

3D-QSAR Study on Apicidin Inhibit Histone Deacetylase

CHEN, Hai-Feng^a(陈海峰) KANG, Jiu-Hong^b(康九红) LI, Qiang^a(李强)
ZENG, Bao-Shan^a(曾宝珊) YAO, Xiao-Jun^a(姚小军) FAN, Bo-Tao^a(范波涛)
YUAN, Shen-Gang^{*a}(袁身刚) Panay, A.^c Doucet, J. P.^c

^aKey Laboratory of Computer Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China

^bSchool of Life Science, Lanzou University, Lanzou, Gansu 730000, China

^cITODYS, CNRS UMR 7086, Université Paris 7, 1, rue Guy de la Brosse, 75005 Paris, France

For Histone Deacetylase (HDAC) Inhibitor, four 3D-QSAR models for four types of different activities, were constructed. The cross-validated q^2 value of CoMFA Model 1 is 0.624 and the noncross-validated r^2 value is 0.939. The cross-validated q^2 value of Model 2 for training set is 0.652 and the noncross-validated r^2 value is 0.963. The cross-validated q^2 value for Model 3 is 0.713, with noncross-validated r^2 value 0.947. The cross-validated q^2 value for Model 4 is 0.566 with noncross-validated r^2 value 0.959. Their predicted abilities were validated by different test sets which did not include in training set. Then the relationship between substituents and activities was analyzed by using these models and the main influence elements in different positions (positions 8 and 14) were found. The polar donor electron group of position 8 could increase the activity of inhibition of HDAC, because it could form chelation with the catalytic Zn. Suitable bulk and positive groups at position 14 are favorable to anti-HDAC activity. These models could well interpret the relationship between inhibition activity and apicidin structure affording us important information for structure-based drug design.

Keywords histone deacetylase inhibitor, CoMFA, 3D-QSAR

Introduction

In eukaryotes, the formation of nucleosomes through DNA binding to core histones and the folding of chromosomes pose tricky obstacles to DNA regulators (for example, transcription factors and RNA polymerase) since both individual nucleosomes and higher order structures of chromosomes limit the access of regulators to their corresponding nucleosomal DNA. Modification of this access generally leads to certain alterations of gene expression, and hence cells behaviors.¹⁻³

In vivo, the modification of histone acetylation plays a critical role in the regulation of nucleosomal DNA access to its regulators. Generally, histone hypoacetylation leads to less accessible DNA and gene suppression, whereas his-

tone hyperacetylation results in more accessible DNA and gene activation.⁴⁻⁶ The state of histone acetylation is regulated by two kinds of competing enzymes, histone acetyltransferases (HAT) and histone deacetylases (HDAC), to keep a fine equilibrium of these two enzymes is important for the normal gene expression and cell cycle, loss of the equilibrium inevitably leads to deranged expression of certain genes, and hence involves in the pathogenesis of diverse diseases.⁷⁻¹³ Among the diseases related with histone acetylation abnormalities, a connection between cancer and histone-modifying enzymes has been clearly shown, and since HDAC are key elements of the multisubunit repressor complexes that lead to the repression of chromatin and silencing of complete sets of target genes, they are considered to be important contributors to malignant transformation.¹⁴⁻¹⁷ In this context, HDAC inhibitors have been proved to be a class of potential anticancer drugs.^{14,17,18}

HDAC inhibitors exhibit their antitumor effects through inducing histone hyperacetylation in chromatin, which in turn usually leads to the relief of the transcriptional repression for a certain subset of genes. Among these are genes encoding for proteins that are crucial for the regulation of cell proliferation, differentiation, or apoptosis, *e.g.*, the cyclin-dependent kinase inhibitor genes p21/WAF1/CIP1, p53, tumor angiogenesis-related genes, telomerase reverse transcriptase mRNA and others.^{16,19-23} Importantly, these inhibitors have demonstrated potential for the prevention and treatment of cancer not only in numerous cell culture²⁴ and animal models,²⁵ but also on humans.^{26,27} Thus there is a great demand for searching effective HDAC inhibitors. Moreover, the recent findings of that HDAC inhibitors are also a relatively new class of potential antiprotozoal drugs^{10,11} and that histone acetylation maybe also involve in the infection of virus^{28,29} enhance the demand and urgency for the search of new HDAC in-

* E-mail: yuansg@mail.sioc.ac.cn

Received April 21, 2003; revised July 7, 2003; accepted August 21, 2003.

Project supported by the Minister of Science and Technology of China (No. 2002AA231011), the National Natural Science Foundation of China (No. 20073058), the Science and Technology Committee of Shanghai (No. 02DJ14013), Chinese Academy of Sciences-National Center of Scientific Research in France Cooperation Program (CNRS/CAS No. 12475) and Embassy of France in China.

hibitors.

While computer simulation techniques offer further means to probe inhibitor's action mechanisms³⁰ especially, 3D-QSAR approaches could build the quantity relationship between the structure and their activity of anti-tumor, and show the influence of substituents on anti-tumor activity. In this paper, the application of CoMFA^{31,32} and CoM-SIA^{33,34} to construct 3D-QSAR models is reported. These 3D-QSAR models can help us to well understand the interaction mechanism between inhibitors and histone deacetylase, and to make quantitative prediction of their inhibitory activities.

Data sets and computational methods

Data sets

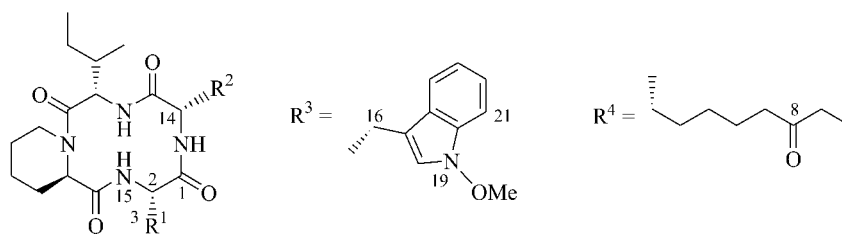
The structures and activities of inhibitors [$-\log(\text{IC}_{50})$] are extracted from the literature^{35,36} and listed in Table 1. HDAC (*Hela*) represents the activity of en-

zyme inhibition of HDAC, which extracts from human *Hela* cells. HDAC (*E. tenella*) represents the activity of enzyme inhibition of HDAC which extracts from *E. tenella* cell. Cell (*P. falciparum*) represents the activity of inhibition cells infected with *P. falciparum*. Cell (*E. tenella*) represents the activity of inhibition cells infected with *E. tenella*. Those compounds with activities which are remarked with "*" consist of their test sets and the others constitute the training sets, respectively.

