Journal of Molecular Graphics and Modelling 29 (2010) 229–239

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Development of highly predictive 3D-QSAR CoMSIA models for anthraquinone and acridone derivatives as telomerase inhibitors targeting G-quadruplex DNA telomere

Vishal P. Zambre, Prashant R. Murumkar, Rajani Giridhar, Mange Ram Yadav*

Pharmacy Department, Faculty of Technology and Engineering, Kalabhavan, The M.S. University of Baroda, Vadodara 390001, Gujarat, India

ARTICLE INFO

Article history:
Received 17 February 2010
Received in revised form 7 July 2010
Accepted 9 July 2010
Available online 15 July 2010

Keywords:
CoMSIA
Anthraquinone
Acridone
Telomerase inhibitors
Intercalators

ABSTRACT

G-quadruplex structures of DNA represent a potentially useful target for anticancer drugs. Telomerase enzyme, involved in immortalization of cancer cells is inhibited by stabilization of G-quadruplex at the ends of chromosomes. Anthraquinone and acridone derivatives are promising G-quadruplex ligands as telomerase inhibitors. So far, optimization of these ligands remained hampered due to the lack of creditable quantitative structure–activity relationships. To understand the structural basis of anthraquinone and acridone derivatives, a predictive 3D-QSAR model has been developed for the first time for telomerase inhibitory activity of G4 ligands, employing comparative molecular similarity indices analysis (CoMSIA). Considering the proposition that the basic nitrogens in these compounds should exist in protonated form at physiological pH the protonated forms of the reported compounds were analyzed and investigated. The QSAR model from conformational template Conf1 exhibited best correlative and predictive properties. The actual predictive abilities of the QSAR model were thoroughly validated through an external validation test set of compounds. The statistics indicate a significantly high prediction power of the best model (r^2, 0.721), supporting the proposed molecular mechanism of DNA G-quadruplex ligands.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Telomerase is a complex ribonucleoprotein reveres transcriptase enzyme. RNA template (hTR) and a catalytic protein domain (hTERT) are two major components of human telomerase [1]. Telomeric DNA consists of repetitive guanine-rich sequences which forms secondary structures based on reverse Hoogsteen-type base pairing involving four guanines in a planar arrangement termed G-tetrad [2,3]. Telomere length progressively shortens in somatic cells with successive rounds of cell division, leading eventually to senescence and apoptosis. In contrast telomere length is maintained in cancer cells and is a major factor in immortalization and tumorigenesis [4]. This happens almost always under activation of the enzyme telomerase [5] which has made it a highly selective target for anti-tumor drug design [6,7]. Various strategies have been proposed to induce telomerase inhibition [8]. A promising strategy is based on the inhibition of telomerase by interacting with G-quadruplex DNA.

G-quadruplex DNA interactive compounds inhibit telomerase *in vitro* by stabilizing single stranded 3′-telomere ends as a quadruplex [9,10]. The best example is a ligand called quarfloxacin (formerly CX-3543) that is now in phase-II clinical trial as an anticancer agent [11].

Substituted acridines and anthraquinones have been reported previously [12,13] to stabilize G-quadruplex. Campbell et al. reported structural basis of DNA quadruplex recognition by an acridine drug (BRACO-19) [14] and also discussed issues related to selectivity in ligand recognition of G-quadruplex loop [15]. Acridone derivatives are another class of ligands that stabilize the G-quadruplex structures possessing comparable telomerase inhibitory activity to acridine derivatives [16].

Harrison et al. [16], Read et al. [17], and Perry et al. [18,19] carried out structure–activity relationship studies of telomerase inhibitors. Effects of amide bond direction on modulation of G-quadruplex recognition and telomerase inhibition by substituted anthracenedione derivatives have studied [20]. Recently, Cuenca et al. postulated that the incorporation of aromatic side chains to the acridone core in the 4 and 5 position would enhance G-quadruplex affinity, although these derivatives did not show telomerase inhibitory activity [21]. A three-dimensional structure–activity relationship (3D-QSAR) as rational basis for the G-quadruplex ligand optimization for anthraquinones and acridones remains unreported. With 3D-QSAR analysis it is possible to analyze the probable structural elements affecting the biological activity of compounds. We have been involved in the computer aided designing of novel acridine derivatives as telomerase inhibitors

* Corresponding author. Tel.: +91 265 2434187; fax: +91 265 2418927.
E-mail addresses: mryadav11@yahoo.co.in, mryadav-phar@msu.ac.in (M.R. Yadav).

