Original article

3D-QSAR studies of HDACs inhibitors using pharmacophore-based alignment

Yadong Chen [a], Huifang Li [a], Wanquan Tang [a], Chengchao Zhu [a], Yongjun Jiang [b], Jianwei Zou [b], Qingsen Yu [b], Qidong You [a,*]

[a] *Department of Medicinal Chemistry, China Pharmaceutical University, 24 Tongjiaxiang, Nanjing 210009, PR China*
[b] *Ningbo Institute of Technology, Zhejiang University, Ningbo 315104, PR China*

A B S T R A C T

Histone deacetylases (HDACs) enzyme is a promising target for the development of anticancer drugs. The enzyme-bound conformation of Trichostatin A (TSA) (PDB ID:1C3R) as an inhibitor of HDACs was used to manually construct a pharmacophore model. This model was then successfully used to identify the bioactive conformation and align flexible and structurally diverse molecules. Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) were performed on hydroxamate-based HDACs inhibitors based on phamacophore alignment. The best predictions were obtained with CoMFA standard model ($q^2 = 0.726$, $r^2 = 0.998$) and CoMSIA model combined with steric, electrostatic, hydrophobic, hydrogen bond donor and acceptor fields ($q^2 = 0.610$, $r^2 = 0.995$). Both of the models were validated by an external test set, which gave a satisfactory predictive r^2 value of 0.800 and 0.732, respectively. Graphical interpretation of the results revealed important structural features of the inhibitors related to the active site of HDACs. The results may be exploited for further design and virtual screening for some novel HDACs inhibitors.

1. Introduction

Histone deacetylases (HDACs) and histone acetyltransferases (HATs) are considered to play an important role in cell cycle control by acting as a transcriptional coactivator or a transcriptional corepressor. HAT-mediated hyperacetylation of positively charged lysine residues in the N-terminal tails of core histones loosens the histone–DNA binding and activates a gene transcription. In contrast, HDAC-catalyzed deacetylation of ε–N acetyl group of lysine residues leads to the tight histone–DNA binding, which restricts the access transcription factors. Perturbation of this balance is often observed in human cancers and inhibition of HDACs has emerged as a novel therapeutic strategy against cancer [1–6].

To date, some structurally distinct compounds targeting HDACs enzyme have been reported [7–11], and with few exceptions, they can be divided into several structural classes, including small-molecular hydroxamic acids, carboxylates, benzamides, electrophilic ketones, and cyclic peptides. Among them, hydroxamic acids were the first and largest class of HDACs inhibitors identified, typified by Trichostatin A (TSA) and suberoylanilide hydroxamic acid (SAHA), which are still rapidly growing. The X-ray crystal structure of HDLP (histone deacetylase-like protein), a bacterial HDAC homologue, complexed with the lead compound TSA, has been resolved and revealed a distinctive mode of protein–ligand interactions [12]. The HDACs catalytic domain consists of a narrow, tube-like pocket spanning the length equivalent to four- to six-carbon straight chains. A Zn^{2+} ion is positioned near the bottom of this enzyme pocket. Accordingly, the structure of the lead compound TSA, represented as hydroxamic acid class of HDACs inhibitors, might be divided into three molecular fragments, each of which interacts with a discrete region of the enzyme pocket. These fragments include a zinc binding group (ZBG), a cap group, and a linker connecting the ZBG and the cap at a proper distance. The crystal structures of human HDAC8 complexed with hydroxamates, reported recently [13,14], also supported such a binding mechanism of HDACs. This three-fragment concept has proven successful in developing structural analogues of TSA as potent HDACs inhibitors [15,16].

CoMFA [17] and CoMSIA [18] are the most popular three-dimensional quantitative structure–activity relationship (3D-QSAR) methods mainly because the graphic results of these models can provide a direct way to visualize the structure–activity relationship. Until now, only a few 3D-QSAR studies for HDACs inhibitors have been reported [19–21]. However, most of these models did not incorporate the structural information of the receptor and investigate the local physicochemical properties of three molecular fragments of inhibitors to their respective interaction region. In order to

gain further insights into the structural and chemical features required for HDACs inhibitory activity, we decided to create CoMFA and CoMSIA models of the HDACs using data accumulated for the hydroxamate-based inhibitors. Due to the flexibility and structural diversity of hydroxamic acid class of HDACs inhibitors, it is difficult to choose a suitable conformation to achieve a meaningful alignment in 3D-QSAR analysis. To effectively superimpose the flexible and structurally diverse molecules and to identify bioactive conformations of these molecules (the two crucial steps of a CoMFA and CoMSIA analysis), an enzyme-bound conformation of TSA determined by X-ray crystallography was firstly used to manually construct a pharmacophore model. A novel strategy, alignment of the best fitting Catalyst conformations (the assumed bioactive conformations) based on this pharmacophore model, was developed and successfully applied. A total of forty-three structurally diverse hydroxamate-based molecules were studied, and two highly predictive CoMFA and CoMSIA models were generated and evaluated. The derived 3D-QSAR CoMFA and CoMSIA models for different chemical series reported as HDACs inhibitors give insight into the influence of various interactive fields on the activity and thus aid in designing and forecasting the HDACs inhibitory activity of novel compounds.

2. Materials and methods

2.1. Molecular modeling

Molecular modeling studies reported here were performed on a Dell Precision 380 workstation with the RHEL 4.0 operating system using Accelrys Catalyst 4.10 [22] and Sybyl 6.9 [23] software packages. All CoMFA and CoMSIA calculations were carried out using Sybyl 6.9.

2.2. Construction of a structure-based pharmacophore model

Pharmacophore models, also termed as hypotheses, can be generated using Catalyst package. A Catalyst pharmacophore model consists of a 3D arrangement of a collection of features necessary for the biological activity of the ligands. Whereas some features (e.g., hydrogen bond acceptor or hydrogen bond donor) are defined as vectors, others (e.g., hydrophobic feature) are located at centroids of the corresponding (e.g., hydrophobic) ligand atoms. The features in Catalyst are associated with location constraints, displayed as colored spheres, which allow a certain spherical tolerance surrounding the ideal position of a particular feature in 3D space.