Molecular modeling and alignment

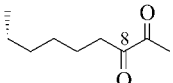
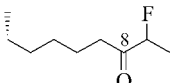
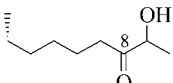
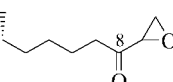
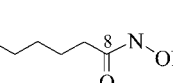
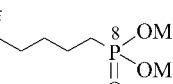
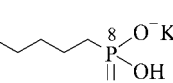
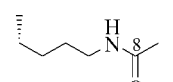
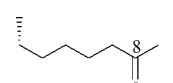
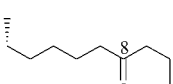
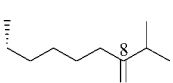
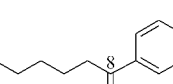
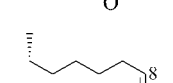
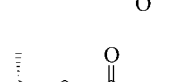
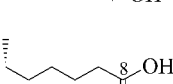
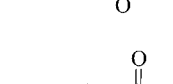
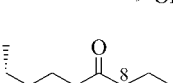
Three-dimensional structure building and all modeling were performed using the SYBYL³⁷ program package. Conformations of compounds in the training set and test set were generated using multisearch model in SYBYL. Energy minimization was performed using Tripos force field³⁸ with a distance-dependent dielectric and the Powell conjugate gradient algorithm with a convergence criterion of $0.42 \text{ kJ}/(\text{mol} \cdot \text{nm}^{-1})$. Partial atomic charges were calculated using the Gasteiger-Hückel method.

Table 1 Structures and activities of inhibitors

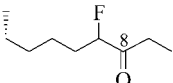
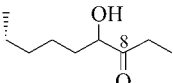
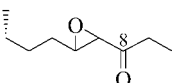
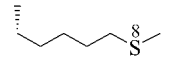
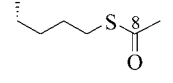
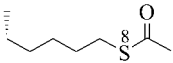
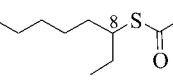
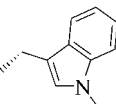
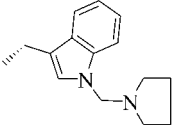
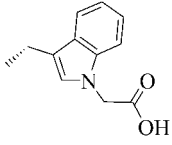
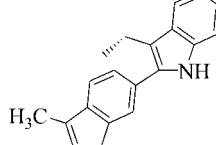
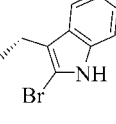
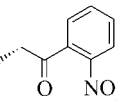
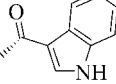


No.	R ¹	R ²	HDAC (<i>Hela</i>)	HDAC (<i>E. tenella</i>)	Cell (<i>P. falciparum</i>)	Cell (<i>E. tenella</i>)
1		R ³	4.00	4.00	2.24	2.03
2		R ³	0.84*	0.82	1.00	1.00
3		R ³	2.21	1.52	1.21	1.00
4		R ³	2.22	—	1.00	—
5		R ³	1.84	—	—	—
6		R ³	2.13	—	—	—
7		R ³	4.40	4.10*	1.68	—
8		R ³	3.15*	—	1.30	—

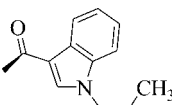
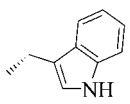
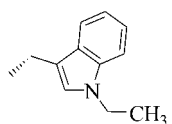
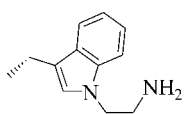
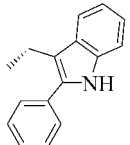
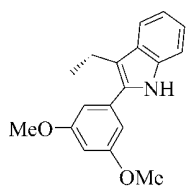
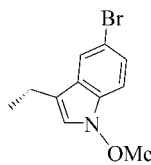
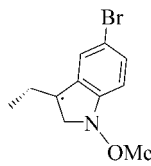
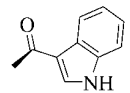
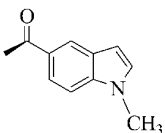
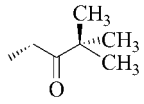
Continued

No.	R ¹	R ²	HDAC (<i>HeLa</i>)	HDAC (<i>E. tenella</i>)	Cell (<i>P. falciparum</i>)	Cell (<i>E. tenella</i>)
9		R ³	2.52	—	1.05	1.60*
10		R ³	4.00	—	2.55*	2.21*
11		R ³	4.52	3.40	2.66*	3.10
12		R ³	5.00	4.00	—	3.52
13		R ³	4.70	3.40*	—	—
14		R ³	1.35*	—	1.07	—
15		R ³	1.74*	—	—	—
16		R ³	0.70*	—	1.00	—
17		R ³	3.70	—	2.26*	1.9*
18		R ³	4.00	—	1.82	1.90*
19		R ³	1.63*	—	1.30	1.12*
20		R ³	1.15*	—	1.46	1.00*
21		R ³	3.22	—	1.30	1.00*
22		R ³	1.55	—	—	—
23		R ³	2.82	2.52	—	1.00
24		R ³	1.07*	—	—	—
25		R ³	2.74	—	1.15	1.50*

Continued

No.	R ¹	R ²	HDAC (<i>Helicoverpa</i>)	HDAC (<i>E. tenella</i>)	Cell (<i>P. falciparum</i>)	Cell (<i>E. tenella</i>)
26		R ³	1.96	—	1.15	1.50*
27		R ³	2.41	—	—	1.00*
28		R ³	3.30	—	1.46	1.72*
29		R ³	1.49	—	—	—
30		R ³	1.96	—	—	—
31		R ³	3.52*	—	1.05	1.00*
32		R ³	2.14	1.96	1.23	1.60
33	R ⁴		—	3.30	2.70	1.90
34	R ⁴		4.70	—	2.46	2.51*
35	R ⁴		—	2.00*	—	—
36	R ⁴		—	2.60	2.00	2.21
37	R ⁴		—	1.70	2.60	2.82
38	R ⁴		2.04*	2.48	1.30*	1.90
39	R ⁴		—	3.00	2.80	1.90