1093-3263/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2010.07.003

using CoMFA (comparative molecular field analysis) and CoMSIA (comparative molecular similarity indices analysis) approaches [22]. Various other targets like tumor-necrosis alpha converting enzyme (TACE) [23,24], farnesyltransferase [25–27] and COX-II (cyclooxygenase-II) [28] have been successfully exploited for rational design of novel enzyme inhibitors. Structural requirements for anthraquinone and acridone derivatives in terms of physicochemical descriptors have been studied in this laboratory for their G-quadruplex stabilizing telomerase inhibitory activity using 3D-QSAR techniques. In this paper, we report 3D-QSAR CoMSIA [29] studies from a panel of 61 telomerase inhibitors obtained from literature [16–19]. Considering that the basic nitrogens in these compounds should remain protonated at physiological pH, the protonated forms were analyzed and investigated in the study. Some remarkable observations were made during this study in relation to the structures of these molecules and their putative binding site of G-quadruplex. All these findings are reported in this work, which is a first successful attempt of its kind for lead optimization using 3D-QSAR CoMSIA technique for telomerase inhibitory activity of anthraquinone and acridone derivatives. The model so developed showed high correlation and prediction of G-quadruplex ligands with telomerase inhibitory activity specifying that telomerase inhibition and G-quadruplex stabilization was directly related. This strongly supported the proposed unified mechanism of telomerase inhibition. CoMSIA studies obtained with hydrophobic, H-bond acceptor and H-bond donor fields (HAD) afforded the best model with good predictive power outlining the dominant role of these fields for telomerase inhibitory activity of the studied G-quadruplex stabilizers. CoMFA [30] studies were also performed for the chosen series of compounds prior to the CoMSIA studies. The CoMFA models (data not shown) obtained were statistically inferior to CoMSIA models, which further confirmed the involvement of hydrophobic interactions as dominant field for anthraquinone and acridone derivatives as telomerase inhibitors.

2. Materials and methods

2.1. Data set

The molecular structures and activities of 61 telomerase inhibitors were taken from the literature [16–19]. The telomerase inhibitory activity of compounds is reported as EC_{50} values in the micro molar (μM) range. The selected compounds cover a wide range of biological activity (0.2–50 μM) and diverse structural features. The reported EC_{50} values were converted into $-\log(EC_{50})$, i.e. pEC_{50} for use in the QSAR studies. The whole set of 61 compounds (Table 1) was divided into training set of 49 and test set of 12 compounds in the process of model refinement for all CoMSIA models reported herein. In the training set, most potent, moderately active and lowly active compounds were included to spread the activity range. The test set compounds were selected in such a manner that at least one structural analog remained in the training set.

2.2. Molecular modeling

A Silicon Graphics Fuel workstation with IRIX 6.5 operating system running SYBYL 7.0 [31] was used for three-dimensional structure building and molecular modeling studies. Three different molecular alignments were carried out in the present study. Initially, two molecular alignments, based on MacroModel Monte Carlo conformational search-derived templates, were derived using centroid and atom-based alignment rule in SYBYL. The two conformational search-based templates were derived as follows. The most active compound **61** was constructed in SYBYL. Its structure was minimized, and then used as a starting point for a Monte Carlo conformational search employing MacroModel version 7.0 (Schrödinger, Inc.). The conformational search was carried out using MMFF94 force field in MacroModel for 5000 iterations using water as a solvent. The global minimum conformation, designated as Conf1, and the second lowest energy conformation from the global minimum, designated as Conf2 were used as templates to construct the rest of the molecules. A third alignment was derived as discussed below:

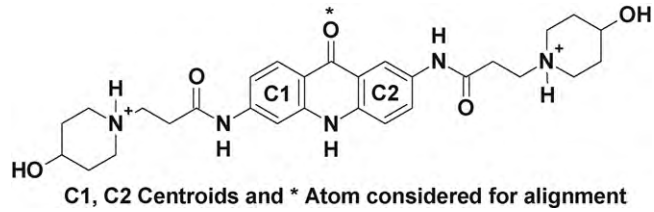


Fig. 1. Template for alignment.

The conformation of a disubstituted acridine derivative (biologically active moiety) was extracted from the co-crystallized structure of G-quadruplex PDB 1L1H [32]. Conf1 of the most active compound **61** was superimposed on the conformation of the acridine derivative, extracted from the co-crystallized structure. As a measure of superimposition reliability, the rmsd between these two conformations was taken into the consideration (0.9 Å obtained herein). It was interesting to note that the resulting conformation (conf3) of compound **61** fully superimposed (rmsd 0.00 Å) on conformation conf1. Hence, conformations conf1 and conf2 only were used as templates for further study. The partial atomic charges required for the electrostatic interactions were computed by the semi empirical molecular orbital method using Molecular Orbital PACkage (MOPAC) [33] with Austin Model 1 (AM1) Hamiltonian [34]. The centroids and atoms considered for alignment are marked with C1, C2 and an asterisk (*), respectively in Fig. 1. The superimposition of all the compounds on template (compound **61**) is shown in Fig. 2.

2.3. pK_a value estimation

Pallas 3.7.1.1 software from CompuDrug International Inc. [35] was used to calculate pK_a values for side chain substituent piperazine, in compounds **11**, **50**, and **54** (Table 2). This program provides an indispensable resource for predicting acidic and basic pK_a values before synthesis in QSAR studies.

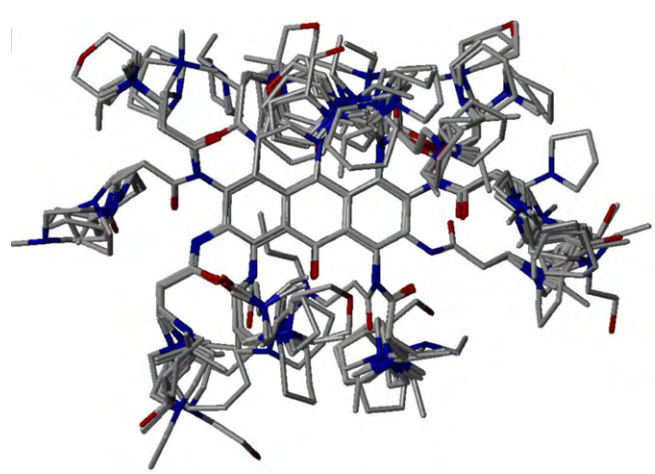


Fig. 2. Alignment of all compounds.