Using the crystal structure of TSA–HDLP complex (PDB code: 1C3R) as a starting point, the bound conformation of TSA was imported to Catalyst. Some selected features were mapped on the corresponding groups of TSA with the feature dictionary in Catalyst View Hypothesis Workbench. The location constraints for each of these features were chosen with the default parameters. At last, all selected features were merged into a single pharmacophore model.

2.3. Conformational model analysis

The conformational analysis was carried out with an energy window high enough to include the bioactive conformation. For the training and test sets molecules, conformational models representing their available conformational space were calculated. All molecules were built using the 2D and 3D sketcher of Catalyst. A conformational set was generated for each molecule using the poling algorithm and the "best-quality" option, based on CHARMm force field. All other parameters used were kept at their default settings. A maximum of 255 conformations within 20 kcal/mol in

energy from the global minimum was generated and saved. The molecules associated with their conformational models were mapped onto the pharmacophore model using the "best fit" option to obtain the bioactive conformation of each molecule.

2.4. Alignment rule

The alignment rule was defined by a chemical function-mapping method and therefore based on a geometric fit of the chemical functions of the molecules to the chemical features of the pharmacophore. All the molecules in the training and test sets were mapped simultaneously onto the pharmacophore model using "best fit" option. The conformation selected for each compound, which was assumed to be the bioactive conformation, corresponded to the conformation which best fit the pharmacophore model. The final aligned molecules were exported to Sybyl for CoMFA and CoMSIA analyses.

2.5. CoMFA/CoMSIA model generation

CoMFA was performed using the QSAR option of Sybyl. The steric and electrostatic field energies were calculated using the Lennard–Jones and the Coulomb potentials, respectively, with a $1/r$ distance-dependent dielectric constant in all intersections of a regularly spaced (0.2 nm) grid. The electrostatic fields were computed using Gasteiger–Huckel charge calculation methods. A sp^3 hybridized carbon atom with a radius of 1.53 Å and a charge of $+1.0$ was used as a probe to calculate the steric and electrostatic energies between the probe and the molecules using the Tripos force field. The standard parameters implemented in Sybyl 6.9 were used. The truncation for both the steric and the electrostatic energies was set to 30 kcal/mol. This indicates that any steric or electrostatic field value that exceeds this value will be replaced with 30 kcal/mol, thus makes a plateau of the fields close to the center of any atom.

The CoMSIA method involves a common probe atom and similarity indices calculated at regularly spaced grid intervals for the prealigned molecules. The similarity indices descriptors were derived with the same lattice box used in CoMFA. CoMSIA calculates hydrophobic, H-bond donor and acceptor fields in addition to steric and electrostatic fields. The distance dependence between the grid point and each atom of molecule was determined by Gaussian function through the similarity indices calculated at all grid points, and a default value of 0.3 was used as an attenuation factor.

The CoMFA and CoMSIA descriptors served as independent variables and pIC_{50} values as dependent variables in PLS regression analysis for deducing 3D-QSAR models. Normally cross-validation is used to check the predictivity of the derived model. The result of analysis corresponds to the regression equation with thousands of coefficients. The predictive values of models were evaluated using leave-one-out (LOO) cross-validation method. The number of components leading to the highest cross-validated r^2 and lowest standard error of estimate (SEE) was set as the optimum number of components (Nc) in PLS analysis. 2.0 kcal/mol was used as the threshold column filtering value in PLS analysis. A set of external test compounds with distinct chemical structures representing the two series of compounds was used to evaluate the CoMFA and CoMSIA models.

3. Results and discussion

3.1. Data sets

The selection of the compounds used for our 3D-QSAR study was done by considering the fact that compounds represent

structural diversity and a range of biological activities. The biological data for hydroxamate-based HDACs inhibitors were collected from the different literature [24–28]. Therefore, it is necessary to normalize the activity data from different sources and make them comparable. Because TSA was identified mainly as an inhibitor of HDACs and was taken as a standard control in every biological assay reported in the references, we took the IC_{50} value of compound TSA from MethylGene [25,26,28] as the standard to normalize the experimental data, which was also adopted by Xie et al. [29] in their QSAR study against HDACs. All of the molecular structures and activity data used for 3D-QSAR study are represented in Tables 1A and 1B, in which inhibitory activity for each compound has been expressed as the normalized IC_{50}. A diverse collection of eight molecules chosen at random was used as the test set from the data set, while the remaining 43 compounds were treated as a training set.

3.2. Pharmacophore model

In the present study, the success of the CoMFA and CoMSIA methods is completely determined by the bioactive conformations and the quality of the alignment of the investigated molecules. Therefore, the first step in our 3D-QSAR study is the generation of a useful and reliable pharmacophore model. A pharmacophore model can be defined as a specific 3D arrangement of chemical groups common to active molecules and essential to their biological activities. Generally, pharmacophore models are generated automatically by HypoGen, HipHop, or manually based on a known conformation of a lead compound in Catalyst package. The third, so-called structure-based approach, was used in this study.

A crystal complex of a lead compound bound to the active site of a protein is sufficient information to start the construction of a structure-based pharmacophore model.

The binding mode of TSA to its target was already well known from previous crystallographic studies [12]. TSA as a lead compound against HDACs is possessed of the appropriate chemical features for binding and has the correct shape to fit into the active site. A pharmacophore model was built manually by mapping the chemical features on the corresponding groups of TSA in its bound conformation. TSA contains 22 features that are close enough for complementary functionalities in the enzyme to participate in binding interactions. If all of these features were included in a single pharmacophore model, the model would be too restrictive to fit structurally diverse molecules. Therefore, we selected a subset of these features including two hydrogen bond acceptors and three hydrophobic features (Fig. 1a), which focus on the three moieties of TSA according to the three-fragment concept.

