Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry 15 (2007) 4470–4481
Bioorganic & Medicinal Chemistry

3D-QSAR studies on malonyl coenzyme A decarboxylase inhibitors

Maulik R. Patel and Tanaji T. Talele*

Department of Pharmaceutical Sciences, College of Pharmacy & Allied Health Professions, St. John's University, Jamaica, NY 11439, USA

Received 8 December 2006; revised 9 April 2007; accepted 15 April 2007
Available online 24 April 2007

Abstract—Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) were performed on a series of Malonyl Co-A decarboxylase (MCD) inhibitors (Cheng et al. *J. Med. Chem.* **2006**, *49*, 1517–1525 and Cheng et al. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 695–700). These inhibitors have shown protective action on the ischemic heart by inhibiting fatty acid oxidation. The CoMFA model produced statistically significant results, with the cross-validated and conventional correlation coefficients being 0.544 and 0.986, respectively. The best results were obtained by combining steric, electrostatic, hydrophobic, and H-bond acceptor fields in CoMSIA, in which case the respective cross-validated and conventional correlation coefficients were 0.551 and 0.918. The predictive ability of CoMFA and CoMSIA, determined using a test set of 24 compounds, gave predictive correlation coefficients of 0.718 and 0.725, respectively. The information obtained from CoMFA and CoMSIA 3D contour maps may be of utility in the design of more potent MCD inhibitors.

1. Introduction

Malonyl-CoA decarboxylase[1] is an enzyme that, by catalyzing the conversion of malonyl-CoA to acetyl-CoA, regulates the intracellular levels of malonyl-CoA. In turn, malonyl-CoA plays a pivotal role in fuel availability in the cardiac muscle[2] by acting as a potent endogenous inhibitor of carnitine palmitoyltransferase-I (CPT-I), a fatty acid transporter localized in the outer mitochondrial membrane. CPT-I is the rate-limiting enzyme in fatty acid oxidation and catalyzes the formation of fatty acylcarnitine from a long fatty acyl-CoA for subsequent transport from the cytosol into the mitochondrial matrix by acylcarnitine translocase. Inside the mitochondrion, the long-chain fatty acids are reconverted to their CoA thioester forms by a complementary enzyme, CPT-II, and enter the β-oxidation pathway for conversion to acetyl-CoA. In the liver, high levels of acetyl-CoA lead to elevated malonyl-CoA levels which will then inhibit CPT-I and, thereby, prevent fatty acid oxidation and favor fatty acid synthesis. Conversely, low malonyl-CoA levels favor fatty acid oxidation by allowing the transport of long-chain fatty acids into the mitochondrion. Therefore, malonyl-CoA is a central metabolite in fatty acid metabolism and plays a key role in balancing fatty acid synthesis and fatty acid oxidation.[3] Moreover, in the heart fatty acid oxidation is a major source of energy, and fatty acid uptake and metabolism down-regulate glucose metabolism. The regulation of intermediary metabolism by circulating fatty acids and glucose comprises the glucose-fatty acid (or Randle) cycle.[4] Under ischemic conditions, the oxygen supply is limited in cardiac tissues and the oxidation of both fatty acids and glucose and the formation of ATP by oxidative phosphorylation are all reduced. Glycolysis increases in an attempt to normalize ATP levels but this compensatory mechanism unfortunately ends in lactic acidosis. In addition, energy is spent to maintain ion homeostasis and myocyte death occurs as a result of abnormally low ATP levels and disrupted osmolarity. This situation is complicated further by a decrease in malonyl-CoA levels due to the activation of a cAMP-dependent protein kinase (AMPK)[5], which phosphorylates and thus inactivates acetyl-CoA carboxylase (ACC). Under these circumstances, both CPT-I activity and fatty acid oxidation are favored over glucose oxidation. It is quite evident that the inhibition of CPT-I activity by increasing malonyl-CoA levels with MCD inhibitors would result in a novel and perhaps a much safer method, compared to other known small-molecule CPT-I inhibitors for the prophylaxis and treatment of ischemic heart diseases.[6]

Keywords: 3D-QSAR; Malonyl coenzyme A decarboxylase.
* Corresponding author. Tel.: +1 718 990 5405; fax: +1 718 990 1877; e-mail: talelet@stjohns.edu

0968-0896/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2007.04.025

Due to a lack of structural information on the MCD enzyme, a ligand-based approach (3D-QSAR) has been used to understand which structural features of MCD inhibitors are relevant to their binding to the active site of MCD. To establish the SAR of these inhibitors, different ligand-based 3D-QSAR methods were employed. Comparative molecular field analysis (CoMFA)[7,8] and comparative molecular similarity indices analysis (CoMSIA)[9] were applied to correlate molecular property fields of aligned inhibitors of MCD using the PLS (partial least squares) method. The resulting contour maps from 3D-QSAR models will help to better understand the steric, electrostatic, hydrophobic, and hydrogen-bond acceptor requirements for MCD inhibition and will serve as a tool for designing more potent MCD inhibitors.

2. Results and discussion

2.1. CoMFA and CoMSIA statistical results

The training set was initially examined for outliers. Usually, if the residual between observed pIC_{50} and predicted pIC_{50} values of an inhibitor is approximately 1 logarithm unit, the inhibitor is considered an outlier. According to this criterion, inhibitors **32** and **43** were found to be outliers. For compounds **32**–**36**, their IC_{50} values somewhat reflected their differing R_2 substituent (i.e., –Me in **32** $IC_{50} = 0.178$ μM, –Me in **33** $IC_{50} = 0.394$ μM, –Et in **34** $IC_{50} = 0.041$ μM, *n*-Pr in **35** $IC_{50} = 0.028$ μM, and *n*-Bu in **36** $IC_{50} = 0.02$ μM). It is quite clear from the training set that increasing the length of alkyl group at R_2 position increases the MCD inhibitory activity until it reaches a maximum when the number of carbons is 6 (e.g., compounds **32**–**36** and **44**–**48**). Since compounds **32** and **43** lack R_2 substituents larger than –Me, they were not predicted well by the CoMFA and CoMSIA models. Among compounds **44**–**53**, *n*-pentyl and methyl pentanoate moieties are optimal at the R_2 position for the potent MCD inhibitory activity as exemplified by compounds **46** ($IC_{50} = 0.019$ μM) and **53** ($IC_{50} = 0.042$ μM). Thus, both CoMFA and CoMSIA models based on the 54 compounds led to a lower cross-validated correlation coefficient q^2 of 0.419 and 0.352, respectively. However, after dropping compounds **32** and **43** from the training set, the 3D-QSAR models were re-performed by using the remaining 52 compounds. In this manner a significantly higher cross-validated q^2 of 0.544 for CoMFA and 0.551 for CoMSIA was possible (Table 1).

The statistical results for CoMFA and CoMSIA modeling are summarized in Table 1. CoMFA using the Sybyl atom fit alignment method gave a cross-validated r^2 of 0.544 with an optimal number of principal components (5) and a conventional r^2 of 0.986 for the non-cross-validated final model. The detailed experimental and predicted pIC_{50} values of the training set using best CoMFA model are presented in Table 2 and Figure 1a.