Continued

No.	R ¹	R ²	HDAC (<i>HeLa</i>)	HDAC (<i>E. tenella</i>)	Cell (<i>P. falciparum</i>)	Cell (<i>E. tenella</i>)
40	R ⁴		2.82	3.40	1.82	1.90
41	R ⁴		3.70	4.00	2.46	2.51
42	R ⁴		—	2.00	1.74	2.30
43	R ⁴		3.89	—	2.00	—
44	R ⁴		3.40	—	1.15*	—
45	R ⁴		4.70	4.70	1.60*	1.90
46	R ⁴		—	1.76	1.26*	1.30
47	R ⁴		—	0.60*	—	1.00*
48	R ⁴		—	2.64	2.15	1.90
49	R ⁴		—	2.52	2.82	2.30
50	R ⁴		1.35	2.35*	1.11*	—

Continued

No.	R ¹	R ²	HDAC (<i>Helicoverma</i>)	HDAC (<i>E. tenella</i>)	Cell (<i>P. falciparum</i>)	Cell (<i>E. tenella</i>)
51	R ⁴		1.59	2.70	1.11	1.30
52	R ⁴		1.38	2.30*	1.07*	—
53	R ⁴		2.47	3.52	—	1.60
54	R ⁴		2.07	3.40	2.40	1.60
55	R ⁴		1.48	2.96*	1.62*	—
56	R ⁴		—	—	2.40	—
57	R ⁴		—	—	2.26	—
58	R ⁴		2.17	3.70	1.65*	1.70
59	R ⁴		2.89	4.10	2.26	1.90

Cross-validated q^2 usually serves as the quantitative measurement of the prediction of CoMFA. Cho *et al.*³⁹ reported that q^2 value was sensitive to the orientation of aligned molecules on the computer terminal and might vary with the orientation by as much as 0.5 q^2 units. For decreasing the influence of structural orientation, we selected the ligand from the complex between histone deacetylase (HDAC) and suberoylanilide hydroxamic acid (SAHA), which was taken from Brookhaven Protein Databank (PDB code: 1C3R), as the template of alignment. The structure of the ligand is shown in Fig. 1. The atoms numbered (2, 3, 1, 15) for compound **1** aligned to atoms (1, 2, 9, 10) of the ligand. Other inhibitors were aligned to compound **1** with the routine SYBYL function of "database align".

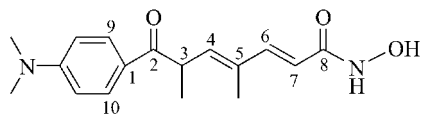


Fig. 1 Ligand extracted from PDB complex.

CoMFA models

After consistently aligning the molecules within a molecular lattice, which extends 0.4 nm unit beyond the aligned molecules in all directions, a probe atom using a sp^3 carbon atom with +1 net charge was employed. The steric and electrostatic interactions between the probe atom and the rest of the molecule were calculated. Electrostatic interactions are modeled using a Coulomb potential and Van der Waals interactions by Lennard-Jones potential. The regression analysis was carried out using the partial least-squares (PLS)⁴⁰ method. The final model was developed with the optimum number of components equal to that yielding the highest q_{cv}^2 . The total set of inhibitors initially was divided into two groups in the approximate ratio 3:1. The selection of the test set and training set compounds was done manually such that low, moderate and high anti-HDAC activity compounds occur in roughly equal proportions in both sets.⁴¹

Results and discussion

3D-QSAR model

The lowest energy conformer of each compound was selected. The aligned diagram of training set (HDAC of *HeLa*) is shown in Fig. 2. For four types of activities, four 3D-QSAR models were constructed. Their statistical parameters are listed in Table 2. The experimental activities (EA), predicted activities (PA) and their residues (RES) are listed in Table 3.



Fig. 2 Alignment of anti-HDAC for training set.

Table 2 PLS Statistics of CoMFA Models

PLS Statistics	Model 1 (HDAC- <i>HeLa</i>)	Model 2 (HDAC- <i>E. Tenella</i>)	Model 3 (<i>Cell-P. falciparum</i>)	Model 4 (<i>Cell-E. etenella</i>)
	36 compounds	24 compounds	34 compounds	24 compounds
q^2	0.624	0.652	0.713	0.566
R^2	0.939	0.963	0.947	0.959
S	0.294	0.210	0.147	0.149
F	120.238	93.866	79.848	65.680
PLS component	4	5	6	6
Steric	0.473	0.521	0.433	0.431
Electrostatic	0.527	0.479	0.567	0.569