Table 1
Molecular structures and telomerase inhibitory activity of anthraquinone and acridone derivatives.

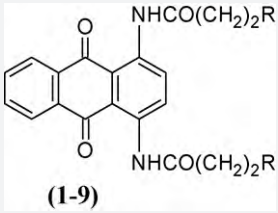
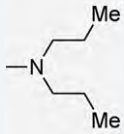
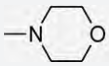
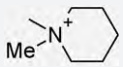
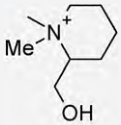
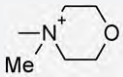
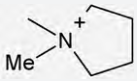
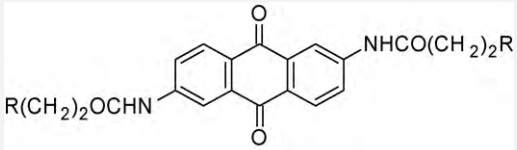
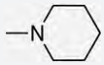
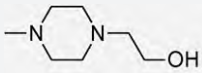
Compound no.	R	Biological activity			Residual	Reference
		Observed		Predicted		
		EC_{50} (μM)	pEC_{50}	pEC_{50}		
						
1		50	4.30	4.220	0.08	[18]
2		33.5	4.47	4.469	0.00	[18]
3		11.1	4.95	5.260	−0.31	[18]
4		9.4	5.03	5.299	−0.27	[18]
5		34.5	4.46	4.336	0.12	[18]
6		1.8	5.74	5.762	−0.02	[19]
7a		5.0	5.30	5.265	0.04	[19]
8		7.0	5.15	5.311	−0.16	[19]
9		3.1	5.51	5.362	0.15	[19]
						
10a		4.5	5.35	5.443	−0.09	[18]
11		16.5	4.78	4.703	0.08	[18]
12a		17.3	4.76	5.033	−0.27	[18]

Table 1 (*Continued*)

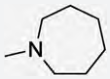
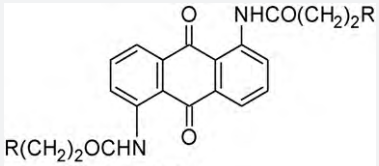
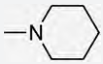
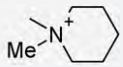
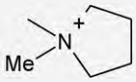
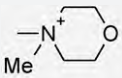
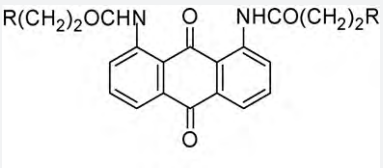
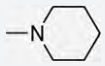
Compound no.	R	Biological activity			Residual	Reference
		Observed		Predicted		
		EC_{50} (μM)	pEC_{50}	pEC_{50}		
13a		1.8	5.74	5.216	0.52	[19]
14		4.1	5.39	5.403	−0.01	[19]
15		3.5	5.46	5.443	0.02	[19]
16		13.0	4.89	5.084	−0.19	[17]
	 (17-24)					
17		2.3	5.64	5.544	0.10	[19]
18a		1.3	5.89	5.444	0.45	[19]
19		2.7	5.57	5.486	0.08	[19]
20		8.6	5.07	5.097	−0.03	[19]
21		8.8	5.06	5.100	−0.04	[19]
22a		14.0	4.85	4.537	0.31	[19]
23		13.2	4.88	4.716	0.16	[19]
24		16.8	4.77	4.682	0.09	[19]
	 (25-32)					
25		3.7	5.43	5.434	0.00	[19]

Table 1 (*Continued*)

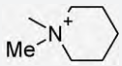
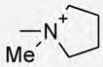
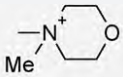
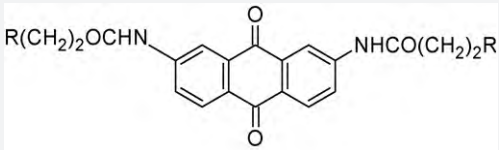
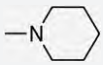
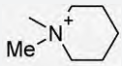
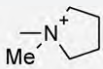
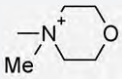
Compound no.	R	Biological activity			Residual	Reference
		Observed		Predicted		
		EC_{50} (μM)	pEC_{50}	pEC_{50}		
26		6.4	5.19	5.368	−0.18	[19]
27		4.2	5.38	5.382	0.00	[19]
28a		7.8	5.11	5.153	−0.04	[19]
29		8.2	5.09	5.152	−0.06	[19]
30		10.0	5.00	4.984	0.02	[19]
31		4.4	5.36	5.174	0.19	[19]
32		7.5	5.12	5.173	−0.05	[19]
		 **(33-40)**				
33		3.1	5.51	5.764	−0.25	[19]
34		2.0	5.70	5.736	−0.04	[19]
35		4.7	5.33	5.154	0.18	[19]
36		4.3	5.37	5.198	0.17	[19]
37a		7.8	5.11	4.986	0.12	[19]
38a		16.0	4.80	4.979	−0.18	[19]
39		16.5	4.78	4.853	−0.07	[19]
40		14.5	4.84	4.812	0.03	[19]