It has been established that the five different descriptor fields are not totally independent of each other and that such dependencies on individual fields usually decrease the statistical significance of the CoMSIA models.[10,11] An evaluation, which one among the five CoMSIA fields is actually needed for the generation of predictive model, was performed by computing all possible combinations of fields (Table 3). The steric and hydrophobic fields yielded the best individual field models with q^2 of 0.405 and 0.402, respectively. The acceptor field was of the lowest internal predictivity. In the combined models, the highest q^2 was obtained by combining steric, electrostatic, hydrophobic, and H-bond acceptor fields ($q^2 = 0.551$ and conventional $r^2 = 0.918$). The detailed experimental and predicted pIC_{50} values based on the selected CoMSIA model for the training set are shown in Table 2 and Figure 1b.

To further assess the robustness and statistical confidence of the derived models, bootstrapping[10] analysis for 30 runs was performed (Table 1). Bootstrapping involves the generation of many new datasets from the original datasets after randomly choosing samples from the original dataset. The bootstrapped r^2 of 0.987 for CoMFA and 0.966 for CoMSIA are suggesting that a good internal consistency exists within the underlying dataset.

Next a cross-validation analysis was applied to the set of compounds in the training set to investigate the stability of the CoMFA and CoMSIA models. The training set model was cross-validated using two (Leave-Half-Out) and five (Leave 20% out) cross-validation groups 50 times each. The average and standard deviation values of q^2 are shown in Table 1. When two cross-validation groups were used, the average q^2 values for CoMFA and CoMSIA were, respectively, 0.155 (standard deviation = 0.459) and 0.326 (standard deviation = 0.420). The use of two cross-validation groups leaves 26 of the 52 training set molecules in the model construction group and predicts the activities of the remaining 26 compounds. By using five cross-validation groups a more consistent cross-validation of training set composition for each run was achieved. Thus for CoMFA the q^2 and SD values were 0.432 and 0.394, those for the CoMSIA were 0.481 and 0.385, respectively.

2.2. Predictivity of 3D QSAR models

The predictive abilities of CoMFA and CoMSIA models were determined from a set of 24 test compounds not included in the model generation. Both CoMFA and CoMSIA models performed well in the prediction of the activities of the test inhibitors. The predicted r^2 from CoMFA and CoMSIA models were found to be 0.718 and 0.725, respectively (Table 1). The detailed experimental and predicted pIC_{50} values based on selected CoMFA and CoMSIA models for the test set compounds are listed in Table 4 and graphically shown in Figures 2a and b. The predicted pIC_{50} residuals of these CoMFA and CoMSIA models were, in all instances, less than 0.5 log unit. Statistically significant predictions support the validity of the derived models in predicting the inhibitory activity of newer compounds.

Table 1. Summary of CoMFA and CoMSIA results

PLS statistics	54-Compound model		52-Compound model	
	CoMFA	CoMSIA	CoMFA	CoMSIA
r^2[a]	0.975	0.793	0.986	0.918
SEE[b]	0.077	0.217	0.057	0.139
F_{test}[c]	369.74	63.67	662.93	131.09
q^2[d]	0.419	0.352	0.544	0.551
SEP[e]	0.390	0.418	0.363	0.367
r^2_{pred}[f]	—	—	0.718	0.725
PLS components[g]	5	3	5	4
Contribution				
Steric	0.54	0.17	0.53	0.16
Electrostatic	0.46	0.18	0.47	0.20
Hydrophobic	—	0.32	—	0.32
H-bond acceptor	—	0.33	—	0.32
r^2_{boot}[h]	—	—	0.987 ± 0.006	0.966 ± 0.011
SEE_{boot}[i]	—	—	0.055 ± 0.035	0.080 ± 0.049
r^2_{LHO}[j]	—	—	0.155	0.326
SD_{LHO}[k]	—	—	0.459	0.420
r^2_{5cv}[l]	—	—	0.432	0.481
$SD_{5\ cv}$[m]	—	—	0.394	0.385

[a] Correlation coefficient.
[b] Standard error of estimate.
[c] Ratio of r^2 explained to unexplained = $r^2/(1 - r^2)$.
[d] Cross-validated correlation coefficient after leave-one-out procedure.
[e] Standard error of prediction.
[f] Predicted correlation coefficient for test set of compounds.
[g] Optimal number of principal components.
[h] Average of correlation coefficient for 30 samplings using bootstrapped procedure.
[i] Average standard error of estimate for 30 samplings using bootstrapped procedure.
[j] Average cross-validated correlation coefficient for 50 runs using leave-half-out (LHO) group.
[k] Standard deviation of average cross-validated correlation coefficient for 50 runs.
[l] Average cross-validated correlation coefficient for 50 runs using five cross-validation groups.
[m] Standard deviation of average cross-validated correlation coefficient for 50 runs.

2.3. CoMFA contour maps

Figure 3a and b shows the steric and electrostatic maps for the CoMFA model on compounds **38**, one of the most potent, and **11**, one of the least active, respectively, serving as reference compounds. Like in the reference compounds, the compounds used in the study have the R_2 substituent located on the right side and the R_1 substituent on the left side except compound **54**, in which both R_1 and R_2 substituents are located on the right side. The green polyhedron locate the regions where bulky substituents would increase the MCD inhibitory activity and the yellow polyhedron indicate regions where steric bulk would not be tolerated. Furthermore, the blue polyhedron depict the sites favoring positively charged groups and the red polyhedron sites favoring negatively charged groups.

The medium size yellow polyhedron in the vicinity of one of the methyls of the isopropyl group indicate that bulky substituents (R_1) at the α-carbon of the amide function (compounds **31**–**41**) would be detrimental. This region of unfavorable steric interactions appears to account for the low inhibitory activity of compounds **31** (IC_{50} = 0.15 μM) and **33** (IC_{50} = 0.394 μM) relative to compound **38** (IC_{50} = 0.008 μM). Whereas in the former two compounds the R_1 substituent is represented by a phenethyl and pyridine-4-yl group, respectively, in the latter it is represented by a sterically more favorable isopropyl group. However, although the R_1 substituent of compounds **43**–**53** is a morpholine ring, which is larger than an isopropyl group, these compounds still show reasonable MCD inhibitory activity probably because the morpholine ring lies outside the sterically restricted yellow polyhedron. Other medium size yellow polyhedron, seen on the upper right side of the image, indicate the presence of unfavorable R_2 bulky substituents such as pyridinyl ring of compound **8**. A cluster of three large yellow polyhedron, underneath the R_2 substituent, indicate occupancy by bulky substituents responsible for the poor inhibitory activity of compounds **11** (IC_{50} = 0.465 μM), **14** (IC_{50} = 0.168 μM), **15** (IC_{50} = 0.155 μM), and **16** (IC_{50} = 0.116 μM). Compounds **11** and **14**–**16** have their R_2 substituents (hydroxybenzyl, cyanopentyloxymethyl, tetrazolylpentyloxymethyl, and tetrazolylbutyloxymethyl, respectively) oriented toward this sterically unfavorable region. It is clear that R_2 substituents have specific requirements to be highly active, for example, substituents such as methyl in compounds **32**, **33**, and **43** and hydroxyethyl in compound **49** are poorly active. Although the structural variation of the R_2 substituent between compounds **44** (R_2 = $-CH_2CH_3$; IC_{50} = 0.074 μM) and **49** (R_2 = $-CH_2CH_2OH$; IC_{50} = 0.166 μM) is minor, they exhibit quite different MCD inhibitory potency since the ethyl group of compound **44** occupies a sterically favorable polyhedron while the

Table 2. Observed and predicted activities for the training set from CoMFA and CoMSIA models