Table 3 Experimental and predicted activities for four models

No.	Model 1			Model 2			Model 3			Model 4		
	EA	PA	RES	EA	PA	RES	EA	PA	RES	EA	PA	RES
1	4.00	3.67	0.33	4.00	3.79	0.21	2.24	1.78	0.46	2.03	1.61	0.42
2	0.84*	0.91	-0.07	0.82	1.21	-0.39	1.00*	1.19	-0.19	1.00	1.25	-0.25
3	2.21	2.62	-0.41	1.52	1.49	0.03	1.21	1.38	-0.17	1.00	0.99	0.01
4	2.22	2.17	0.05	—	—	—	1.00	1.12	-0.12	—	—	—
5	1.84	2.01	-0.17	—	—	—	—	—	—	—	—	—
6	2.13	1.68	0.45	—	—	—	—	—	—	—	—	—
7	4.40	3.87	0.53	4.10*	4.08	0.02	1.68*	1.14	0.54	—	—	—
8	3.15*	3.72	-0.57	—	—	—	1.30*	1.04	0.26	—	—	—
9	2.52	2.79	-0.27	—	—	—	1.05	1.05	0.00	1.60*	1.60	0.00
10	4.00	4.11	-0.11	—	—	—	2.55	2.57	-0.02	2.21*	2.18	0.03
11	4.52	4.45	0.071	3.40	3.39	0.01	2.66*	2.35	0.31	3.10	3.19	-0.09
12	5.00	4.53	0.47	4.00	3.69	0.31	—	—	—	3.52	3.52	0.00
13	4.70	4.33	0.37	3.40*	3.45	-0.05	—	—	—	—	—	—
14	1.35*	1.26	0.09	—	—	—	1.07*	1.59	-0.52	—	—	—
15	1.74*	1.34	0.40	—	—	—	—	—	—	—	—	—
16	0.70*	0.75	-0.05	—	—	—	1.00	1.09	-0.09	—	—	—
17	3.70	3.46	0.24	—	—	—	2.26	2.20	0.06	1.90*	1.89	0.01
18	4.00	3.81	0.19	—	—	—	1.82	1.75	0.07	1.90*	1.82	0.08
19	1.63*	1.63	0.00	—	—	—	1.30	1.17	0.13	1.12*	1.27	-0.15
20	1.15*	1.01	0.14	—	—	-1.46	1.42	0.04	1.00*	1.18	-0.18	—
21	3.22	3.68	-0.46	—	—	—	1.30	1.33	-0.03	1.00*	1.13	-0.13

Continued

No.	Model 1			Model 2			Model 3			Model 4		
	EA	PA	RES	EA	PA	RES	EA	PA	RES	EA	PA	RES
22	1.55	1.65	-0.10	—	—	—	—	—	—	—	—	—
23	2.82	3.38	-0.56	2.52	2.85	-0.33	—	—	—	1.00	0.99	0.01
24	1.07*	0.97	0.10	—	—	—	—	—	—	—	—	—
25	2.74	2.68	0.06	—	—	—	1.15	1.10	0.05	1.50*	1.51	-0.01
26	1.96	1.94	0.02	—	—	—	1.15	1.24	-0.09	1.50*	1.48	0.02
27	2.41*	2.66	-0.25	—	—	—	—	—	—	1.00*	1.00	0.00
28	3.30	3.41	-0.11	—	—	—	1.46	1.49	-0.03	1.72*	1.72	0.00
29	1.49	1.83	-0.34	—	—	—	—	—	—	—	—	—
30	1.96	1.87	0.09	—	—	—	—	—	—	—	—	—
31	3.52*	3.04	0.48	—	—	—	1.05	1.04	0.01	1.00*	1.17	-0.17
32	2.14	2.17	-0.03	1.96	1.77	0.19	1.23	1.22	0.01	1.60	1.73	-0.13
33	—	—	—	3.30	3.63	-0.33	2.70	2.42	0.28	1.90	2.11	-0.21
34	4.70	4.86	-0.16	—	—	—	2.46*	1.98	0.48	2.51*	2.25	0.26
35	—	—	—	2.00*	1.77	0.23	—	—	—	—	—	—
36	—	—	—	2.60	2.67	-0.07	2.00	2.12	-0.12	2.21	2.25	-0.04
37	—	—	—	1.70	1.75	-0.05	2.60	2.55	0.05	2.82	2.65	0.17
38	2.04*	1.96	0.08	2.48	2.25	0.23	1.30*	1.70	-0.40	1.90	1.94	-0.04
39	—	—	—	3.00	2.89	0.11	2.80	2.92	-0.12	1.90	2.02	-0.12
40	2.82	2.64	0.18	3.40	3.49	-0.09	1.82	1.90	-0.08	1.90	1.88	0.02
41	3.70	3.31	0.39	4.00	3.99	0.01	2.46	2.55	-0.09	2.51	2.44	0.07
42	—	—	—	2.00	1.68	0.32	1.74	2.13	-0.39	2.30	2.25	0.05
43	3.89	4.36	-0.47	—	—	—	2.00	2.01	-0.01	—	—	—
44	3.40	3.63	-0.23	—	—	—	1.15	1.12	0.03	—	—	—
45	4.70	4.89	-0.19	4.70	4.64	0.06	1.60	1.53	0.07	1.90	1.83	0.07
46	—	—	—	1.76	1.85	-0.09	1.26	1.28	-0.02	1.30	1.29	0.01
47	—	—	—	0.60*	1.53	-0.93	—	—	—	1.00*	1.16	-0.16
48	—	—	—	2.64	2.75	-0.11	2.15	2.12	0.03	1.90	1.93	-0.03
49	—	—	—	2.52	2.47	0.05	2.82*	1.89	0.92	2.30	2.29	0.01
50	1.35	1.19	0.16	2.35*	3.08	-0.73	1.11	1.16	-0.05	—	—	—
51	1.59	1.58	0.01	2.70	2.58	0.12	1.11	1.07	0.04	1.30	1.26	0.04
52	1.38	1.19	0.19	2.30*	2.83	-0.53	1.07*	1.37	-0.30	—	—	—
53	2.47	2.53	-0.06	3.52	3.55	-0.03	—	—	—	1.60	1.64	-0.04
54	2.07	1.99	0.08	3.40	3.37	0.03	2.40*	2.43	-0.03	1.60	1.55	0.05
55	1.48	1.73	-0.25	2.96*	3.02	-0.06	1.62 ⁷ *	1.64	-0.02	—	—	—
56	—	—	—	—	—	—	2.40	2.34	0.06	—	—	—
57	—	—	—	—	—	—	2.26	2.19	0.07	—	—	—
58	2.17	2.27	-0.10	3.70	3.71	-0.01	1.65	1.66	-0.01	1.70	1.71	-0.01
59	2.89	2.77	0.12	4.10	4.28	-0.18	2.26	2.28	-0.02	1.90	1.89	0.01