Table 1 (*Continued*)

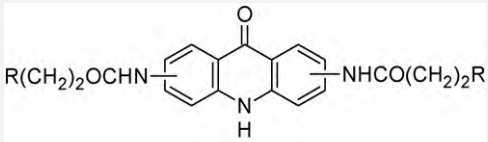
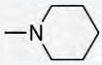
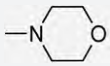
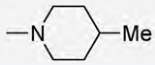
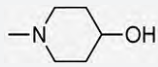
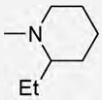
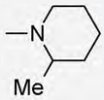
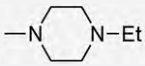
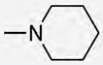
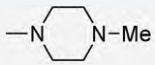
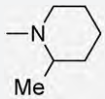
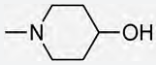
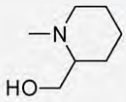
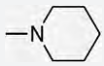
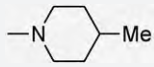
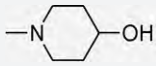
Compound no.	Position	R	Biological activity			Residual	Reference
			Observed		Predicted		
			EC_{50} (μM)	pEC_{50}	pEC_{50}		
		 (41-61)					
41	3,6		8.1	5.09	5.326	−0.24	[16]
42	3,6		5.9	5.23	5.460	−0.23	[16]
43	3,6		4.3	5.37	5.381	−0.01	[16]
44	3,6		5.7	5.24	5.416	−0.18	[16]
45	3,6		49	4.31	4.532	−0.22	[16]
46	3,6		1.7	5.77	5.536	0.23	[16]
47	3,6		1.7	5.77	5.428	0.34	[16]
48	3,6		2.3	5.64	5.557	0.08	[16]
49a	3,6		2.3	5.64	5.507	0.13	[16]
50	3,6		1.9	5.72	5.634	0.09	[16]
51	2,7		5.8	5.24	5.359	−0.12	[16]
52	2,7		1.9	5.72	5.882	−0.16	[16]
53	2,7		0.6	6.22	5.906	0.31	[16]
54a	2,7		1.9	5.72	5.643	0.08	[16]
55	2,7		1.5	5.82	5.934	−0.11	[16]
56	2,7		2.3	5.64	5.716	−0.08	[16]
57	2,6		0.2	6.70	6.638	0.06	[16]

Table 1 (*Continued*)

Compound no.	Position	R	Biological activity			Residual	Reference
			Observed		Predicted		
			EC_{50} (μM)	pEC_{50}	pEC_{50}		
58	2,6		0.7	6.15	6.175	−0.03	[16]
59	2,6		0.4	6.40	6.164	0.24	[16]
60a	2,6		0.2	6.70	6.202	0.5	[16]
61	2,6		0.2	6.70	6.716	−0.02	[16]

[a]Test set compounds.

2.4. CoMSIA study

Comparative molecular similarity indices analysis (CoMSIA) was performed to evaluate steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor properties of molecules by employing the standard options in SYBYL. The steric, electrostatic, hydrophobic, H-bond donor and H-bond acceptor fields were calculated separately using the sp^3 carbon atom probe with a charge of +1 provided in SYBYL 7.0. Similar to CoMFA, a data table has been constructed from similarity indices calculated at the intersections of regularly spaced lattice (2 Å spacing). Similarity indices $A_{F,k}$ between the compounds of interest and a probe atom have been calculated according to equation

$$A^q_{F,k}(j) = -\sum_{i=1}^{n} \omega_{\text{probe},k}\omega_{ik}e^{-\alpha r^2_{iq}}$$

where q is the grid point for molecule j; ω_{ik} is the actual value of the physicochemical property k of atom i; $\omega_{\text{probe},k}$ indicates probe atom with charge +1, radius 1 Å, hydrophobicity +1, H-bond donor and acceptor property +1; α is an attenuation factor; and r_{iq} is the mutual distance between the probe atom at grid point q and atom i of the test molecule. The default value of α is 0.3.

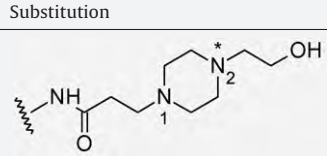
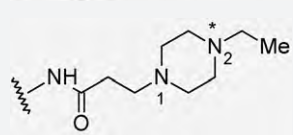
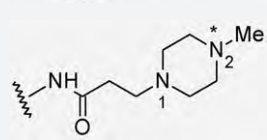
PLS method was used to linearly correlate the CoMSIA fields to the inhibitory activity. Cross-validation [36,37] analysis was performed using the leave one out (LOO) method in which one compound is removed from the data set and its activity is predicted using the model derived from the rest of the dataset. The cross-validated r^2 that resulted in optimum number of components and lowest standard error of prediction were considered for further analysis. To speed up the analysis and reduce noise, a minimum filter value (σ) of 2.0 kcal mol^{-1} was used. Final analysis was performed to calculate conventional r^2 using the optimum number of components.