Compound	Observed pIC_{50}	CoMFA		CoMSIA	
		Predicted	Residual	Predicted	Residual
1	0.33	0.41	0.08	0.57	0.24
2	0.41	0.42	0.01	0.46	0.05
3	0.38	0.38	0.00	0.36	−0.02
4	0.60	0.58	−0.02	0.40	0.20
5	0.47	0.42	−0.05	0.36	−0.11
6	0.39	0.46	0.07	0.41	0.02
7	0.54	0.53	−0.01	0.39	−0.15
8	0.41	0.41	0.00	0.43	0.02
9	1.13	1.16	0.03	1.15	0.02
10	1.23	1.27	0.04	1.36	0.13
11	0.33	0.31	−0.02	0.44	0.11
12	0.56	0.52	−0.04	0.57	0.01
13	0.97	0.95	−0.02	1.14	0.17
14	0.77	0.76	−0.01	0.60	−0.17
15	0.81	0.85	0.04	0.71	−0.10
16	0.94	0.89	−0.05	1.04	0.10
17	0.47	0.37	−0.10	0.46	−0.01
18	0.82	0.76	−0.06	0.61	−0.21
19	0.93	0.83	−0.10	0.82	−0.11
20	0.68	0.68	0.00	0.81	0.13
21	1.36	1.48	0.12	1.38	0.02
22	0.72	0.77	0.05	0.67	−0.05
23	0.50	0.50	0.00	0.62	0.12
24	0.85	0.83	−0.02	0.91	0.06
25	1.19	1.29	0.10	1.10	−0.09
26	1.43	1.38	−0.05	1.43	0.00
27	0.62	0.65	0.03	0.72	0.10
28	0.79	0.90	0.11	0.92	0.13
29	1.06	1.08	0.02	1.05	−0.01
30	1.39	1.44	0.05	1.30	−0.09
31	0.82	0.84	0.02	1.16	0.34
32	0.75	1.48	0.73	1.52	0.77
33	0.40	0.46	0.06	0.68	0.28
34	1.39	1.43	0.04	1.40	0.01
35	1.55	1.51	−0.04	1.56	0.01
36	1.70	1.69	−0.01	1.51	−0.19
37	2.10	2.07	−0.03	2.12	0.02
38	2.10	2.07	−0.03	2.07	−0.03
39	1.70	1.60	−0.04	1.57	−0.13
40	1.43	1.36	−0.07	1.31	−0.12
41	1.02	0.91	−0.11	0.92	−0.10
42	1.22	1.21	−0.01	1.34	0.12
43	0.49	1.25	0.76	1.40	0.91
44	1.13	1.24	0.11	1.16	0.03
45	1.47	1.41	−0.06	1.49	0.02
46	1.72	1.71	−0.01	1.38	−0.34
47	1.51	1.50	−0.01	1.47	−0.04
48	1.14	1.17	0.03	1.26	0.12
49	0.78	0.78	0.00	0.87	0.09
50	1.00	1.09	0.09	0.91	−0.09
51	0.61	0.55	−0.06	0.78	0.17
52	1.38	1.37	−0.01	1.19	−0.19
53	1.38	1.38	0.00	1.48	0.10
54	0.54	0.55	0.01	0.39	−0.15

hydroxyethyl group of compound **49** does not. The R_2 groups favoring a strong inhibition of MCD contain a 4–5-carbon fragment in an extended conformation (e.g., compounds **37** and **38**). The most potent activity is attained when the R_2 substituent is a 4-carbon fragment with a terminal carbon bearing an acidic or electron withdrawing group such as –COOH (compound **37**), tetrazole (compound **38**), and CN (compound **39**). A small green polyhedron near one of the methyls of the isopropyl group, representing the R_1 substituent, points to the existence of a directional steric requirement in that region of the receptor since the other methyl group of isopropyl moiety is surrounded by yellow polyhedron. Larger size green polyhedron located near the anilino nitrogen atom can accommodate a sterically bulky R_2 group that would increase the inhibitory

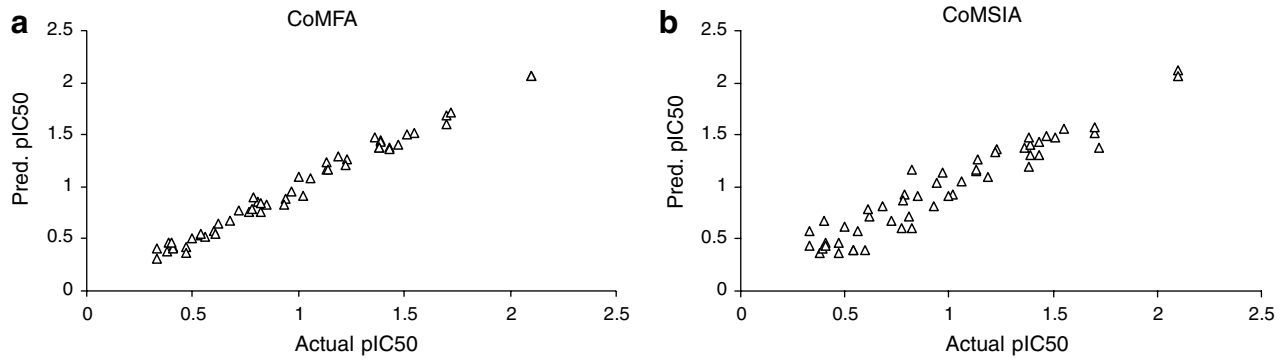


Figure 1. Plot of predicted versus actual pIC_{50} values for the training set molecules based on CoMFA (a) and CoMSIA (b) models.

Table 3. Combinations of different CoMSIA fields and their results[a]

Field(s)	q^2	Components	SEP	r^2	SEE	F value	Contribution				
							S	E	H	D	A
S	0.405	5	0.392	0.904	0.151	87.00	1.00	—	—	—	—
E	0.110	4	0.463	0.783	0.226	42.30	—	1.00	—	—	—
H	0.402	5	0.418	0.883	0.168	69.11	—	—	1.00	—	—
D	0.162	3	0.451	0.562	0.317	20.55	—	—	—	1.00	—
A	0.064	4	0.481	0.676	0.276	24.55	—	—	—	—	1.0
SE	0.417	4	0.388	0.906	0.149	112.78	0.45	0.55	—	—	—
SH	0.524	4	0.375	0.903	0.151	109.69	0.36	—	0.64	—	—
SD	0.386	5	0.413	0.890	0.163	74.30	0.40	—	—	0.60	—
SA	0.408	5	0.421	0.931	0.129	124.34	0.36	—	—	—	0.64
EH	0.396	4	0.399	0.879	0.168	85.54	—	0.41	0.59	—	—
ED	0.202	3	0.443	0.677	0.272	33.53	—	0.30	—	0.70	—
EA	0.311	4	0.428	0.809	0.212	49.67	—	0.38	—	—	0.62
HD	0.346	4	0.406	0.785	0.225	42.89	—	—	0.45	0.55	—
HA	0.388	3	0.409	0.803	0.213	65.01	—	—	0.48	—	0.52
DA	0.265	4	0.436	0.761	0.237	37.44	—	—	—	0.53	0.47
SEH	0.545	4	0.364	0.918	0.139	131.70	0.25	0.30	0.45	—	—
SED	0.392	5	0.406	0.912	0.145	95.30	0.28	0.29	—	0.43	—
SEA	0.470	4	0.390	0.893	0.158	98.15	0.24	0.28	—	—	0.48
SHD	0.455	5	0.384	0.922	0.137	108.66	0.22	—	0.38	0.40	—
SHA	0.494	5	0.382	0.936	0.124	133.79	0.21	—	0.40	—	0.39
SDA	0.400	4	0.408	0.858	0.183	70.87	0.20	—	—	0.43	0.37
EHD	0.452	5	0.386	0.915	0.143	98.53	—	0.25	0.38	0.37	—
EHA	0.491	3	0.387	0.859	0.180	97.32	—	0.24	0.37	—	0.39
EDA	0.341	4	0.424	0.813	0.210	51.00	—	0.21	—	0.41	0.38
HDA	0.397	5	0.402	0.875	0.173	64.56	—	—	0.32	0.37	0.31
SEHD	0.531	5	0.366	0.939	0.121	140.55	0.17	0.20	0.31	0.32	—
SEHA	0.551	4	0.367	0.918	0.139	131.09	0.16	0.20	0.31	—	0.32
SHDA	0.471	5	0.387	0.915	0.143	99.26	0.15	—	0.27	0.32	0.26
SEDA	0.446	4	0.398	0.877	0.170	83.57	0.16	0.18	—	0.35	0.32
EHDA	0.457	5	0.387	0.900	0.155	82.56	—	0.17	0.27	0.30	0.26
SEHDA	0.523	5	0.373	0.928	0.132	118.19	0.12	0.15	0.24	0.27	0.22