The selection of chemical features was based on structural observations about parts (atom, functional groups) of TSA which are common for hydroxamic acid class of inhibitors and essential for activity. These important features could be identified from the interaction information of the TSA–HDLP complex, which reflect the primary interactions between the protein and ligand. An essential part of the pharmacophore is the ZBG (hydroxamic acid), represented by two hydrogen bond acceptors (HBA1 and HBA2 in Fig. 1a). Though the metal-binding function is not included in the Catalyst feature dictionary, it can be replaced by a standard hydrogen bond acceptor or donor function [30]. The hydrophobic interactions between ligand and HDACs enzyme may equally play a crucial role. The hydrophobic feature (Hydro1 in Fig. 1a), located at 4-methyl group on TSA, is used to model the hydrophobic interaction between the linker of TSA and the two juxtaposed phenyl moieties of Phe141 and Phe-198 at the active site, the other two hydrophobic features (Hydro2 and Hydro3 in Fig. 1a) mapped to the cap groups (dimethylamino and phenyl groups, respectively)

of TSA, which make contacts on the surface of the active pocket. The location constraints for each of these features were chosen with the default parameters.

At last, a pharmacophore model, which represents a defined binding mode of hydroxamic acid class of inhibitors to HDACs, was developed (Fig. 1b) and highlighted the necessary 3D arrangement of features that are required for binding. The pharmacophore model based on the bound conformation of TSA was mapped to the active site of the HDLP crystal structure (Fig. 1c), which is in good agreement with the structural information derived from X-ray structure of the binding domain of HDLP. We have recently defined a 3D-QSAR pharmacophore for HDACs inhibitors by HypoGen [31], which also substantiates this similar pharmacophore model.

3.3. Bioactive conformation determination

Choice of the assumed bioactive conformation for each compound has been the most difficult by far, especially when dealing with highly flexible compounds. In some cases, there is no right way to determine the bioactive conformations of some molecules. However, a reliable pharmacophore model may be used to determine the bioactive conformations of the ligands that share the same binding mode. The conformation selected for each compound, assumed as the bioactive conformation, corresponds to the conformation which best fits the pharmacophore.

To verify whether the method finds the correct bioactive conformation, we applied the method to two structurally different compounds, whose bioactive conformations are known from X-ray structures of the ligand–enzyme complex. SAHA, which represents flexible HDACs inhibitors containing aliphatic linker, and NHB (N-Hydroxy-4-(Methyl{[5-(2-Pyridinyl)-2-Thienyl]Sulfonyl}Amino)-Benzamide), which represents relatively rigid HDACs inhibitors containing aromatic linker, have the crystal structures in complex with human HDAC8 (PDB ID:1T69 and 1W22, respectively). Their bound conformations were respectively mapped onto the pharmacophore model using the "best fit "option and meanwhile superimposed to the best fitting Catalyst conformations (Fig. 2). The root mean square distance (RMSD) between the heavy atom positions of the bound and the best fitting Catalyst conformation were 1.51 Å and 1.13 Å for SAHA and NHB, respectively. It is important to note that the best fitting Catalyst conformations of the flexible ligands are in good agreement with their bioactive conformations. Hence, the results showed that Catalyst's conformational model generation is capable of reproducing the bioactive conformation from the Protein Data Bank. These findings support our choice for the bioactive conformation obtained from the best fitting Catalyst conformation of the calculated ensemble rather than the commonly used the energy minimization and docking approaches.

3.4. Alignment of molecules in the training and test sets

One of the most fundamental problems, when trying to develop a good and predictive 3D-QSAR model, is how to align the investigated compounds. This becomes especially critical when one is dealing with a set of structurally flexible and diverse compounds. Considerable effort has been devoted to investigate and develop better protocols for alignment of compounds. There are several methods available, such as cocrystallized conformation-based alignment, docked conformation-based alignment, and global minimum energy conformation-based alignment. A pharmacophore model also constitutes a useful tool to guide the alignment of compounds in 3D-QSAR CoMFA/CoMSIA study. Compared with the scaffold alignment based on the atom RMS fitting, which is commonly used in CoMFA/CoMSIA study, the pharmacophore-based alignment approach is more advantageous in aligning

Table 1A
Structures and normalized inhibitory activity (IC$_{50}$) of compounds investigated hydroxamate-based HDACs inhibitors containing aliphatic chain linker
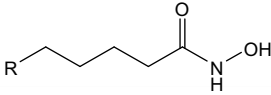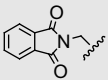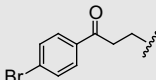

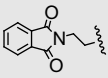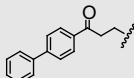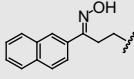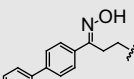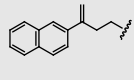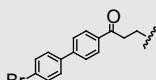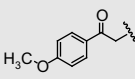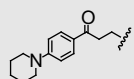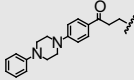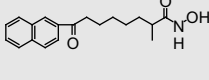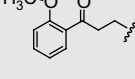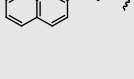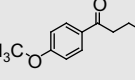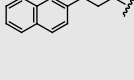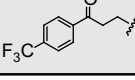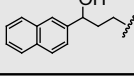| cmpd.	R	IC$_{50}$ (µM)	Ref.	cmpd.	R	IC$_{50}$ (µM)	Ref.
1		0.103	[24]	13		0.045	[25]
2		0.0714	[24]	14		0.005	[25]
3		1.5	[25]	15		0.0085	[25]
4		0.5	[25]	16		0.004	[25]
5		0.065	[25]	17		0.008	[25]
6		0.135	[25]	18		0.002	[25]
7		0.45	[25]	19		0.0095	[25]
8		0.035	[25]	20		0.0045	[25]
9		0.153	[25]	21		0.6	[25]
10		0.095	[25]	22		0.006	[25]
11		0.015	[25]	23		0.025	[25]
12		0.045	[25]	24		0.035	[25]

Table 1B
Structures and normalized inhibitory activity (IC$_{50}$) of compounds investigated hydroxamate-based HDACs inhibitors containing aromatic linker

cmpd.	R$_1$	R$_2$	R$_3$	IC$_{50}$ (μM)	Ref.
25	–	H	(phenyl)	0.9	[26]
26	–CH$_2$–	H	(phenyl)	1	[26]
27	–(CH$_2$)$_2$–	H	(phenyl)	0.1	[26]
28	–(CH$_2$)$_3$–	H	(phenyl)	1	[26]
29	–CH=CH–	H	(phenyl)	0.2	[26]
30	(isobutenyl)	H	(phenyl)	2	[26]
31	(isobutyl)	H	(phenyl)	17	[26]
32	–CH=CH–	CH$_3$	(phenyl)	0.6	[26]
33	–CH=CH–	H	4-Cl-phenyl	0.075	[26]
34	–CH=CH–	H	3,4-diCl-phenyl	0.1	[26]
35	–CH=CH–	H	2,4-diCl-phenyl	0.3	[26]
36	–CH=CH–	H	4-F$_3$CO-phenyl	0.6	[26]
37	–CH=CH–	H	3-F$_3$C-phenyl	0.1	[26]
38	–CH=CH–	H	4-H$_3$C-phenyl	0.3	[26]