2.5. Model validation

The primary objective of a typical quantitative structure–activity relationship (QSAR) study is to develop a mathematical model that can be used for the prediction of considered biological activity of new untested compounds from the compounds' molecular structures, because the real utility of QSAR model is in its ability to predict correctly modeled property for new compounds. In the present study many CoMSIA models were developed, most of which were having good primary statistics represented in terms of high values of cross-validation r^2, Fischer statistics (F-test), etc. Although, these values are indicative of the significance of the model, tests like Y-randomization and bootstrapping are required for determining its robustness.

The predictive power of CoMSIA models was further validated by using an external test set (inhibitors marked with 'a' in Table 1).

Table 2
pK_a values of basic nitrogens of piperazine ring system in compounds.

Compound no.	Anthraquinone/acridone	Substitution	pK_a 1	pK_a 2
11	2,6-Disubstituted anthraquinone		3.49	7.75
50	3,6-Disubstituted acridone		3.49	8.61
54	2,7-Disubstituted acridone		3.50	7.59

*Nitrogens used for protonation strategy in the study.

Table 3
Summery of CoMSIA for Conf1.

Parameter	**HAD**	SED	SEA	SHE	EHD	SEDA	SEHD	SEHA	SEHDA
r^2_{cv}	**0.489**	0.403	0.336	0.347	0.425	0.420	0.404	0.329	0.410
ONC	**5**	4	5	4	4	4	4	4	4
SEP	**0.425**	0.433	0.456	0.453	0.425	0.427	0.433	0.459	0.430
r^2_{ncv}	**0.923**	0.785	0.846	0.822	0.832	0.860	0.821	0.855	0.861
SEE	**0.157**	0.260	0.220	0.237	0.230	0.210	0.237	0.213	0.209
F-value	**103.221**	40.25	60.256	50.713	54.517	67.353	50.386	64.994	67.999
$Pr^2 = 0$	**0**	0	0	0	0	0	0	0	0
r^2_{pred}	**0.721**	0.657	0.789	0.633	0.665	0.721	0.662	0.666	0.701
r^2_{bs}	**0.934**	–	–	–		–	–	–	–
SD_{bs}	**0.020**	–	–	–		–	–	–	–
Contribution (fraction) (%)									
S	–	20.4	23.0	20.7	–	15.4	14.4	16.0	11.8
E	–	38.6	42.3	41.9	33.4	27.9	28.5	30.4	22.2
H	**31.7**	–	–	37.4	31.2	–	26.3	27.4	20.6
D	**40.1**	41.0	–	–	35.4	31.4	30.7	–	25.1
A	**28.2**	–	34.8	–	–	25.2	–	26.2	20.3

r^2_{cv}, cross-validated correlation coefficient; ONC, optimum number of components from PLS analysis; SEP, standard error of prediction; r^2_{ncv}, non-crossvalidated correlation coefficient; SEE, standard error of estimate; *F*, Fischer ratio; *S*, steric; *E*, electrostatic; *H*, hydrophobic; *D*, donor; *A*, acceptor; r^2_{pred}, predictive correlation coefficient; r^2_{bs}, bootstrapping correlation; SD_{bs}, bootstrapping standard deviation.

The inhibitors in the test set were given exactly the same pretreatment as the inhibitors in the corresponding training set. The correlation between the experimental and predicted activity for models was calculated as a predictive r^2 value.

3. Results and discussion

3.1. CoMSIA results

A representative set of 61 anthraquinone and acridone derivatives was extracted from the literature [16–19]. Great care was taken while building the structures of compounds to determine physiological ionization state and to check *in situ* protonation sites present on ligands.

Since the molecules contained many nitrogen atoms in their structures, it was speculated that some of them could exist in protonated form in a biological environment. Hence, compounds were built in protonated forms and investigated considering that the basic nitrogen in a compound would get protonated at physiological pH. Nitrogen atoms of the amides were spared from the protonation as these were not supposed to be protonated at physiological pH due to their lower basicity. As there were structures (compounds **11**, **50**, **54**) with two nitrogen atoms in piperazine ring, the decision about the right protonated nitrogen was taken on the basis of pK_a values. Table 2 contains structures and pK_a values of the studied compounds. CoMSIA models were developed using individual single fields and also various fields in combination. Qualities of the models so developed were gauzed by the statistics obtained for each model. Data of CoMSIA models using different fields for Conf1 and Conf2 is presented in Tables 3 and 4, respectively. It was observed that CoMSIA model for Conf1 (Table 3), involving a combination of hydrophobic, acceptor and donor (HAD) fields, played significant role in the prediction of biological activity, and offered a highly predictive 3D-model for substituted anthraquinone and acridone derivatives. For Conf1 an excellent value of 0.721 (test set) for r^2 prediction and 0.489 for r^2 cross-validation with 5 optimum numbers of components were obtained for this model. A good r^2_{ncv} 0.923 (training set) was observed for internal prediction of the model. The contribution of hydrophobic, acceptor and donor fields of this model were 31.7%, 28.2%, and 40.1%, respectively. CoMSIA PLS analysis for Conf1 (Table 3) using hydrophobic, acceptor, and donor descriptors was found to be statistically acceptable. However, CoMSIA models for Conf2 (Table 4) was statistically inferior to models developed for Conf1 (Table 3). It was observed from the statistical quality of two different conformational models that Conf1 afforded the bioactive conformation for the development of predictive CoMSIA model for anthraquinones and acridones as telomerase inhibitors.