[a] Abbreviations: S (steric); E (electrostatic); H (hydrophobic); D (H-bond donor); A (H-bond acceptor).

activity. This is the case of compounds **35** (IC_{50} = 0.028 μM) and **36** (IC_{50} = 0.020 μM), whose R_2 substituents (*n*-propyl and *n*-butyl, respectively) are oriented toward this green polyhedron. The three small yellow polyhedron observed around 1,1,1,3,3,3-hexafluoro-2-propanol subunit (common structure in all the compounds studied here), two being located around each of the $–CF_3$ groups and one around the hydroxyl-bearing carbon, show that increase in steric bulk in these regions will decrease the activity. In contrast, two small green polyhedron on the right side of the image indicate the presence of sterically favorable bulky R_2 substituents (like carboxybutyl in compound **37** and tetrazolylbutyl in compound **38**) with a strong positive effect on MCD inhibition.

A red polyhedraon near the aniline nitrogen atom is occupied by an electronegative group which is favorable for MCD inhibitory activity. Most of the highly active compounds having nitrogen atom in this position explain the requirement for electronegative group. Other red polyhedra on the right side of the image suggest that electronegative groups are favorable for MCD inhibitory activity. An electronegative functional group like ester

Table 4. Observed and predicted activities for the test set by CoMFA and CoMSIA models

Compound	Observed pIC_{50}	CoMFA		CoMSIA	
		Predicted	Residual	Predicted	Residual
1	0.70	0.71	0.01	0.70	0.00
2	0.83	0.91	0.08	0.88	0.05
3	0.94	0.81	−0.13	0.89	−0.05
4	0.52	0.67	0.15	0.77	0.25
5	0.46	0.80	0.34	0.64	0.18
6	0.65	0.90	0.25	0.92	0.27
7	1.59	1.20	−0.39	1.28	−0.31
8	1.23	1.33	0.10	1.28	0.05
9	1.52	1.30	−0.22	1.27	−0.25
10	1.37	1.45	0.08	1.55	0.18
11	1.12	1.28	0.16	1.14	0.02
12	1.15	0.98	−0.17	0.94	−0.21
13	1.70	1.56	−0.14	1.24	−0.46
14	1.27	1.58	0.31	1.41	0.14
15	1.38	1.39	0.01	1.19	−0.19
16	0.75	0.94	0.19	0.95	0.20
17	0.95	0.94	−0.01	0.92	−0.03
18	0.93	0.63	−0.30	0.99	0.06
19	0.82	0.64	−0.18	0.64	−0.18
20	0.62	0.30	−0.32	0.90	0.28
21	0.52	0.73	0.21	0.83	0.31
22	0.25	0.35	0.10	0.36	0.11
23	0.81	0.68	−0.13	0.80	−0.01
24	0.67	0.52	−0.15	0.72	0.05

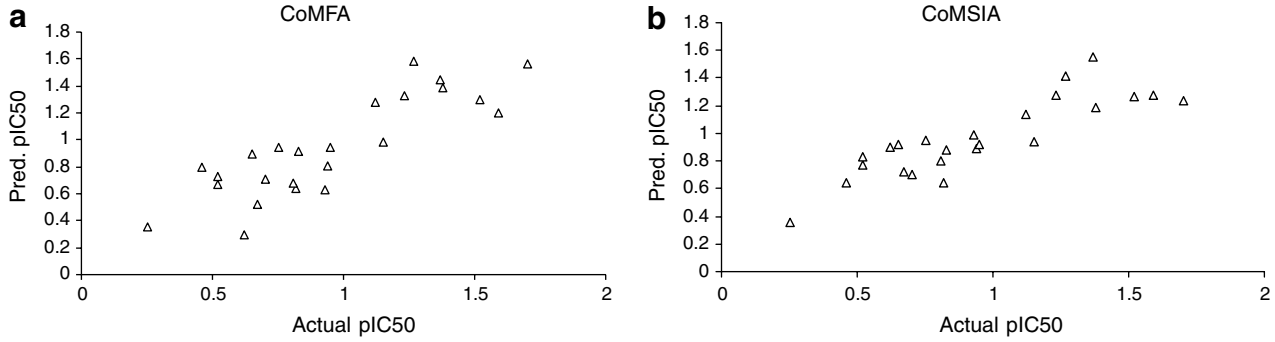


Figure 2. Plot of predicted versus actual pIC_{50} values for the test set molecules based on CoMFA (a) and CoMSIA (b) models.

at R_2 substituent in compound **53** (IC_{50} = 0.042 μM) is oriented in this region. A blue polyhedra surrounded by a yellow one is present in the vicinity of one of the methyls of the isopropyl group and is associated with an electropositive R_1 substituent with limited steric bulk and a positive influence on MCD inhibition. This situation is exemplified by the pyridine ring nitrogen of the moderately active compound **22**. Another blue polyhedron on the right side of the image shows that electropositive groups would favor the MCD inhibitory activity. The cyano group of R_2 substituent in moderately active compounds **7** and **24** is located in this blue polyhedron.

2.4. CoMSIA contour maps

The CoMSIA steric and electrostatic contour maps (figures not shown) are rather similar to the corresponding CoMFA contour maps. Figures 4a and b are hydrophobic maps for the CoMSIA model for compound **38**, one of the most potent compounds, and for compound **11**, one of the least potent respectively. The orange and white polyhedra highlight areas where hydrophobic and hydrophilic properties are preferred, respectively. The large orange polyhedron seen in the vicinity of isopropyl group at R_1 position indicate that hydrophobic groups are preferred for MCD inhibitory activity. Most of the highly active compounds bear an isopropyl substituent at R_1 position. In contrast, the poor MCD inhibitory activity of compound **11** may stem from the presence of a hydroxyl group in the small orange polyhedron seen on the right side of the image. By lacking a hydrophobic R_2 substituent near this favorable orange polyhedra, compound **23** (IC_{50} = 0.318 μM) manifest a very poor inhibitory activity. Small white polyhedra at the aniline nitrogen atom indicate the preference for favorable hydrophilic/polar nitrogen containing groups (i.e., nitrogen atom of amide, urea or imidazole) in this region as in the case of compounds **9**–**53**. In contrast, the weakly active compounds **1**–**8** exhibit a carbon instead of a nitrogen atom in the same region.