Evaluation of four 3D-QSAR models

For four types of different activities, four 3D-QSAR models were constructed. Model 1 was constructed with a training set of 36 molecules. The cross-validated q^2 value of CoMFA model of training set is 0.624 with four principal components. The noncross-validated r^2 value is 0.939

with standard errors (SE) 0.294. The corresponding correlation coefficient r^2 value between EA and PA of test set is 0.916 with standard errors (SE) 0.287. The correlation between EA and PA is shown in Fig. 3 (a). It is shown that model 1 is excellent. Model 2 was constructed with a training set of 24 molecules. The cross-validated q^2 value of training set is 0.652 with five principal compo-

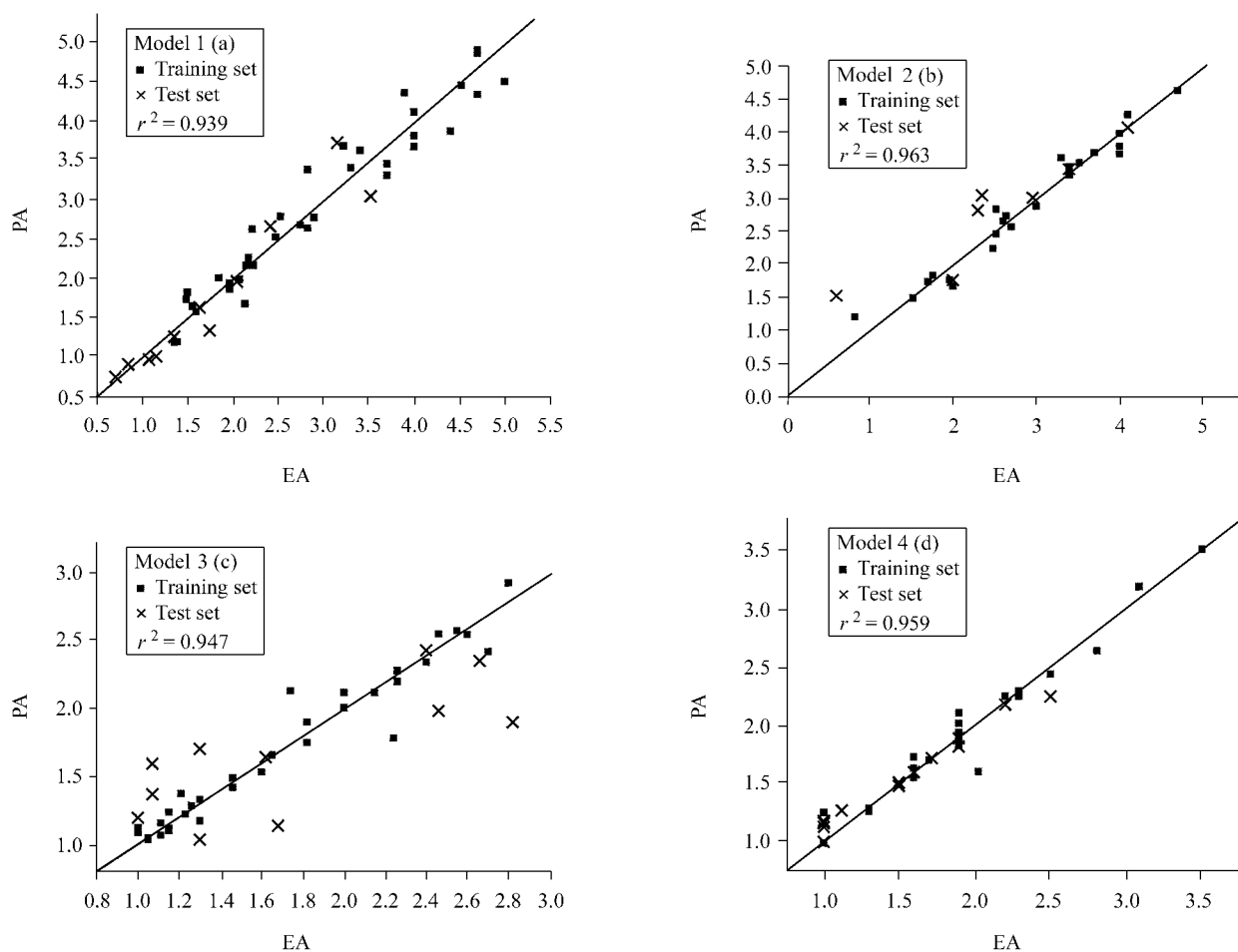


Fig. 3 Correlations between EA and PA for models 1–4.

nents. The noncross-validated r^2 value is 0.963, with standard errors (SE) equal to 0.210. The corresponding correlation coefficient r^2 between EA and PA of test set is 0.867, with standard errors (SE) 0.359. The correlation between EA and PA is shown in Fig. 3(b). It is shown that model 2 is a useable model. Model 3 was built with a training set of 34 compounds. The cross-validated q^2 value of training set is 0.713, with six principal components. The noncross-validated r^2 value is 0.947 with standard errors (SE) 0.147. The corresponding correlation coefficient r^2 value between EA and PA of test set is 0.592, with standard errors (SE) 0.408. The correlation between EA and PA is shown in Fig. 3(c). It is shown that CoMFA model has certain predicted ability. Model 4 includes 24 compounds in training set. The cross-validated q^2 value of training set is 0.566 with six principal components. The noncross-validated r^2 value is 0.959 with standard errors (SE) 0.149. The corresponding correlation coefficient r^2 value between EA and PA of test set is 0.981 with standard errors (SE) 0.059. The correlation between EA and PA is shown in Fig. 3(d). It is shown that model 4 is a robust and good model.

Analysis of CoMFA models

The PLS statistics of CoMFA are summarized in Table

2. The steric and electrostatic field contributions of model 1 are 0.473 and 0.527, model 2 are 0.521 and 0.479, model 3 are 0.433 and 0.567 and model 4 are 0.431 and 0.569, respectively. Exception model 2, electrostatic field is the main influence factor to activity for other models.