Table 1B (continued)

cmpd.	R$_1$	R$_2$	R$_3$	IC$_{50}$ (μM)	Ref.
39	–CH=CH–	H	3-H$_3$C-phenyl	0.1	[26]
40	–CH=CH–	H	4-tBu-phenyl	0.1	[26]
41	–CH=CH–	H	2,4,6-triisopropyl-phenyl	0.6	[26]
42	–CH=CH–	H	4-H$_3$CO-phenyl	0.06	[26]
43	–CH=CH–	H	3,4-di(H$_3$CO)-phenyl	0.09	[26]

cmpd.	R$_4$	X	IC$_{50}$ (μM)	Ref.
44	H	CH$_2$–CH$_2$	52.56	[27]
45	H	CH=CH	12.25	[27]
46	Cl	CH=CH	9.65	[27]
47	CH$_3$	CH=CH	11.25	[27]
48	(furan-based structure)		2.4	[28]
49	(chlorophenyl urea structure)		0.84	[28]

cmpd.	R$_5$	IC$_{50}$ (μM)	Ref.
50	(phenylsulfonyl)	2.0	[28]
51	(phenyl)	1.0	[28]

flexible and diverse molecules. Catalyst aligns molecules according to their chemical functionality, so structurally different molecules that have the same 3D chemical features can be aligned together without necessity of atom-by-atom comparison. This method provides a different but straightforward way of representing ligand–receptor interactions that avoids the bias that can be introduced by relying on atom-based alignment. Molecules that belong to different structural families can be considered conveniently because different chemical groups or atoms can belong to the definition of a given function. If a predictive model can be successfully developed, then this may serve as an indication that the choice of alignment scheme was a reasonable one.

Fig. 3 shows an alignment of all molecules in the training and test sets by the pharmacophore model. It seemed that the alignment was perfect when the bioactive conformations were automatically
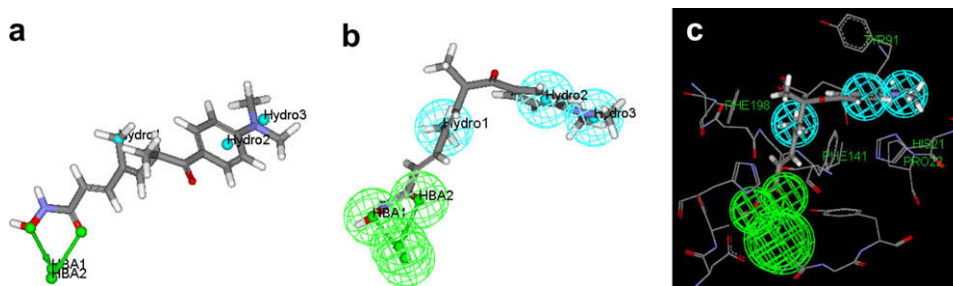
Fig. 1. (a) The enzyme-bound conformation of TSA, populated with Catalyst chemical features. Blue dots represent centers of hydrophobic regions; green arrows represent hydrogen bond acceptors. The arrows locate the heavy atom (tail) and the position on the enzyme participating in the interaction (head) and also provide a visual sense of directionality. (b) The pharmacophore model based on the bound conformation of TSA. The location constraints (dots and arrows) represent the ideal relative positions in space for the chemical functions, and the tolerance spheres (colored spheres) represent relative areas around the location constraints where chemical functions can still participate in mapping to the pharmacophore. (c) The pharmacophore model mapped to an HDLP crystal structure. [For interpretation of color in this figure legend the reader is referred to web version of the article.]

aligned to the pharmacophore model. From the chemical intuition, the three common core structures of hydroxamic acid derivates should adopt similar orientation. In the investigation of the alignment of molecules, one criterion was adopted. The CO and OH groups in hydroxamic acid moiety should superimpose or locate near the two hydrogen bond acceptor features to force all compounds to take similar space orientations, which represents that the ZBG could coordinate with the zinc ion at the active site. For these compounds with good inhibitory activities, they can produce good fits with all five features in the pharmacophore model. While for those compounds with poor inhibitory activity, they can only produce relatively good fits with three or four features. If we carefully investigate the alignment between the pharmocophore model and the studied compounds, we can find that the alignment based on the pharmacophore model is obviously reasonable for all of the compounds.

3.5. 3D-QSAR models

The CoMFA and CoMSIA descriptors were calculated and 3D-QSAR models were analyzed and optimized using the partial least squares (PLS) method based on these descriptors. The statistical results of CoMFA and CoMSIA studied are summarized in Table 2. These analyses were based on the clusters of molecules that were aligned by pharmacophore model. The regression coefficient (r^2) and the cross-validation coefficient (q^2) of the 3D-QSAR model constructed by CoMFA are 0.998 and 0.726, respectively. The two coefficients for the CoMSIA model are 0.995 and 0.610, respectively. The q^2 and r^2 values of the 3D-QSAR models indicate a statistically significant and stable model. In terms of the coefficient values, the CoMFA model appears slightly better than CoMSIA model. Both models exhibit good predictive capabilities as shown by the cross-validation method. The predicted vs experimental pIC$_{50}$ values for the training and test set obtained by the two methods are shown in

Table 3. In both cases, the predicted values do not deviate significantly from the experimental IC$_{50}$ values (generally with less than one logarithmic unit). The predicted vs experimental pIC50 values for the training set are graphically depicted in Fig. 4. Compared with others' work based on hydroxamic acid class of HDACs inhibitors [19,20], the pharmacophore-based alignment produced better results than the traditional atom fit and docking alignment.