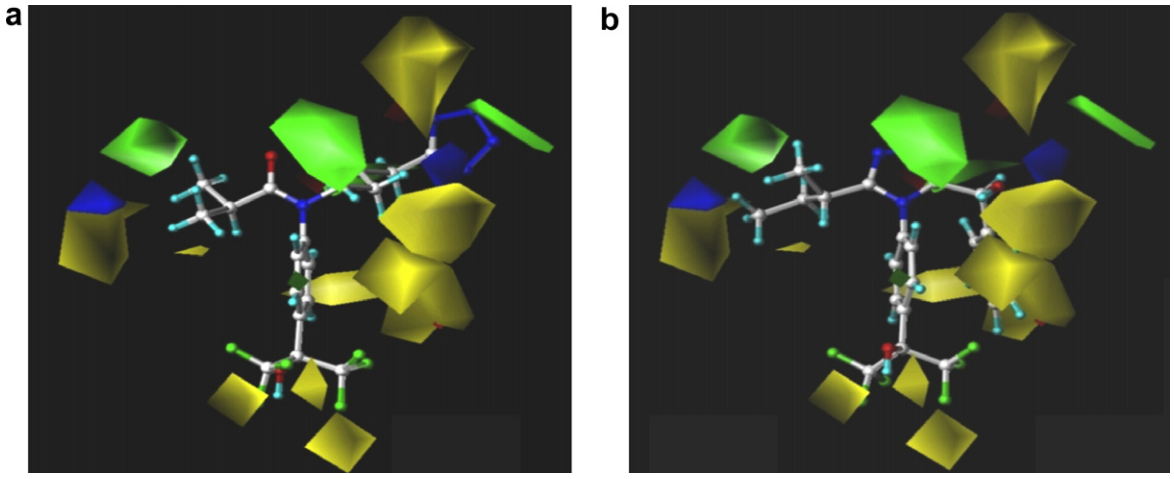


Figure 3. CoMFA coefficient contour map (STDEV*COEFF) for steric and electrostatic fields. Compounds **38** (a) and **11** (b) are displayed as references. Steric fields: green contours indicate regions where bulky groups increase activity; yellow contours indicate regions where bulky groups decrease activity. Electrostatic fields: blue contours indicate regions where electropositive groups increase activity; red contours indicate regions where electronegative groups increase activity.

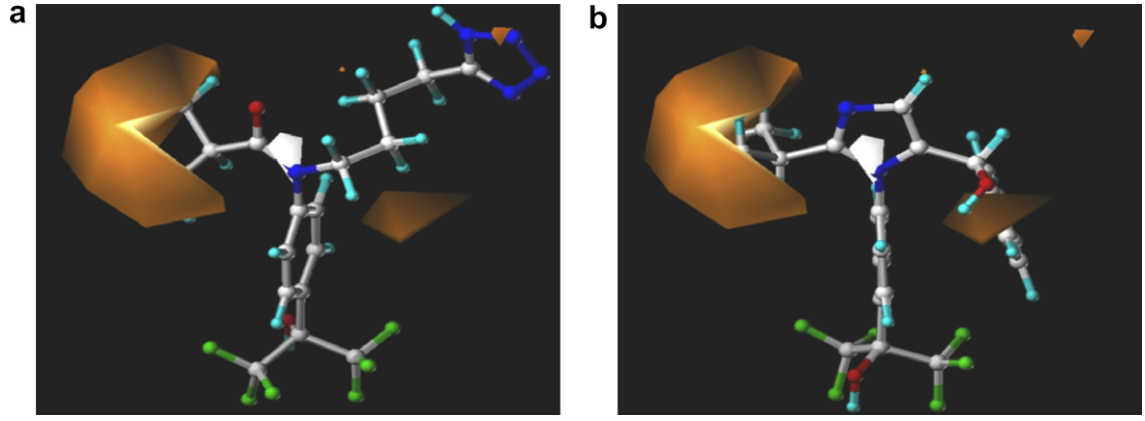


Figure 4. CoMSIA coefficient contour map (STDEV*COEFF) for hydrophobic fields. Compounds **38** (a) and **11** (b) are displayed as references. Contour maps in orange and white indicate favorable and unfavorable hydrophobic groups, respectively, for MCD inhibitory activity.

The hydrogen-bond acceptor maps for the CoMSIA model for compounds **38** and **11**, used as prototype, are shown in Figure 5a and b, respectively. The magenta and cyan polyhedra indicate areas occupied by hydrogen-bond acceptor groups exerting favorable and unfavorable effects on MCD inhibition, respectively. A large magenta polyhedron in the vicinity of the R_2 substituents indicates the location of hydrogen-bond acceptor groups having an enhancing effect on MCD inhibition (carboxyl in compound **37** and tetrazolyl in compound **38**). This type of substituent is missing from compound **11**, the weakly active compound. Another medium size magenta polyhedron near the isopropyl group suggests the presence of hydrogen-bond acceptor groups in this region would increase MCD inhibitory activity. The morpholine ring oxygen of moderately active compounds **45**–**47** and **52**–**53** was mapped onto same polyhedron. The cyan polyhedron located onto –OH group of 1,1,1,3,3,3-hexafluoro-2-propanol subunit indicate the presence of hydrogen-bond acceptor in this region is unfavorable for MCD inhibitory activity. The cyan polyhedron seen on the lower right side of the image shows that hydrogen-bond acceptor in this region will decrease the MCD inhibitory activity. The cyano group of moderately active compounds **7** and **24** is located within these unfavorable cyan polyhedron. Another cyan polyhedron seen on the upper right side of the image indicates the presence of hydrogen-bond acceptor in this region is unfavorable for the MCD inhibitory activity. Isoxazole ring nitrogen of the poorly active compounds **1**–**8** is surrounded by these cyan polyhedron.

In short, based on the results of the 3D-QSAR analyses undertaken in the present study, it is quite evident that a close agreement exists between the structural features of the inhibitors examined and their contour maps. The absence of colored regions in some areas of the contour map does not necessarily imply that these regions do not contribute to MCD inhibitory activity but, rather, they contain the same or similar functionality.

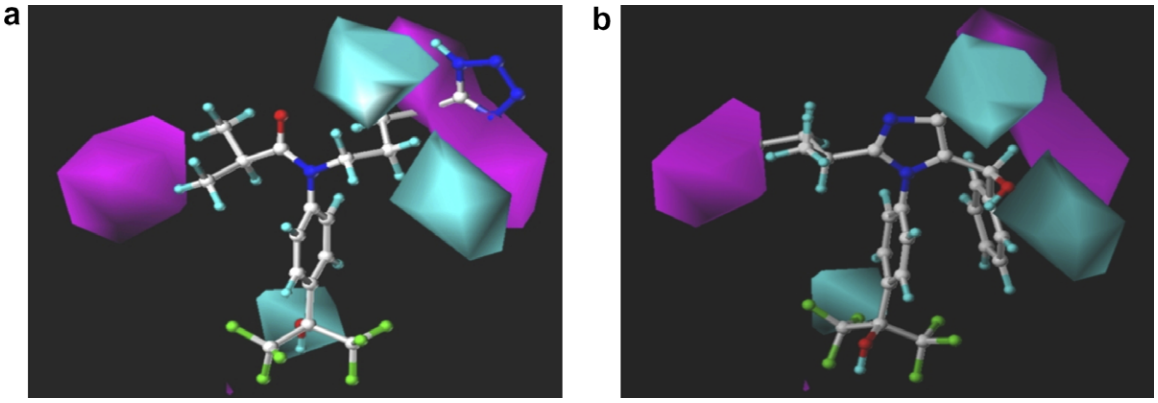


Figure 5. CoMSIA coefficient contour map (STDEV*COEFF) for hydrogen-bond acceptor fields. Compounds **38** (a) and **11** (b) are displayed as references. Contour maps in magenta and cyan indicate favorable and unfavorable hydrogen-bond acceptor groups, respectively, for MCD inhibitory activity.