Table 4
Summery of CoMSIA for Conf2.

Parameter	HAD	SED	SEA	SHE	EHD	SEDA	SEHD	SEHA	SEHDA
r^2_{cv}	0.302	0.329	0.338	0.256	0.356	0.294	0.273	0.307	0.290
ONC	2	4	2	2	2	2	2	2	2
SEP	0.457	0.459	0.446	0.473	0.440	0.460	0.467	0.456	0.462
r^2_{ncv}	0.663	0.737	0.627	0.607	0.632	0.629	0.620	0.648	0.644
SEE	0.318	0.287	0.335	0.344	0.332	0.334	0.338	0.325	0.327
F-value	45.247	30.776	38.649	35.455	39.517	39.018	37.479	42.400	41.644
$Pr^2 = 0$	0	0	0	0	0	0	0	0	0
r^2_{pred}	0.649	0.645	0.513	0.606	0.642	0.573	0.621	0.588	0.610
r^2_{bs}	–	–	–	–		–	–	–	–
SD_{bs}	–	–	–	–		–	–	–	–
Contribution (fraction) (%)									
S	–	21.2	22.1	22.8	–	15.8	16.2	15.4	12.1
E	–	33.9	37.1	36.3	30.7	26.1	25.6	25.6	19.9
H	34.0	–	–	40.9	34.8	–	29.0	28.8	22.4
D	33.4	44.9	–	–	34.5	29.5	29.3	–	22.9
A	32.6	–	40.8	–	–	28.5	–	30.2	22.7

r^2_{cv}, cross-validated correlation coefficient; ONC, optimum number of components from PLS analysis; SEP, standard error of prediction; r^2_{ncv}, non-crossvalidated correlation coefficient; SEE, standard error of estimate; *F*, Fischer ratio; *S*, steric; *E*, electrostatic; *H*, hydrophobic; *D*, donor; *A*, acceptor; r^2_{pred}, predictive correlation coefficient; r^2_{bs}, bootstrapping correlation; SD_{bs}, bootstrapping standard deviation.

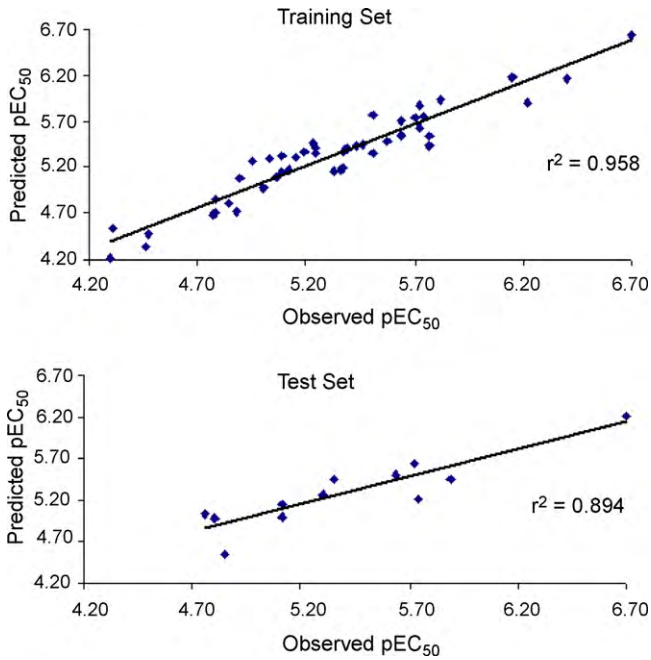


Fig. 3. Graph of observed vs. predicted activities for training and rest set molecules from the best predicitive CoMSIA model (HAD).

Above findings indicate that hydrophobic, hydrogen bond acceptor and donor fields can be more favorable for G-quadruplex stabilization in case of anthraquinone and acridone derivatives, and they could make an important contribution to the binding energy of association. The graphs of observed vs. predicted activities for training and test set molecules from the best CoMSIA model (hydrophobic, hydrogen bond acceptor and hydrogen bond donor descriptors) are shown in Fig. 3.

3.2. Model validation of best CoMSIA model (Conf1)

3.2.1. Internal validation (goodness of fit)

In model development we examined the internal predictive power of the models and their ability to reproduce biological activities of the compounds for the training set. The computed telomerase inhibitory activity from CoMSIA model showed a good correlation (r^2_{ncv} 0.923) with experimental telomerase inhibitory activity.

3.2.2. Fischer statistics (F-test)

Fischer value (F) is the ratio between explained and unexplained variance for a given number of degree of freedom. F-test is variance related statistics that compares two models differing by one or more variable to see if the more complex model is more reliable than the less complex one. The model is supposed to be good if the F-test is above a threshold value, i.e. tabulated value. The larger the value of F, the greater the probability that the QSAR model is. The F-value for the best CoMSIA model was 103.221 [F_{05} (5, 60) = 2.3688 (Tab)] at 95% confidence level, which suggests that the model is statistically significant.