Besides internal validation with the training set, we used an external test set of eight compounds to validate the predictive power of the derived models. The external validation process can be considered the most valuable validation method, as these compounds are completely excluded during the training of the model. Prior to prediction, the test set compounds were processed identically to the training set compounds, as described in the Methods section. All of the test set compounds, which represent the different structural diversity involving in the training set, are well predicted without any apparent outliers. The chosen external test set yielded a predictive r^2 of 0.800 and 0.732 for CoMFA and CoMSIA models, respectively. A Standard error value of test set is 0.364 and 0.362, respectively. These values indicated a good predictive power and were in agreement with the statistical values from the internal validation procedures. The predicted vs experimental pIC$_{50}$ values for the test set compounds are listed and plotted in Table 3 (marked with an *) and Fig. 4, respectively. The high predictive power of CoMFA and CoMSIA models for sets of structurally diverse compounds suggested that these models possess a high accommodating capacity.

3.6. Interpretation of the 3D-QSAR maps with respect to ligand binding domain

The statistically relevant results from the 3D-QSAR analysis are visualized as 3D contour maps. The coefficient contour maps and
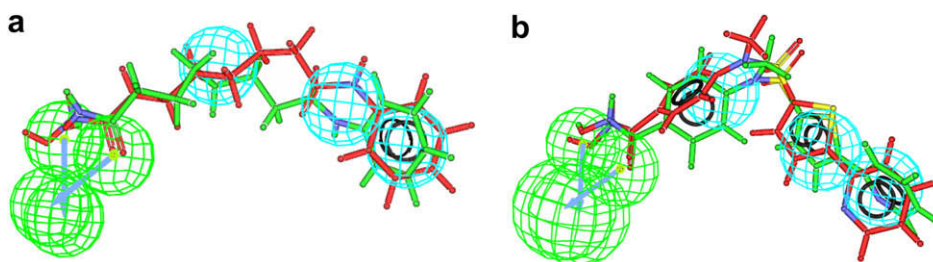

Fig. 2. The best fitting Catalyst conformation (green stick model) and the bound conformation in the crystal structure (red stick model) are superimposed on the pharmacophore model. Comparison of both conformations of compounds SAHA (a) and NHB (b), respectively. [For interpretation of color in this figure legend the reader is referred to web version of the article.]
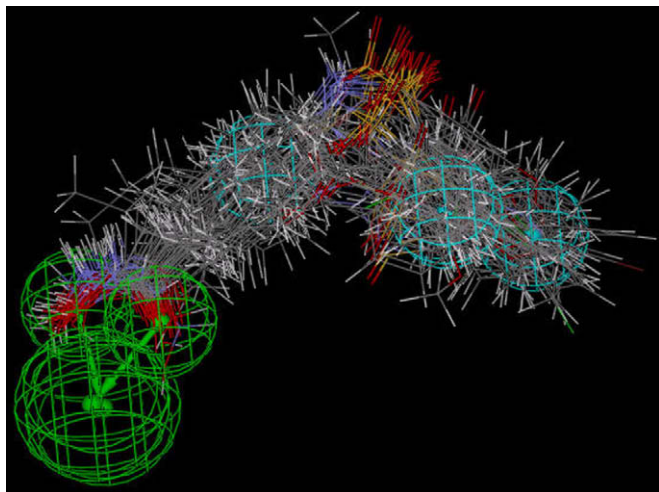
Fig. 3. Molecular alignment used in the present study, obtained from the pharmacophore model by a chemical function-mapping method with best fit option.

the active site of the HDLP (pdb ID:1c3r) were superimposed and compared with each other. It provides an opportunity to interpret features indicated in the contour maps with respect to the receptor environment. The electrostatic and steric contour maps from the CoMSIA model are in accordance with the field distributions of CoMFA maps, and thus the contour maps of CoMSIA model are only displayed with the binding site residues and with the alignment conformation of the most active compound 18 as a reference structure in Fig. 5. For the analysis of the CoMFA and CoMSIA contour maps, we divided roughly the HDACs inhibitor into three sections to further explore the hypothetical interaction features of a ligand with its receptor: the ZBG domain, the linker domain, and the cap group domain.

3.6.1. The ZBG domain

The hydroxamic acid functional group is important for enzyme inhibitory activity, as it coordinates the zinc ion through CO and OH groups at the active site. It also forms hydrogen bonds between its NH and His131/Asp166, OH groups and His132/Asp173, and its CO and the Tyr297 hydroxyl group, respectively. There are three hydrogen bond donor favorable cyan contours around the hydroxamic acid

Table 2
Summary of CoMFA and CoMSIA statistics for the training and test sets.

	CoMFA	CoMSIA
q^2 (Leave-One-Out)	0.726	0.610
r^2	0.998	0.995
F values	2913.973	1197.780
SEE[a]	0.051	0.080
Nc[b]	6	6
Field Contribution		
steric	0.562	0.278
electrostatic	0.438	0.193
hydrophobic		0.167
HB donor		0.192
HB acceptor		0.171
Test Set		
No. of compounds	8	8
r^2 of test set	0.800	0.732
Standard error of test set	0.364	0.362
P value	0.00270	0.00673

[a] Standard error of estimate.
[b] Optimum number of components.

Table 3
Experimental and predicted activities (pIC_{50}) of the training and test set compounds by CoMFA and CoMSIA models.

