves to be an important factor in binding free energy, indicating that a structural modification of an inhibitor should be made in such a way that strengthening of the enzyme-inhibitor interaction overcomes any increased desolvation cost. Introduction of a hydrogen bond donor at the hydrophobic head is found to increase the inhibitory activity due to the formation of a hydrogen bond with the side chain of Asp99. To optimize the inhibitory strength, more importantly, the hydrophobic head should be elongated so that hydrophobic interactions with the flexible loop at the top of the active site can be facilitated. Further study is needed to find additional factors for maximizing the inhibitor potency.

Acknowledgements

This work was supported by Grant No. R02-2001-000-00022-0 from the Basic Research Program of the Korea Science and Engineering Foundation.

References

1. Jenuwein, T. and Allis, C.D., Science, 293 (2001) 1074.
2. Johnstone, R.W., Nat. Rev. Drug Discov., 1 (2002) 287.
3. Marks, P.A., Rifkind, R.A., Richon, V.M., Breslow, R., Miller, T. and Kelly, W.K., Nat. Rev. Cancer, 1 (2001) 194.
4. Gray, S.G. and Ekstrom, T.J., Exp. Cell Res., 262 (2001) 75.
5. Imai, S., Armstrong, C.M., Kaeberlein, M. and Guarente, L., Nature, 403 (2000) 795.
6. Finnin, M.S., Donigian, J.R. and Pavletich, N.P., Nat. Struct. Biol., 8 (2001) 621.
7. Finnin, M.S., Donigian, J.R., Cohen, A., Richon, V.M., Rifkind, R.A., Marks, P.A., Breslow, R. and Pavletich, N.P., Nature, 401 (1999) 188.
8. Lipscomb, W.N. and Sträter, N., Chem. Rev., 96 (1996) 2375.
9. Yoshida, M., Kijima, M., Akita, M. and Beppu, T., J. Biol. Chem., 265 (1990) 17174.
10. Kijima, M., Yoshida, M., Sugita, K., Horinouchi, S. and Beppu, T., J. Biol. Chem., 268 (1993) 22429.
11. Nakajima, H., Kim, Y.B., Terano, H., Yoshida, M. and Horinouchi, S., Exp. Cell. Res., 241 (1988) 126.
12. Darkin-Rattray, S., Gurnett, A.M., Myres, R.W., Dulski, P.M., Crumely, K.M., Allocco, J.J., Cannova, C., Meinke, P.T., Colletti, S.L., Bednarek, M.A., Singh, S.B., Goetz, M.A., Dombrowski, A.W., Polishook, J.D. and Schimatz, D.M., Proc. Natl. Acad. Sci. USA, 93 (1996) 13143.
13. Kelly, W.K., O'Connor, O.A. and Marks, P.A., Exp. Opin. Investig. Drugs, 11 (2001) 1695.
14. Dai, Y., Guo, Y., Guo, J., Pease, L.J., Li, J., Marcotte, P.A., Glaser, K.B., Tapang, P., Albert, D.H., Richardson, P.L., Davidsen, S.K. and Michaelides, M.R., Bioorg. Med. Chem. Lett., 13 (2003) 1897.
15. Bouchain, G., Leit, S., Frechette, S., Khalil, E.A., Lavoie, R., Moradei, O., Woo, S.H., Fournel, M., Yan, P.T., Kalita, A., Trachy-Bourget, M.-C., Beaulieu, C., Li, Z., Robert, M.F., MacLeod, A.R., Besterman, J.M. and Delorme, D., J. Med. Chem., 46 (2003) 820.
16. Mai, A., Massa, S., Ragno, R., Cerbara, I., Jesacher, F., Loidl, P. and Brosch, G., J. Med. Chem., 46 (2003) 512.
17. Uesato, S., Kitagawa, M., Nagaoka, Y., Maeda, T., Kuwajima, H. and Yamori, T., Bioorg. Med. Chem. Lett., 12 (2002) 1347.
18. Remiszewski, S.W., Sambucetti, L.C., Atadja, P., Bair, K.W., Cornell, W.D., Green, M.A., Howell, K.L., Jung, M., Kwon, P., Trogani, N. and Walker, H., J. Med. Chem., 45 (2002) 753.
19. Mai, A., Massa, S., Ragno, R., Esposito, M., Sbardella, G., Nocca, G., Scatena, R., Jesacher, F., Loidl, P. and Brosch, G., J. Med. Chem., 45 (2002) 1778.
20. Wittich, S., Scherf, H., Xie, C., Brosch, G., Loidl, P., Gerhauser, C. and Jung, M., J. Med. Chem., 45 (2002) 3296.
21. Phiel, C.J., Zhang, F., Huang, E.Y., Guenther, M.G., Lazar, M.A. and Klein, P.S., J. Biol. Chem., 276 (2001) 36734.
22. Göttlicher, M., Minucci, S., Zhu, P., Krämer, O.H., Schimpf, A., Giavara, S., Sleeman, J.P., Coco, F.L., Nervi, C., Pelicci, P.G. and Heinzel, T., EMBO J., 20 (2001) 6969.
23. Saito, A., Yamashita, T., Mariko, Y., Nosaka, Y., Tsuchiya, K., Ando, T., Suzuki, T., Tsuruo, T. and Nakanishi, O., Proc. Natl. Acad. Sci. USA, 96 (1999) 4592.
24. Frey, R.R., Wada, C.K., Garland, R.B., Curtin, M.L., Michaelides, M.R., Li, J., Pease, L.J., Glaser, K.B., Marcotte, P.A., Bouska, J.J., Murphy, S.S. and Davison, S.K., Bioorg. Med. Chem. Lett., 12 (2002) 3443.
25. Kapustin, G.V., Fejer, G., Gronlund, J.L., McCafferty, D.G., Seto, E. and Etzkorn, F.A., Org. Lett., 5 (2003) 3053.
26. Pina, I.C., Gautschi, J.T., Wang, G.-Y.-S., Sanders, M.L., Schmitz, F.J., France, D., Cornell-Kennon, S., Sambucetti, L.C., Remiszewski, S.W., Perez, L.B., Bair, K.W. and Crews, P., J. Org. Chem., 68 (2003) 3866.
27. Wada, C.K., Frey, R.R., Ji, Z., Curtin, M.L., Garland, R.B., Holms, J.H., Li, J., Pease, L.J., Guo, J., Glaser, K.B., Marcotte, P.A., Richardson, P.L., Murphy, S.S., Bouska, J.J., Tapang, P., Magoc, T.J., Albert, D.H., Davidsen, S.K. and Michaelides, M.R., Bioorg. Med. Chem. Lett., 13 (2003) 3331.
28. Baker, D. and Sali, A., Science, 294 (2001) 93.
29. Evers, A. and Klebe, G., Angew. Chem. Int. Ed. Engl., 43 (2004) 248.
30. Bairoch, A. and Apweiler, R., Nucleic Acids Res., 27 (1999) 49.
31. Sonnhammer, E.L.L., Eddy, S.R. and Durbin, R., Proteins, 28 (1997) 405.
32. Thompson, J.D., Higgins, D.G. and Gibson, T.J., Nucleic Acids Res., 22 (1994) 4673.
33. Sali, A. and Blundell, T.L., J. Mol. Biol., 234 (1993) 779.
34. Fiser, A., Do, R.K.G. and Sali, A., Protein Sci., 9 (2000) 1753.

35. Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K. and Olson, A.J., J. Comput. Chem., 19 (1998) 1639.
36. Hoops, S.C., Anderson, K.W. and Merz, K.M., Jr., J. Am. Chem. Soc., 113 (1991) 8262.
37. Ryde, U., Proteins, 21 (1995) 40.
38. Stote, R.H. and Karplus, M., Proteins, 23 (1995) 12.
39. Fox, T. and Kollman, P.A., Phys. Chem. B, 102 (1998) 8070.
40. Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Jr., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W. and Kollman, P.A., J. Am. Chem. Soc., 117 (1995) 5179.
41. Bayly, C.A., Cieplak, P., Cornell, W.D. and Kollman, P.A., J. Phys. Chem., 97 (1993) 10269.
42. Vanommeslaeghe, K., Alsenoy, C.V., Proft, F.D., Martins, J.C., Tourwe, D. and Geerlings, P., Org. Biol. Chem., 1 (2003) 2951.
43. Case, D.A., Pearlman, D.A., Caldwell, J.W., Cheatham, T.E., Ross, W.S., Simmerling, C., Darden, T., Merz, K.M., Jr., Stanton, R.V., Cheng, A., Vincent, J.J., Crowley, M., Tsui, V., Radmer, R., Duan, Y., Pitera, J., Massova, I., Seibel, G.L., Singh, U.C., Weiner, P. and Kollman, P.A., AMBER 7, University of California, San Francisco, 2002.
44. Zwanzig, R.J., J. Chem. Phys., 22 (1954) 1420.
45. Beveridge, D.L. and DiCapua, F.M., Annu. Rev. Biophys. Biophys. Chem., 18 (1989) 431.
46. Bash, P.A., Singh, U.C., Brown, F.K., Langridge, R. and Kollman, P.A., Science, 235 (1987) 574.
47. Merz, K.M., Jr. and Kollman, P.A., J. Am. Chem. Soc., 111 (1989) 5649.
48. Rao, B.G., Tilton, R.F. and Singh, U.C., J. Am. Chem. Soc., 114 (1992) 4447.
49. Rastelli, G., Thomas, B., Kollman, P.A. and Santi, D.V., J. Am. Chem. Soc., 117 (1995) 7213.
50. Essex, J.W., Severance, D.L., Tirado-Rives, J. and Jorgensen, W.L., J. Phys. Chem. B., 101 (1997) 9663.
51. Reddy, M.R. and Erion, M.D., J. Am. Chem. Soc., 123 (2001) 6246.
52. Guimaraes, C.R.W. and Bicca de Alencastro, R., J. Med. Chem., 45 (2002) 4995.
53. Berendsen, H.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. and Haak, J.R., J. Chem. Phys., 81 (1984) 3684.
54. Ryckaert, J.P., Ciccotti, G. and Berendsen, H.C., J. Comput. Phys., 23 (1977) 327.
55. Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M., J. Appl. Crystallogr., 26 (1993) 283.
56. Munagala, N., Basus, V.J. and Wang, C.C., Biochemistry, 40 (2001) 4303.
57. Gill, H.S., Pfluegl, G.M.U. and Eisenberg, D., Biochemistry, 41 (2002) 9863.
58. Huntley, J.J.A., Scrofani, S.D.B., Osborne, M.J., Wright, P.E. and Dyson, H.J., Biochemistry, 39 (2000) 13356.