3.2.3. Y-Randomization test

In this test, the dependent variable is randomly shuffled and a new QSAR model is developed using the original independent variable matrix. The process is repeated several times, and the poor values of r^2 and q^2 for the new models ensure the robustness of the original QSAR models [38–41]. In the present case, 10 random trials were run for best CoMSIA model (HAD). None of the random trials could match the original model. The non-cross-validated and cross-validated r^2 for the model from randomized data was found in the range of 0.182–0.238 and −0.581 to −0.143, respectively. The results of randomization test have been presented in Table 5.

Table 5
Randomization test results for CoMSIA model (HAD).

Parameter	Value
No. of random trials	10
No. of trials with r^2_{ncv} and r^2_{cv} greater than non-random trial	0
No. of trials with r^2_{ncv} and r^2_{cv} lesser than non-random trials	10
Range of r^2_{ncv} from randomized data	0.182–0.238
Range of r^2_{cv} from randomized data	−0.581 to −0.143

3.2.4. Bootstrapping

Further, to establish strength and robustness of the model bootstrapping [42,43] (100 runs) was performed. A statistical confidence was determined by the standard deviation of the statistics obtained from bootstrapping. A bootstrapped value (Table 3) (r^2_{bs} of 0.934) and bootstrapped standard deviation (SD_{bs} of 0.020) are obtained for CoMSIA (HAD) model. The results indicate that a reliable model has been successfully constructed.

3.2.5. External validation

LOO cross-validation alone is not a reliable indicator of the prediction capability of QSAR model for the two well-known reasons. LOO tends to reflect more of model robustness than its prediction power and LOO yields increasingly optimistic results with increasing data size (simply because removing one compound from a large data set will generally result in little if any variation of its structural domain) [44]. Moreover, LOO cross-validation does not show the correct symptomatic trend with increasing number of observations [45]. Certainly, the high value of LOO cross-validation appears to be the necessary but not the sufficient condition for the models to have good predictive power. Golbraikh and Tropsha ever emphasized that the actual predictive ability of a QSAR model can only be estimated using an external test set of compounds that were not used for building the model [46].

The value of r^2_{pred} was calculated for the test set and gave the best results for CoMSIA model with hydrophobic, acceptor and donor as descriptors with the values 0.721. Thus, the CoMSIA model (HAD) displays higher predictivity both in regular cross-validation and in the prediction of the test compounds.

3.3. Interpretation of the best CoMSIA model

The contour plot of the CoMSIA hydrophobic, H-bond acceptor, and H-bond donor fields (stdev*coeff) are presented in Fig. 4. Favored and disfavored levels fixed at 80% and 20%, respectively, were used. The most active compound **61** ($^{tel}EC_{50}$ 0.2) from the series is shown inside the fields.

Hydrophobic maps shown in Fig. 4a indicate that the lipophilic favorable yellow region is found surrounding the aromatic planar chromophore and alkyl side chains. This suggests that lipophilicity of this portion of the molecule is an important factor for the activity. The white region at the end of the side chains indicates that hydrophilic substitution in this zone favor the inhibitory activity. Aromatic chromophore of anthraquinone and acridone derivatives fits well in favorable yellow hydrophobic region. This indicates the importance of lipophilicity for telomerase inhibitory activity. Alkylamido side chains of compounds **57** ($^{tel}EC_{50}$ 0.2), **60** ($^{tel}EC_{50}$ 0.2) and **61** ($^{tel}EC_{50}$ 0.2) are perfectly entrenched in favorable yellow region, hence, found to be most potent in the series. Side chains of compound **1** ($^{tel}EC_{50}$ 4.30) were directed away from the favorable hydrophobic region. This could be one of the reasons why 1,4-substituted anthraquinone derivatives are less potent than 2,6-