Because HDAC analogues have different substituents at positions 2 and 14 (Table 1), we selected these positions to investigate the relationship between substituents and anti-HDAC activity in detail. There is not steric and electrostatic field around the skeleton ring of these models. Fig. 4A shows the contour plots of CoMFA model with structure 11 of model 1. The meanings of the different colour areas are shown in the legend of this Figure. Yellow-coloured and red-coloured regions near position 8 show that small volume and negative substituents could increase inhibitor activity. The anti-HDAC activities for the following three compounds are in the order of **20** < **19** < **18**. The bulk volumes of their corresponding substituents have the same order CHMe_2 (0.3118 nm^3) < $\text{CH}_2\text{CH}_2\text{Me}$ (0.13163 nm^3) < phenyl (0.3254 nm^3). The anti-HDAC activities for the following five compounds are in the order of **1** < **7** < **11** < **14** < **13**, agreeing with the order of the ability of donor charge (CH_2CH_3 < OCH_3 < OH < NHOH < epoxide). Because some red-coloured regions are just near position 8, the anti-HDAC activity of compounds (carbonyl is at position 7 or 9) will decrease. So the activities of these com-

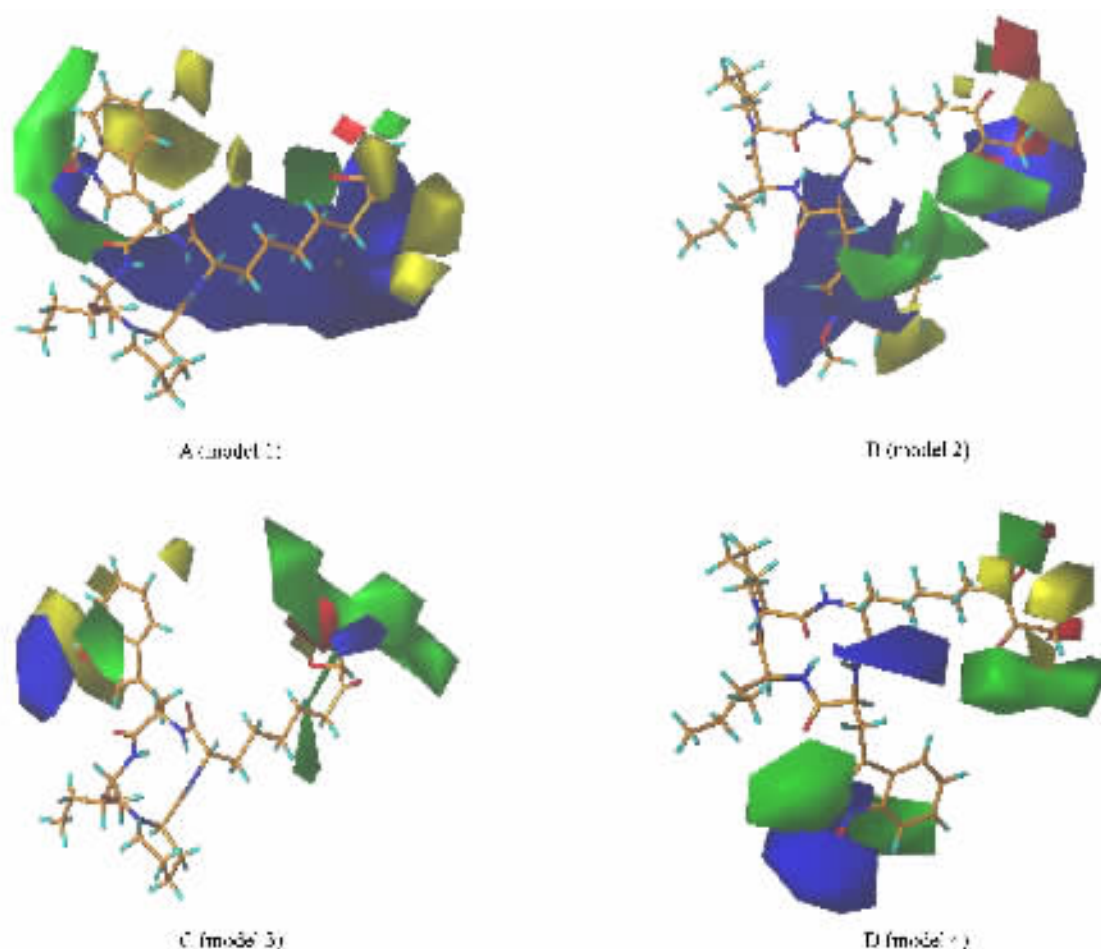


Fig. 4 CoMFA contour plots. Green contours indicate regions where bulky groups increase activity, whereas yellow contours indicate regions where bulky groups decrease activity. Blue contours indicate regions where positive groups increase activity, whereas red contours indicate regions where negative charge increases activity.

pounds have the order of **7** (C₈-ester) > **8** (C₉-ester) > **6** (C₇-ester), **23** (C₈-acid) > **22** (C₇-acid) > **24** (C₉-acid) and **31** (C₈-thiol) > **32** (C₉-thiol) > **30** (C₇-thiol). Finnin *et al.*⁴² reported that carbonyl of position 8 of SAHA compound has chelating interaction with Zn, and carbonyl of position 8 for apicidin occupies the same place by aligned apicidin with SAHA. It explains the compound of carbonyl at position 8 has higher activity. As the substituent of position 14, yellow-coloured near benzene ring and green-coloured near pyrrole or pyridine ring show that small volume substituent of benzene is favourable for the anti-HDAC activity and bulky substituent of pyrrole may increase activity. The anti-HDAC activities for the following three compounds are in the order of **41** < **1** < **34** and **39** < **40**. The bulk volumes of their corresponding substituents at N₁₉ have the same order of H (0.0437 nm³) < OCH₃ (0.1722 nm³) < CH₂⁻ pyrrolidine (0.3194 nm³). This is in agreement with the conclusion from CoMFA model.

Fig. 4C illustrates the CoMFA contour plot of model 3 with structure **11**. Red-coloured region near position 8 at the direction of vertical chain shows that negative charge could increase inhibitor activity. The substituents of compounds **1**, **2** and **3** are carbonyl (C=O), ethane (C=C) and hydrogen (C-H), their abilities of donor charge have

the order of **1** > **3** > **2**. This is consistent with the order of activity. At the direction of chain, green-coloured region near position 8 shows that bulk groups are favourable to the activity. The volumes of substituents for **11** [-CH(OH)CH₃], **10** [-CH(F)CH₃], **1** (-CH₂CH₃) and **7** (-OCH₃) are 0.2349 nm³, 0.2131 nm³, 0.2092 nm³ and 0.1725 nm³, respectively. Their anti-HDAC active values are ordered as **11** > **10** > **1** > **7**.