3. Conclusions

3D-QSAR models for structurally diverse 1,1,1,3,3,3-hexafluoro-2-propanol derivatives with statistical significance and predictive abilities by using CoMFA and CoMSIA methods were developed and validated. A high bootstrapped r^2 value, 0.987 and 0.966 for CoMFA and CoMSIA, respectively, with a small standard deviation indicates the existence of a similar relationship among all compounds used to build the model. In addition to steric and electrostatic fields, hydrophobic and hydrogen-bond acceptor fields are also found to be important for MCD inhibitory activity since the predictive power of the CoMSIA model appeared to be superior to that of CoMFA model. The significant predictive ability of 3D QSAR models observed for the external test-set of molecules makes these models useful for designing new compounds with good MCD inhibitory properties.

4. Materials and methods

4.1. Data sets and biological activity

The training (52 compounds) and test (24 compounds) sets used in this study comprise a series of 1,1,1,3,3,3-hexafluoro-2-propanol derivatives, previously shown to be potent inhibitors of MCD.[12–15] The IC_{50} values were converted to pIC_{50} ($-\log IC_{50}$, in units of μM) and used as dependent variables in CoMFA and CoMSIA QSAR analyses. The compounds in the test set were novel in terms of structural diversity, range of biological activity, as well as it allows to use one test compound per two training set compounds, thus resulting in a rigorous validation of the training model than would be possible with a smaller test set. The distribution of biological activities of training set and test set versus % of compounds is shown in Figure 6. The structures of the training and test set of compounds are shown in Tables 5 and 6, respectively.

4.2. Molecular modeling and alignment

All computational studies were performed using SYBYL version 7.2[16] on a Dell Precision 470 n workstation

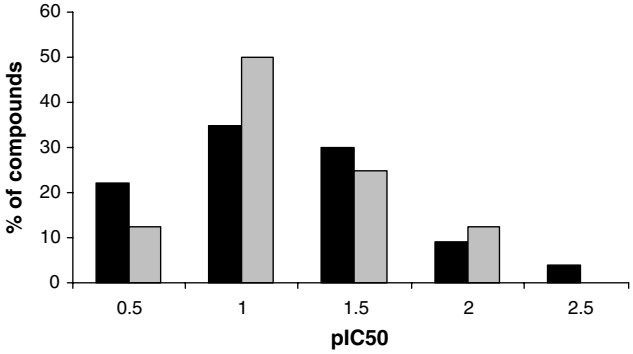
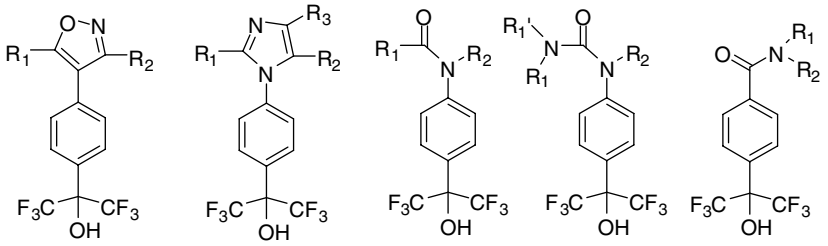


Figure 6. Distribution of biological activities for training set (black bars) and test set (gray bars) versus % of compounds.

with the RHEL 4.0 operating system. The molecular conformation and alignment of molecules are two sensitive input parameters affecting the CoMFA model. The 3D structures of the compounds in the training and test sets were constructed using the Sketch Molecule function in SYBYL. Energy minimizations were performed with the conjugate gradient method using the Tripos force field[17] and Gasteiger-Marsili charges[18] with a convergence criterion of 0.001 kcal/mol Å. Each structure was then subjected to simulated annealing as it enables the rapid identification of the good solutions, ideally the global minimum.[19] The system was heated at 1000 K for 1 ps and then cooled at 200 K for 1 ps. The exponential annealing function was used and 10 such cycles were run. Using this method, the conformer in the training and test set compounds with the least energy content was determined and subsequently subjected to further minimization with the same criteria as mentioned above and used for alignment.

CoMFA and CoMSIA studies require that the 3D structures of the compounds to be analyzed should be aligned according to a suitable conformational template, which is assumed to be a bioactive conformation.[7] Since at present structural information on these inhibitor–protein complexes is not available; one of the most potent molecules (compound **38**) was chosen as a template to fit the rest of the training and test set compounds by

Table 5. Structures of 54 MCD inhibitors used as the training set[a]