cmpd.[a]	Experimental pIC_{50}	CoMFA model		CoMSIA model	
		Predicted pIC_{50}	Residual	Predicted pIC_{50}	Residual
1	6.99	6.94	0.05	6.92	0.07
2	7.15	7.14	0.01	7.24	−0.09
3	5.82	5.86	−0.04	5.86	−0.04
4	6.30	6.29	0.01	6.27	0.03
5	7.19	7.24	−0.05	7.23	−0.04
6*	6.87	6.02	0.85	6.48	0.39
7	6.35	6.33	0.02	6.29	0.06
8*	7.46	7.46	0.00	7.17	0.29
9	6.82	6.86	−0.04	6.79	0.03
10	7.02	7.05	−0.03	7.06	−0.04
11	7.82	7.88	−0.06	7.82	0.00
12	7.35	7.39	−0.04	7.43	−0.08
13*	7.35	6.91	0.44	7.19	0.16
14	8.30	8.31	−0.01	8.25	0.05
15	8.07	8.09	−0.02	8.12	−0.05
16	8.40	8.38	0.02	8.38	0.02
17	8.10	8.05	0.05	8.09	0.01
18	8.70	8.73	−0.03	8.66	0.04
19	8.02	7.97	0.05	8.04	−0.02
20	8.35	8.32	0.03	8.38	−0.03
21	6.22	6.23	−0.01	6.24	−0.02
22	8.22	8.12	0.10	8.08	0.14
23	7.60	7.64	−0.04	7.70	−0.10
24*	7.46	7.10	0.36	6.87	0.59
25	6.05	6.03	0.02	6.03	0.02
26	6.00	6.01	−0.01	6.03	−0.03
27	7.00	7.01	−0.01	7.08	−0.08
28	6.00	6.00	0.00	5.95	0.05
29	6.70	6.66	0.04	6.55	0.15
30	5.70	5.67	0.03	5.63	0.07
31	4.77	4.75	0.02	4.80	−0.03
32	6.22	6.29	−0.07	6.33	−0.11
33	7.12	7.04	0.08	6.94	0.18
34	7.00	6.94	0.06	6.91	0.09
35	6.52	6.48	0.04	6.54	−0.02
36	6.22	6.23	−0.01	6.24	−0.02
37*	7.00	7.12	−0.12	6.95	0.05
38	6.52	6.63	−0.11	6.57	−0.05
39	7.00	6.93	0.07	6.82	0.18
40	7.00	7.06	−0.06	7.02	−0.02
41*	6.22	6.45	−0.23	6.68	−0.46
42	7.22	7.32	−0.10	7.35	−0.13
43	7.05	7.03	0.02	7.15	−0.10
44	4.28	4.28	0.00	4.32	−0.04
45	4.92	4.87	0.05	4.87	0.05
46	5.01	5.04	−0.03	5.03	−0.02
47*	4.95	5.19	−0.24	5.17	−0.22
48	5.62	5.59	0.03	5.60	0.02
49	6.08	6.06	0.02	6.09	−0.01
50*	5.70	6.04	−0.34	6.63	−0.93
51	6.00	6.08	−0.08	6.06	−0.06

[a] The test set compounds are labeled with an '*' symbol.

functional group in hydrogen bond donor contour map (Fig. 5d). The result is consistent with the observation from X-ray crystallographic studies. Two large white contours representing unfavorable hydrophobic interaction are seen at the ZBG domain in the hydrophobic contour map (Fig. 5c). This suggests that the hydrophobic groups in the domain may be unfavorable for the coordinating interaction between the zinc ion and hydroxamic acid group.

3.6.2. The linker domain

The linker of HDACs inhibitors locates at the tube-like hydrophobic pocket which is surrounded by Phe-198 and Phe141 at the active site. A hydrophobic favorable yellow contour in hydrophobic contour map (Fig. 5c) and the steric favorable green contours in the steric contour map (Fig. 5a) are observed in the linker domain, respectively. The receptor and 3D-QSAR model consistently
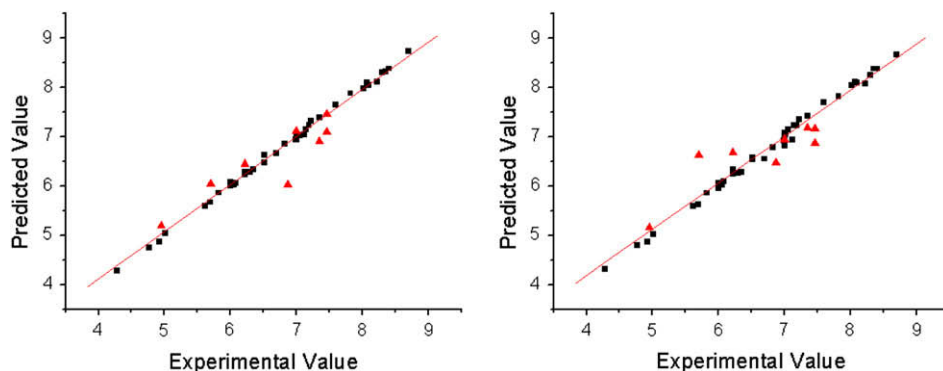
Fig. 4. Graphs of experimental value vs predicted values for training and test set compounds. (a) CoMFA, (b) CoMSIA (■ training set; ▲ test set).
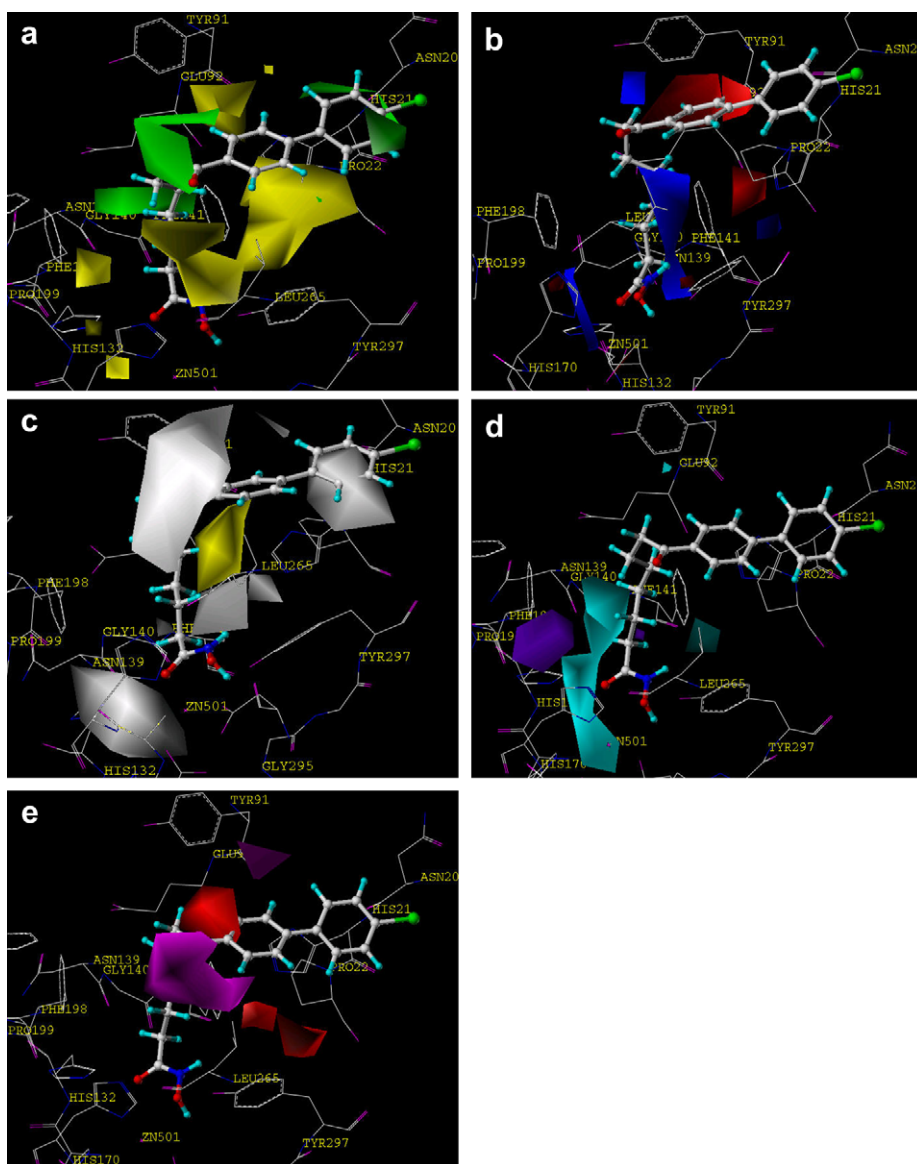