For model 2 and model 4, the same training sets were selected to construct their CoMFA models, so we could better find their relationships. The CoMFA contour plot of model 2 is shown in Fig. 4B with structure **11**. Red-coloured, green-coloured and blue-coloured regions near position 8 show that negative and suitable bulk groups are favourable to the activity. This could explain why the compounds with carbonyl substituent have higher activity than other substituents at position 8. At the chain end, there are green-coloured, yellow-coloured and blue-coloured regions. It shows that suitable bulk and positive groups are favourable to the anti-HDAC activity. This explains that the activities of compounds **7** (-OCH₃), **12** (epoxide) and **11** [-CH(OH)CH₃] have the order of **7** > **12** > **11**. Blue-coloured, yellow-coloured and green-coloured regions are near position 14, these show that positive charge

groups could increase activity. The substituents of compound **1**, **33**, **40**, **42**, **45** and **49** are the donor group of electron. Their activities are better than those of compounds **35** and **46**. The suitable bulk substituents could increase anti-HDAC activity. The activities of compounds **45** and **36** testify this.

Fig. 4D illustrates the CoMFA contour plot of model 4. There are green-coloured and red-coloured regions just near position 8. It is shown that bulk and negative groups are favourable to the activity and this is the same as model 2. Unlike model 2, green-coloured, yellow-coloured and red-coloured regions are near the substituents of the end-chain, these compounds with suitable bulk and negative substituents easily penetrate the cell membrane. They have higher anti-cell activity. It is consistent with the experimental result. It causes the difference between the activity of enzyme and that of cell. There are green-coloured regions near the indole ring of position 14. It is shown that bulk substituents of this position could increase activity. Blue-coloured regions near the indole ring of position 14 show that positive groups of position N₁₉ are favourable to the activity. The activities of compounds **42** (- C₂H₅), **1** (- OCH₃), **43** (- CH₂CH₂NH₂) have the order of **42** > **1** > **43**.

For model 2 and model 4, suitable bulk groups are favourable to their inhibitor activities at position 8. Positive and suitable bulk groups could increase the activity of enzyme, while negative charge groups at position 2 are favourable to the activity of anti-cell. Because the substituents of suitable bulk and negative charge could easily penetrate the cell membrane, these compounds are more favourable to cell activity. Consideration of four models, these compounds with suitable bulk and positive substituents are favourable to anti-HDAC activity at position 14.

Conclusion

For four types of different activities, four 3D-QSAR models were constructed. The cross-validated q^2 value of CoMFA Model 1 is 0.624 and the noncross-validated r^2 value is 0.939. The cross-validated q^2 value of CoMFA Model 2 is 0.652 and noncross-validated r^2 value is 0.963. The cross-validated q^2 value of CoMFA Model 3 is 0.713 with non cross-validated r^2 value 0.947. The cross-validated q^2 value of Model 4 is 0.566 with noncross-validated r^2 value 0.959. Their predicted abilities are validated by different test sets. The corresponding correlation coefficient r^2 value between EA and PA of test set for model 1 is 0.916, with standard errors (SE) 0.287. The corresponding correlation coefficient r^2 value between EA and PA of test set for model 2 is 0.867 with standard errors (SE) 0.359. The corresponding correlation coefficient r^2 value between EA and PA of test set for model 3 is 0.592 with standard errors (SE) 0.408. The corresponding correlation coefficient r^2 value between EA and PA of test set for model 4 is 0.959 with standard errors (SE) 0.149.

Then the relationship between substituents and activities was analyzed using these models and main influence elements in different positions (positions 8 and 14) were found. The polar donor electron group of position 8 could increase the activity of inhibition of HDAC, because it could form chelation with the catalytic Zn. Suitable bulk and positive groups at position 14 are favorable to anti-HDAC activity. These models could interpret the relationship between inhibition activity and apicidin structure well affording us important information for structure-based drug design.

References

- 1 Kouzarides, T. *Curr. Opin. Genet. Dev.* **1999**, *9*, 40.
- 2 Spencer, V. A.; Davie, J. R. *Gene* **1999**, *240*, 1.
- 3 Klochendler, Y. A.; Moshe, Y. *Biochim. Biophys. Acta* **2001**, *1551*, M1.
- 4 Fry, C. J.; Peterson, C. L. *Science* **2002**, *295*, 1847.
- 5 Archer, S. Y.; Hodin, R. A. *Curr. Opin. Genet. Dev.* **1999**, *9*, 171.
- 6 Grunstein, M. *Nature* **1997**, *389*, 349.
- 7 Wolffe, A. P.; Pruss, D. *Cell* **1996**, *84*, 817.
- 8 Won, J.; Yim, J.; Kim, T. K. *J. Biol. Chem.* **2002**, *277*, 38230.
- 9 Giles, R. H.; Peters, D. J. M.; Breuning, M. H. *Trends Genet.* **1998**, *14*, 178.
- 10 Singh, S. B.; Zink, D. L.; Liesch, J. M.; Dombrowski, A. W.; Darkin, R. S. J.; Schmatz, D. M.; Goetz, M. A. *Org. Lett.* **2001**, *3*, 2815.
- 11 Joshi, M. B.; Lin, D. T.; Chiang, P. H.; Goldman, N. D.; Fujioka, H.; Aikawa, M.; Syin, C. *Mol. Biochem. Parasitol.* **1999**, *99*, 11.
- 12 Timmermann, S.; Lehrmann, H.; Polesskaya, A.; Harel, B. A. *Cell Mol. Life Sci.* **2001**, *58*, 728.
- 13 Zhang, Y.; Dufau, M. L. *J. Biol. Chem.* **2002**, *277*, 33431.
- 14 Marks, P.; Rifkind, R. A.; Richon, V. M.; Breslow, R.; Miller, T.; Kelly, W. K. *Nat. Rev. Cancer* **2001**, *1*, 194.
- 15 Fuks, F.; Burgers, W. A.; Godin, N.; Kasai, M.; Kouzarides, T. *EMBO J.* **2001**, *20*, 2536.
- 16 Cameron, E. E.; Bachman, K. E.; Myohanen, S.; Herman, J. G.; Baylin, S. B. *Nat. Genet.* **1999**, *21*, 103.
- 17 Johnstone, R. W. *Nat. Rev. Drug Discov.* **2002**, *1*, 287.
- 18 Weidle, U. H.; Grossmann, A. *Anticancer Res.* **2000**, *20*, 1471.
- 19 Lavelle, D.; Chen, Y. H.; Hankewych, M.; DeSimone, J. *Am. J. Hematol.* **2001**, *68*, 170.
- 20 Sambucetti, L. C.; Fischer, D. D.; Zabludoff, S.; Kwon, P. O.; Chamberlin, H.; Trogani, N.; Xu, H.; Cohen, D. *J. Biol. Chem.* **1999**, *274*, 34940.
- 21 Yu, X.; Guo, Z. S.; Marcu, M. G.; Neckers, L.; Nguyen, D. M.; Chen, G. A.; Schrupp, D. S. *J. Nat. Cancer Inst.* **2002**, *94*, 504.
- 22 Kwon, H. J.; Kim, M. S.; Kim, M. J.; Nakajima, H.; Kim, K. W. *Int. J. Cancer* **2002**, *97*, 290.
- 23 Suenaga, M.; Soda, H.; Oka, M.; Yamaguchi, A.; Nakatomi, K.; Shiozawa, K.; Kawabata, S.; Kasai, T.;