Supporting Information

The summary of docking simulation results with Auto-Dock for the top five conformational clusters of the three inhibitor scaffolds under consideration.

1. Scaffold 1

	1a			**1b**			**1d**	
Cluster rank	% population	Binding energy (kcal/mol)	Cluster rank	% population	Binding energy (kcal/mol)	Cluster rank	% population	Binding energy (kcal/mol)
1	21.1	−19.06	1	15.0	−19.13	1	34.8	−19.24
2	25.3	−18.43	2	18.8	−18.48	2	13.0	−18.32
3	10.8	−18.29	3	6.3	−18.28	3	4.3	−17.68
4	6.3	−17.93	4	20.3	−17.65	4	17.4	−17.11
5	15.8	−16.80	5	12.5	−17.01	5	8.7	−16.78

2. Scaffold 2

	2b			**2c**			**2d**	
Cluster rank	% population	Binding energy (kcal/mol)	Cluster rank	% population	Binding energy (kcal/mol)	Cluster rank	% population	Binding energy (kcal/mol)
1	44.8	−19.72	1	56.8	−23.63	1	68.5	−22.17
2	12.1	−19.31	2	18.9	−21.36	2	14.3	−20.88
3	31.0	−18.23	3	8.1	−20.51	3	3.8	−20.75
4	3.4	−18.05	4	3.5	−19.34	4	7.6	−20.30
5	1.8	−17.30	5	6.2	−18.16	5	5.7	−19.14

3. Scaffold 3

	3a			**3b**	
Cluster rank	% population	Binding energy (kcal/mol)	Cluster rank	% population	Binding energy (kcal/mol)
1	36.7	−19.80	1	11.4	−19.07
2	31.4	−18.85	2	13.6	−18.57
3	3.6	−17.67	3	15.9	−18.09
4	10.8	−17.52	4	2.3	−17.96
5	6.0	−16.55	5	6.8	−17.59