Fig. 5. Contour maps of the pharmacophore-based CoMSIA model (StDev × Coeff) superimposed into the active binding site of HDLP. Favored and disfavored contour levels for CoMSIA were fixed at 80% and 20%, respectively, for all contour maps. The most active compound 18 as a reference structure is shown in stick model. (a) CoMSIA steric contour map. Green indicates regions where bulky groups increase activity; yellow indicates regions where bulky groups lead to decreasing activity. (b) CoMSIA electrostatic contour map. Blue indicates regions where positive charges increase activity or negative charges decrease it; red indicates region where negative charges increase activity or positive charges decrease it. (c) CoMSIA hydrophobic contour map. Yellow indicates regions where hydrophobic groups enhance activity; white indicates regions where hydrophilic groups decrease activity. (d) CoMSIA Hydrogen bond donor contour map: cyan indicates regions where hydrogen bond donor groups enhance activity; purple indicates regions where hydrogen bond donor groups decrease activity. (e) CoMSIA Hydrogen bond acceptor contour map: magenta indicates regions where hydrogen bond acceptor groups enhance activity; red indicates regions where hydrogen bond acceptor groups decrease activity. [For interpretation of color in this figure legend the reader is referred to web version of the article.]

indicate the hydrophobic and steric interactions between the linker of inhibitors and the tube-like pocket at the active site were beneficial to inhibitory activity. In addition, the steric unfavorable yellow contours lie close to the carbon next to the hydroxamic acid functional group (Fig. 5a), suggesting that any bulky substituted group at this position is unfavorable for inhibitory activity. This is in agreement with the fact that the inhibitory activities of compounds 30 and 31 with more bulky groups in the carbon next to the hydroxamic acid functional group are low. The linker fell into a positive charge favorable blue contour (Fig. 5b), which indicated that the linkers with low electron density were preferred in the tube-like pocket. This can be attributed to the partial electrostatic compensation from the phenyl groups (Phe-198 and Phe141) with high electron density. In summary, these findings indicate that the linkers composed of bulky, electropositive, and hydrophobic molecular fragments may increase the activity.

3.6.3. The cap group domain

The cap group of HDACs inhibitors makes good contact with the residues in the rim space at the active site and is an extremely variable moiety. The hydrophobic unfavorable white contours (Fig. 5c) near the His21 and Pro22 indicate that hydrophilicity of the cap group would contribute to the inhibitory activity in this region. It is worth noting that there is not a hydrophobic favorable yellow contour in the cap group domain. This finding indicates that the introduction of appropriate hydrophilic groups in the cap group domain is favorable for inhibitory activity. This may be because the cap groups are exposed to the solvent. Another hydrophobic unfavorable white contour (Fig. 5c) lies close to the entrance of the tube-like pocket, where the hydrophobic interaction makes inhibitor unfavorably enter the tube-like hydrophobic pocket and is unbeneficial to inhibitory activity. The steric yellow unfavorable contours (Fig. 5a) in vicinity of the Glu92 suggest that the bulky groups will conflict with residues in this region and are unfavorable for steric interactions. There are two steric green favorable contours (Fig. 5a) around the His21. The green contours indicate the bulky group in the region is favorable for steric interactions and beneficial to inhibitory activity. In addition, the magenta contours (Fig. 5e) around the entrance of tube-like pocket indicate that hydrogen bond acceptor is favored in this region.

4. Conclusions

In our study, a successful strategy of pharmacophore-based alignment of the best fitting Catalyst conformations was designed and applied to build 3D-QSAR model. Firstly, a chemical feature-based pharmacophore model was manually generated based on the bound conformation of a potent inhibitor TSA. We used the pharmacophore model to identify the bioactive conformations, and to superimpose the structurally diverse hydroxamate-based compounds. Two highly predictive and statistically significant CoMFA and CoMSIA models were derived from pharmacophore-based alignment. The study produced equally good models in terms of the q^2 values. The predictive powers of the derived models were demonstrated to be reliable.

The CoMFA model provided the most significant correlation of steric and electrostatic fields with the inhibitory activities. Furthermore, the CoMFA method provided better statistical models than CoMSIA, which suggested that that steric and electrostatic interactions play important roles in the binding of the hydroxamate-based inhibitors at the HDACs active site. The CoMSIA steric and electrostatic contour maps are in accordance with field distribution of CoMFA contour maps and consistent with structure–activity relationships. The comparison of 3D-QSAR models with the active site of HDACs revealed the interactions of amino acid residues with the three fragments of HDACs inhibitors with steric, electrostatic, hydrophobic, hydrogen bond fields around them. Such pharmacophore model may be applied to virtually screen commercial database to find some hits and then calculate the activities of some novel hits using CoMFA and CoMSIA models. The identified possible lead compounds will be valuable for further synthesis.