- Yamada, Y. ; Kamihira, S. ; Tei, C. ; Kohno, S. *Int. J. Cancer* **2002**, *97*, 621.
- 24 Jung, M. In *Transcription Factors and Other Nuclear Proteins as Targets for Cancer Chemotherapy*, Eds. : Thangue, N. B. ; Bandara, L. R. , Humana Press, Totowa, **2002**, pp. 123—144.
- 25 Coffey, D. C. ; Kutko, M. C. ; Glick, R. D. ; Butler, L. M. ; Heller, G. ; Rifkind, R. A. ; Marks, P. A. ; Richon, V. M. ; La Quaglia, M. P. *Cancer Res.* **2001**, *61*, 3591.
- 26 Warrell, R. P. J. ; He, L. Z. ; Richon, V. ; Calleja, E. ; Pandolfi, P. P. *J. Nat. Cancer Inst.* **1998**, *90*, 1621.
- 27 Piekarz, R. L. ; Robey, R. ; Sandor, V. ; Bakke, S. ; Wilson, W. H. ; Dahmouh, L. ; Kingma, D. M. ; Turner, M. L. ; Altemus, R. ; Bates, S. E. *Blood* **2001**, *98*, 2865.
- 28 Li, M. ; Damania, B. ; Alvarez, X. ; Ogryzko, V. ; Ozato, K. ; Jung, J. U. *Mol. Cell. Biol.* **2000**, *20*, 218254.
- 29 Col, E. ; Gilquin, B. ; Caron, C. ; Khochbin, S. *J. Biol. Chem.* **2002**, *277*, 37955.
- 30 Kireev, D. B. ; Chretien, J. R. ; Grierson, D. S. ; Monneret, C. *J. Med. Chem.* **1997**, *40*, 4257.
- 31 Cramer, III R. D. ; Patterson, D. E. ; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
- 32 Clark, M. ; Cramer, III R. D. ; Jones, D. M. ; Patterson, D. E. ; Simeroth, P. E. *Tetrahedron Comput. Tech.* **1990**, *3*, 47.
- 33 Klebe, G. ; Abraham, U. ; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 4130.
- 34 Böhm, M. ; Stürzebecher, J. ; Klebe, G. *J. Med. Chem.* **1999**, *42*, 458.
- 35 Colletti, S. L. ; Myers, R. W. ; Darkin, R. S. J. ; Gurnett, A. M. ; Dulski, P. M. ; Galuska, S. ; Allocco, J. J. ; Ayer, M. B. ; Li, C. ; Lim, J. ; Crumley, T. M. ; Cannova, C. ; Schmatz, D. M. ; Wyratt, M. J. ; Fisher, M. H. ; Meinke, P. T. *Bioorg. Med. Chem. Lett.* **2001**, *22*, 107.
- 36 Colletti, S. L. ; Myers, R. W. ; Darkin, R. S. J. ; Gurnett, A. M. ; Dulski, P. M. ; Galuska, S. ; Allocco, J. J. ; Ayer, M. B. ; Li, C. ; Lim, J. ; Crumley, T. M. ; Cannova, C. ; Schmatz, D. M. ; Wyratt, M. J. ; Fisher, M. H. ; Meinke, P. T. *Bioorg. Med. Chem. Lett.* **2001**, *22*, 113.
- 37 SYBYL Program Package, Version 6.8, St. Louis (MO), Tripos Associates Inc, **2001**.
- 38 Clark, M. ; Cramer, III R. D. ; van Opdenbosch, N. *J. Comput. Chem.* **1989**, *10*, 982.
- 39 Cho, S. J. ; Tropsha, A. *J. Med. Chem.* **1995**, *38*, 1060.
- 40 Clark, M. ; Cramer, III R. D. *Quant. Struct.-Act. Relat.* **1993**, *12*, 137.
- 41 Desiraju, G. R. ; Gopalakrishnan, B. ; Jetti, R. K. R. ; Nagaraju, A. ; Raveendra, D. ; Sarma, J. A. R. P. ; Sobhia, M. E. ; Thilagavathi, R. *J. Med. Chem.* **2002**, *45*, 4847.
- 42 Finin, M. S. ; Donigian, J. R. ; Cohen, A. ; Richon, V. M. ; Rifkind, R. A. ; Marks, P. A. ; Breslow, R. ; Paletich, N. P. *Nature* **1999**, *401*, 188.

(E0304218 SONG, J. P. ; FAN, Y. Y.)