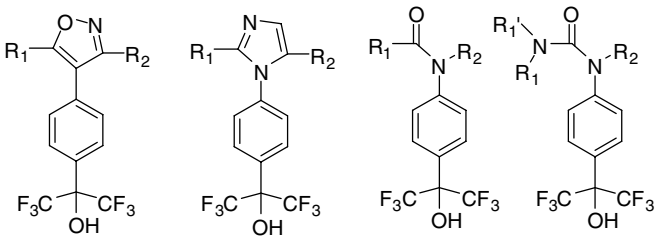


1-8 **9-30** **31-41** **42-53** **54**

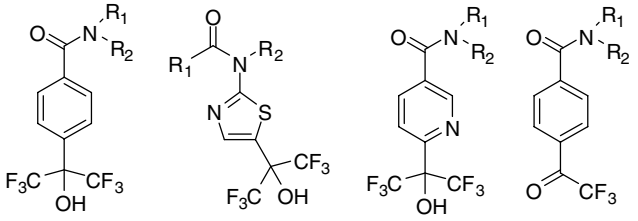
Compound	R_1	R_2	R_3	IC_{50} (μM)
1	H_2N–	–*i*-Pr	—	0.460
2	4-(CN-Ph)CONH–	–*i*-Pr	—	0.390
3	PhCH(Et)CONH–	–*i*-Pr	—	0.420
4	*i*-PrCONH–	–Pyridin-4-yl	—	0.250
5	4-PyCONH–	*i*-Pr	—	0.340
6	Me–	–Me	—	0.410
7	Me–	–*trans*-CH=CH–CN	—	0.290
8	Me–	–NHCO(4-Py)	—	0.390
9	s-Bu–	–CH(CO)(*n*-Pr)	–H	0.074
10	s-Bu–	–CH(OH)(*i*-Pr)	–H	0.059
11	s-Bu–	–CH(OH)Ph	–H	0.465
12	*i*-Pr–	$–CH_2OH$	–H	0.274
13	*i*-Pr–	$–CH_2O(CH_2)_4CN$	–H	0.107
14	*i*-Pr	$–CH_2O(CH_2)_5CN$	–H	0.168
15	*i*-Pr–	$–CH_2O(CH_2)_5$-(5-tetrazole)	–H	0.155
16	*i*-Pr–	$–CH_2O(CH_2)_4$-(5-tetrazole)	–H	0.116
17	s-Bu–	$–CH_2NHCH_2CH_2OMe$	–H	0.340
18	*i*-Pr–	=N–OH	–H	0.152
19	*i*-Pr–	=N–OEt	–H	0.118
20	*i*-Pr–	=N–OPh	–H	0.209
21	*i*-Pr–	=N–O(t-Bu)	–H	0.044
22	4-Py–	–H	–COOEt	0.191
23	s-Bu	–H	–COOEt	0.318
24	*i*-Pr–	–*trans*-CH=CH–CN	–H	0.140
25	*i*-Pr–	$–CH_2CH_2COOMe$	–H	0.065
26	*i*-Pr–	$–CH_2CH(Me)COOEt$	–H	0.037
27	*i*-Pr–	$–CH_2CH_2COOH$	–H	0.239
28	s-Bu–	$–CH_2CH_2COOH$	–H	0.163
29	s-Bu–	–*trans*-CH=CH–COOMe	–H	0.088
30	s-Bu–	$–CH_2CH_2COOMe$	–H	0.041
31	$PhCH_2CH_2$–	–Cyclohexyl	—	0.150
32	*i*-Pr	–Me	—	0.178
33	4-Pyridinyl	–Me	—	0.394
34	*i*-Pr–	–Et	—	0.041
35	*i*-Pr–	–*n*-Pr	—	0.028
36	*i*-Pr–	–*n*-Bu	—	0.020
37	*i*-Pr–	$–(CH_2)_4COOH$	—	0.008
38	*i*-Pr–	$–(CH_2)_4$-(5-1H-tetrazole)	—	0.008
39	*i*-Pr–	$–(CH_2)_4CN$	—	0.020
40	-p-CNPh	–Et	—	0.037
41	4-Pyridinyl	–Et	—	0.095
42	–Et, $R_1' = –(CH_2)_4OH$, R_1 = –H	–Et	—	0.060
43	$–CH_2CH_2OCH_2CH_2$–	–Me	—	0.321
44	$–CH_2CH_2OCH_2CH_2$–	–Et	—	0.074
45	$–CH_2CH_2OCH_2CH_2$–	–*n*-Pr–	—	0.034
46	$–CH_2CH_2OCH_2CH_2$–	–*n*-pentyl	—	0.019
47	$–CH_2CH_2OCH_2CH_2$–	–*n*-hexyl	—	0.031
48	$–CH_2CH_2OCH_2CH_2$–	–*n*-heptyl	—	0.073
49	$–CH_2CH_2OCH_2CH_2$–	$–CH_2CH_2OH$	—	0.166
50	$–CH_2CH_2OCH_2CH_2$–	$–CH_2CH_2COOMe$	—	0.099
51	$–CH_2CH_2OCH_2CH_2$–	$–CH_2CH_2COOH$	—	0.247
52	$–CH_2CH_2OCH_2CH_2$–	$–(CH_2)_3CN$	—	0.040
53	$–CH_2CH_2OCH_2CH_2$–	$–(CH_2)_4COOMe$	—	0.042
54	–Et	$–CH_2CH_2CN$	—	0.290

[a] In compounds **43**–**53**, R_1 and R_1' are cyclized into a morpholine ring.

Table 6. Structures of 24 MCD inhibitors used as the test set[a]



1 **2-6** **7** **8-11**



12 **13-17** **18** **19-24**

Compound	R_1	R_2	IC_{50} (μM)
1	*n*-BuCH(Et)CONH–	–*i*-Pr	0.200
2	s-Bu–	–CH(OH)CH_3	0.149
3	Pyridin-4-yl	–CH_2OH	0.114
4	s-Bu–	–CH_2NH(*n*-Bu)	0.303
5	*i*-Pr–	=N–*O*-(*i*-Bu)	0.350
6	*i*-Pr–	–CH_2CH(Me)COOH	0.225
7	*i*-Pr–	–$(CH_2)_4$COOMe	0.026
8	–Me, R_1' = –Me, R_1 = –H	–Et	0.059
9	–$CH_2CH_2OCH_2CH_2$–	–*n*-Bu	0.030
10	–$CH_2CH_2OCH_2CH_2$–	–CH_2CH_2Ph	0.043
11	–$CH_2CH_2OCH_2CH_2$–	–$CH_2CH_2CH_2$COOMe	0.075
12	*i*-Bu-	–*i*-Bu	0.070
13	*i*-Pr–	4-CNBn	0.020
14	*i*-Pr–	4-MeO_2CBn–	0.054
15	*i*-Pr–	4-Pyridinylmethyl	0.042
16	Me	–*n*-Bu	0.177
17	4-Pyridinyl	–*n*-Bu	0.111
18	*n*-Pr–	Cyclopropylmethyl	0.117
19	*n*-Pr–	–*n*-Pr	0.153
20	H	–CH(CH_3)CH_2CH(CH_3)$_2$	0.242
21	H	2-Methylcyclohexyl	0.304
22	H	–CH_2CH_2Ph	0.566
23	Ph	–CH_2CH_2CN	0.155
24	Et	–CH_2CH_2CN	0.214

[a] In compounds **9**–**11**, R_1 and R_1' are cyclized into a morpholine ring.

using the Sybyl fit atoms function. The reference atoms in compound **38** used for alignment were: (i) C-1, C-2, and C-3 of the 1,1,1,3,3,3-hexafluoro-2-propanol moiety; (ii) C-4 and C-5 of the phenyl ring; (iii) N-6, C-7, O-8, and C-9 of the 2-methylpropanamide moiety; and (iv) C-10 and C-11 of the tetrazolylbutyl substituent on the amide nitrogen (Fig. 7a). The alignment of the training and test set compounds is shown in Figure 7b.

4.3. CoMFA and CoMSIA 3D QSAR models

In deriving the CoMFA and CoMSIA descriptor fields, a 3D cubic lattice with grid spacing of 2 Å in *x*, *y*, and *z* directions was created to encompass the aligned molecules. CoMFA descriptors were calculated using an sp^3

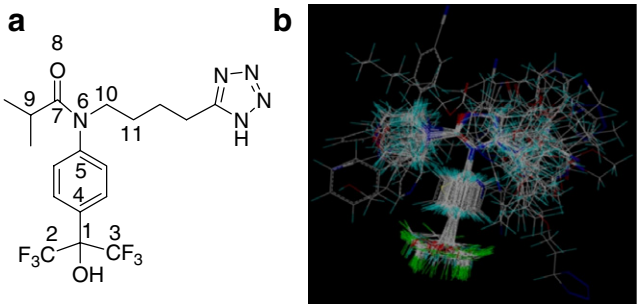


Figure 7. (a) Compound **38** used as a template for atom based alignment. The atoms for alignment are numbered 1–11. Note that the atom numbering does not follow the IUPAC rules. (b) Training and test sets aligned on minimum energy conformation of compound **38**.

carbon probe atom with a van der Waals radius of 1.52 Å and a charge of +1.0 to generate steric (Lennard-Jones 6–12 potential) field energies and electrostatic (Coulombic potential) fields with a distance-dependent dielectric at each lattice point. Values of steric and electrostatic energy were truncated at 30 kcal/mol. The CoMFA steric and electrostatic fields thus generated were scaled by the CoMFA-STD method in SYBYL.