References

[1] P.A. Wade, Hum. Mol. Genet. 10 (2001) 693–698.
[2] W.D. Cress, E. Seto, J. Cell Physiol. 184 (2000) 1–16.
[3] P. Marks, R.A. Rifkind, V.M. Richon, R. Breslow, T. Miller, W.K. Kelly, Nat. Rev. Cancer 1 (2001) 194–202.
[4] C.M. Grozinger, S.L. Schreiber, Chem. Biol. 9 (2002) 3–16.
[5] R.W. Johnstone, Nat. Rev. Drug Discov. 1 (2002) 287–299.
[6] O.H. Kramer, M. Gottlicher, T. Heinzel, Trends Endocrinol. Metab. 12 (2001) 294–300.
[7] M. Jung, Curr. Med. Chem. 8 (2001) 1505–1511.
[8] J.S. Chen, D.V. Faller, R.A. Spanjaard, Curr. Cancer Drug Targets 3 (2003) 219–236.
[9] T. Suzuki, T. Ando, K. Tsuchiya, N. Fukazawa, A. Saito, Y. Mariko, T. Yamashita, O. Nakanishi, J. Med. Chem. 42 (1999) 3001–3003.
[10] R.B.S. Breslow, L. Gershell, T.A.Miller, P.A. Marks, et al., Class of cytodifferentiating agents and histone deacetylase inhibitors, and methods of use thereof, U.S. Patent 6,511,990B1, 2003.
[11] J.L. Marshall, N. Rizvi, J. Kauh, W. Dahut, M. Figuera, M.H. Kang, W.D. Figg, I. Wainer, C. Chaissang, M.Z. Li, M.J.J. Hawkins, Exp. Ther. Oncol. 2 (2002) 325–332.
[12] M.S. Finnin, J.R. Donigian, A. Cohen, V.M. Richon, R.A. Rifkind, P.A. Marks, R. Breslow, N.P. Pavletich, Nature 401 (1999) 188–193.
[13] J.R. Somoza, R.J. Skene, B.A. Katz, C. Mol, J.D. Ho, A.J. Jennings, C. Luong, A. Arvai, J.J. Buggy, E. Chi, J. Tang, B.C. Sang, E. Verner, R. Wynands, E.M. Leahy, D.R. Dougan, G. Snell, M. Navre, M.W. Knuth, R.V. Swanson, D.E. McRee, L.W. Tari, Structure 12 (2004) 1325–1334.
[14] A. Vannini, C. Volpari, G. Filocamo, E.C. Casavola, M. Brunetti, D. Renzoni, P. Chakravarty, C. Paolini, R. De Francesco, P. Gallinari, C. Steinkuhler, S. Di Marco, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 15064–15069.
[15] M. Jung, G. Brosch, D. Kolle, H. Scherf, C. Gerhauser, P.J. Loidl, Med. Chem. 42 (1999) 4669–4679.
[16] S.M. Sternson, J.C. Wong, C.M. Grozinger, S.L. Schreiber, Org. Lett. 3 (2001) 4239–4242.
[17] R.D. Cramer, D.E. Patterson, J.D. Bunce, J. Am. Chem. Soc. 110 (1988) 5959–5967.
[18] G. Klebe, U. Abraham, T. Mietzner, J. Med. Chem. 37 (1994) 4130–4146.
[19] D.C. Juvale, V.V. Kulkarni, H.S. Deokar, N.K. Wagh, S.B. Padhye, V.M. Kulkarni, Org. Biomol. Chem. 4 (2006) 2858–2868.
[20] Y. Guo, J. Xiao, Z. Guo, F. Chu, Y. Cheng, S. Wu, Bioorg. Med. Chem. 13 (2005) 5424–5434.
[21] R. Ragno, S. Simeoni, S. Valente, S. Massa, A.J. Mai, Chem. Inf. Model 46 (2006) 1420–1430.
[22] Catalyst, Version 4.10 (Software Package), Accelrys, Inc., San Diego,, 2004.
[23] Sybyl Molecular Modeling Software Packages; ver.7.0, TRIPOS, Associates, Inc., St Louis, MO 63144, 2004.
[24] S.W. Remiszewski, L.C. Sambucetti, P. Atadja, K.W. Bair, W.D. Cornell, M.A. Green, K.L. Howell, M. Jung, P. Kwon, N. Trogani, H. Walker, J. Med. Chem. 45 (2002) 753–757.
[25] S.H. Woo, S. Frechette, E. Abou Khalil, G. Bouchain, A. Vaisburg, N. Bernstein, O. Moradei, S. Leit, M. Allan, M. Fournel, M.C. Trachy-Bourget, Z. Li, J.M. Besterman, D. Delorme, J. Med. Chem. 45 (2002) 2877–2885.
[26] R. Lavoie, G. Bouchain, S. Frechette, S.H. Woo, E. Abou-Khalil, S. Leit, M. Fournel, P.T. Yan, M.C. Trachy-Bourget, C. Beaulieu, Z. Li, J. Besterman, D. Delorme, Bioorg. Med. Chem. Lett. 11 (2001) 2847–2850.
[27] Silvio Massa, A. Mai, Gianluca Sbardella, Monica Esposito, Rino Ragno, Peter Loidl, Gerald Brosch, J. Med. Chem. 44 (2001) 2069–2072.
[28] Delorme Daniel, Ruel Rejean, Lavoie Rico, Thibault Carl, Abou-Khalil, Elie. Inhibitors of Histone Deacetylase, MethylGene, Inc. Int. Patent Appl. WO 01/38322 2001.
[29] A. Xie, C. Liao, Z. Li, Z. Ning, W. Hu, X. Lu, L. Shi, J. Zhou, Curr. Med. Chem. Anticancer Agents 4 (2004) 273–299.
[30] H. Li, J. Sutter, R. Hoffman, HypoGen: an automated system for generating 3D predictive pharmacophore models, Pharmacophore Perception, Development, and Use in Drug Design, International University Line, La Jolla, CA, 2000.
[31] Ya-Dong Chen, Yong-Jun Jiang, Jian-Wei Zhou, Qing-Sen Yu, Qi-Dong You, J. Mol. Graph. Model 26 (2008) 1160–1168.