CoMSIA descriptors were derived according to Klebe et al.[9] with the same lattice box as that used for the CoMFA calculations, with a grid spacing of 2 Å and employing a C^{1+} probe atom with a radius of 1.0 Å as implemented in SYBYL. CoMSIA similarity indices (A_F) for a molecule j with atoms i at a grid point q were calculated using Eq. 1:

$$A^{q}_{F,k}(j) = -\sum \omega_{probe,k}\omega_{ik} e^{-\alpha r}_{iq2} \quad (1)$$

Being an extension to the CoMFA approach which has two fields, the CoMSIA method incorporates five different physicochemical properties (steric, electrostatic, hydrophobic, hydrogen bond donor, and hydrogen bond acceptor), were evaluated using the probe atom. A Gaussian type distance-dependence was which used between the grid point q and each atom i in the molecule. A default value of 0.3 was used as the attenuation factor α. In CoMSIA, the steric indices are related to the third power of the atomic radii, the electrostatic descriptors are derived from partial atomic charges, the hydrophobic fields are derived from atom based parameters,[20] and the hydrogen-bond donor and acceptor atoms within a putative protein environment are derived from experimental values.[21,22]

The CoMFA and CoMSIA descriptors were used as independent variables and the pIC_{50} values were used as dependent variables in PLS regression analyses[23,24] to derive 3D-QSAR models using the standard implementation in the SYBYL package. The predictive value of the models was evaluated first by leave-one-out (LOO) cross-validation.[8,25] The cross-validated coefficient, q^2, was calculated using Eq. 2:

$$q^2 = 1 - \frac{\sum(Y_{predicted} - Y_{observed})^2}{\sum(Y_{observed} - Y_{mean})^2} \quad (2)$$

where $Y_{predicted}$, $Y_{observed}$, and Y_{mean} are predicted, actual, and mean values of the target property (pIC_{50}), respectively. $\sum(Y_{predicted} - Y_{observed})^2$ is the predictive residual sum of squares (PRESS).

The boot strapping analysis[26] and the number of cross-validations (e.g., 2 and 5) were carried out and confirmed by the average value for 50 runs from each cross-validation. To test the utility of the model as a predictive tool, an external set of compounds with known activities but not used in model generation (the test set) were predicted. The predictive r^2, calculated by using Eq. 3, was based on molecules from the test set and it was used to evaluate the predictive power of the CoMFA and CoMSIA models.

$$\text{Predictive } r^2 = 1 - (\text{'press'}/\text{SD}) \quad (3)$$

where SD is the sum of the squared deviations between the actual activities of the compounds in the test set and the mean activity of the compounds in the training set and press is the sum of the squared deviations between predicted and actual activities for every compound in the test set. A predictive r^2 value of 1 signifies that the CoMFA model is perfectly predictive for the test set, while prediction of a mean value of the training set for every member in the test set yields a predictive $r^2 = 0$. The activity of the test set was predicted by the CoMFA and CoMSIA models using the predict command. CoMFA and CoMSIA coefficient maps were generated by interpolation of the pair-wise products between the PLS coefficients and the standard deviations of the corresponding CoMFA or CoMSIA descriptor values.

Acknowledgments

T.T. wishes to thank Dean Robert Mangione and Dr. Louis Trombetta for providing start-up funds to purchase Sybyl software package. We thank Drs. Cesar Lau-Cam, Santosh Kulkarni, and Vijay Gokhale for helpful comments and suggestions.

References and notes

1. (a) Kim, Y. S.; Kolattukudy, P. E. *Arch. Biochem. Biophys.* **1978**, *190*, 234; (b) Kim, Y. S.; Kolattukudy, P. E. *Arch. Biochem. Biophys.* **1978**, *190*, 585; (c) Kim, Y. S.; Kolattukudy, P. E. *Biochim. Biophys. Acta* **1978**, *531*, 187; (d) Kim, Y. S.; Kolattukudy, P. E.; Boos, A. *Int. J. Biochem.* **1979**, *10*, 551; (e) Kim, Y. S.; Kolattukudy, P. E.; Boos, A. *Arch. Biochem. Biophys.* **1979**, *196*, 543; (f) Hunaiti, A. R.; Kolattukudy, P. E. *Arch. Biochem. Biophys.* **1984**, *229*, 426; (g) Voilley, N.; Roduit, R.; Vicaretti, R.; Bonny, C.; Waeber, G.; Dyck, J. R. B.; Lopaschuk, G. D.; Prentki, M. *Biochem. J.* **1999**, *340*, 213.
2. Hamilton, C.; Saggerson, E. D. *Biochem. J.* **2000**, *350*, 61.
3. Zammit, V. A. *Biochem. J.* **1999**, *343*, 505.
4. (a) Randle, P. J.; Garland, P. B.; Hales, C. N.; Newsholme, E. A. *Lancet* **1963**, *1*, 785; (b) Randle, P. J. *Diabetes Metab. Rev.* **1998**, *14*, 263.
5. Vavvas, D.; Apazdis, A.; Saha, A. K.; Gamble, J.; Patel, A.; Kemp, B. E.; Witters, L. A.; Ruderman, N. B. *J. Biol. Chem.* **1997**, *272*, 13255.
6. Dyck, J. D.; Cheng, J. F.; Stanley, W. C.; Barr, R.; Chandler, M. P.; Brown, S.; Wallace, D.; Arrhenius, T.; Harmon, C.; Yang, G.; Nadzan, A.; Lopaschuk, G. D. *Circ. Res.* **2004**, *94*, 78.
7. Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
8. Cramer, R. D., III; Bunce, J. D.; Patterson, D. E.; Frank, I. E. *Quant. Struct.-Act. Relat.* **1988**, *7*, 18.
9. Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 4130.
10. Bringmann, G.; Rummey, C. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 304.
11. Bohm, M.; Strzebecher, J.; Klebe, G. *J. Med. Chem.* **1999**, *42*, 458.

12. Cheng, J. F.; Chen, M.; Wallace, D.; Tith, S.; Haramura, M.; Liu, B.; Mak, C. C.; Arrhenius, T.; Reily, S.; Brown, S.; Thorn, V.; Harmon, C.; Barr, R.; Dyck, J. D.; Lopaschuk, G. D.; Nadzan, A. *J. Med. Chem.* **2006**, *49*, 1517.
13. Cheng, J. F.; Chen, M.; Liu, B.; Hou, Z.; Arrhenius, T.; Nadzan, A. M. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 695.
14. Wallace, D. M.; Haramura, M.; Cheng, J. F.; Arrhenius, T.; Nadzan, A. M. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 1127.
15. Cheng, J. F.; Mak, C. C.; Huang, Y.; Penuliar, R.; Nishimoto, M.; Zhang, L.; Chen, M.; Wallace, D.; Arrhenius, T.; Chu, D.; Yang, G.; Barbosa, M.; Barr, R.; Dyck, J. R. B.; Lopaschuk, G. D.; Nadzan, A. M. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3484.
16. SYBYL 7.2, Tripos Inc., 1699 South Hanley Rd., St. Louis, Missouri, 63144.
17. Clark, M.; Cramer, R. D., III; Van Opdenbosch, N. *J. Comput. Chem.* **1989**, *10*, 982.
18. Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.
19. Barakat, M. T.; Dean, P. M. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 295.
20. Viswanadhan, V. N.; Ghose, A. K.; Revenkar, G. R.; Robins, R. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 163.
21. Klebe, G. *J. Mol. Biol.* **1994**, *237*, 212.
22. Klebe, G.; Mietzner, T.; Weber, F. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 35.
23. Wold, S.; Albano, C.; Dunn, W. J.; Edlund, U.; Esbenson, K.; Geladi, P.; Hellberg, S.; Lindburg, W.; Sjostorm, M. In *Chemometrics*; Kowalski, B., Ed.; Reidel: Dordrecht, The Netherlands, 1984; 17.
24. Stahle, L.; Wold, S. *J. Chemom.* **1987**, *1*, 185.
25. Wold, S. *Technometrics* **1978**, *4*, 397.
26. Vong, R.; Geladi, P.; Wold, S.; Esbensen, K. *J. Chemom.* **1988**, *2*, 281.