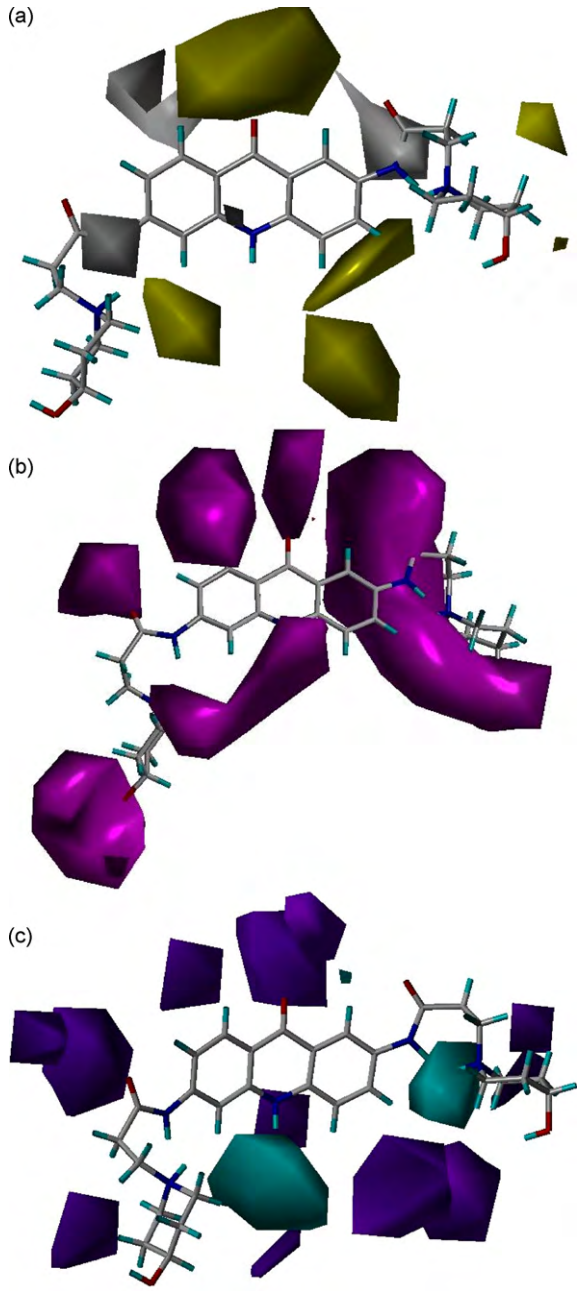


Fig. 4. CoMSIA stdev*coeff contour maps. Compound **61** is shown inside the field. (a) Hydrophobic field, (b) H-bond acceptor field and (c) H-bond donor field. Yellow and white contours indicate regions where hydrophobic groups favored and disfavored the activity, respectively. Magenta contours represent areas where H-bond acceptors are favored. Cyan and purple contours represent areas where H-bond donors are favored and disfavored, respectively.

substituted. In case of compounds **11** ($^{tel}EC_{50}$ 16.5), **50** ($^{tel}EC_{50}$ 1.9) and **54** ($^{tel}EC_{50}$ 1.9), the well defined white region is associated with nitrogen atom in piperazine moiety and indicates that more hydrophilic piperazine moiety is preferable. This result explains why nitrogen belonging to the piperazine has positive contribution to telomerase inhibitory activity in 3D-QSAR CoMSIA analysis.

In principle, H-bond donor and acceptor fields should highlight areas where ligands and putative hydrogen partners in the receptor site could form hydrogen bonds that significantly influence binding affinity. Basically, the acceptor field contains information about where hydrogen bond donating groups should be on the receptor whereas the donating field depicts where hydrogen bond acceptor group should be located on the receptor. CoMSIA H-bond acceptor field (Fig. 4b) shows the regions where H-bond donors in the binding site are predicted to enhance (magenta) the binding of ligands to G-quadruplex.

According to Lewis definition of acidity and basicity, where acids accept and bases donate electron density, the H-bond acceptor is a Lewis base since it donates electron density. It is interesting to observe that the nitrogen atom of the acridone ring system functions as a hydrogen bond acceptor as it carries non-bonding pair of electrons. There is prominent magenta contour surrounding the central nitrogen atom of acridone, which indicates that an H-bond acceptor is favored at this position. In anthraquinons, the nitrogen of acridone is replaced by a ketonic oxygen. This oxygen is also indicated as a favorable region for H-bond acceptor group along with the amide keto oxygens of the side chains.

It may be inferred from the above observations that the central nitrogen atom of acridone ring should have more electron density, so that it can act as H-bond acceptor and the substitutions should be made accordingly at this position without diminishing its relative basicity (i.e., strength) as hydrogen bond acceptor group. Molecular orbital considerations suggest that the acridone chromophoric group is not a fully delocalized system. Hence, central ring acridone nitrogen atom does not contribute to effective π–π staking interaction with adjacent base pairs, as seen in acridine derivatives. However, it has not been made very clear whether its interaction with the G-quadruplex is ionic or not [16]. The Current study reveals that the acridone ring nitrogen remains non-protonated and acts as H-bond acceptor group in the stacking process. Compound **61** is most active as its oxygen atom of carbonyl group at position-9 and central ring nitrogen atom of acridone chromophore fall exactly in the region of favorable H-bond acceptor contours. H-bond donor alkyl side chains of compound **1** ($^{tel}EC_{50}$ 4.30) are oriented away from favorable H-bond donor contour region. This is the reason it has got lowest activity in the series.

The graphical interpretation of the field's contributions of the H-bond donor properties is shown in Fig. 4c. A cyan colored H-bond donor region surrounds the piperidine ring nitrogen atoms showing the favorable regions for H-bond donor fields. A prominent purple disfavorable region was found in the vicinity of oxygen atom of carbonyl group present on C9 position of acridone chromophore. In compound **1** ($^{tel}EC_{50}$ 4.3) NH of amide groups substituted at C1 and C4 position of anthraquinone were perfectly entrenched towards H-bond donor disfavorable region along with oxygen atom of amide protruding towards H-bond donor favorable region hence it showed poor telomerase inhibitory activity. This explains the involvement of amide functionality present in anthraquinone derivatives for stabilization of G-quadruplex. In case of compound **41** ($^{tel}EC_{50}$ 8.1), NH of amide groups in alkyl side chains of acridone ring are oriented away from the favorable H-bond donor region. This could be the reason for its lower activity among other 3,6-disubstituted acridone derivatives.

It is remarkable to note that the nitrogens on the basic side chains acts as H-bond donor after protonation. Compounds with simple ammonium groups (**14**, **17**, **18**, **19**, **25**, **27**, **33**, **34**, and **35**) possess higher biological activity over the corresponding quarternary ammonium compounds (**12**, **20**, **23**, **24**, **28**, **32**, **37**, **38**, and **40**) because the quarternary ammonium group can not act as H-bond donor.

4. Conclusion

3D-QSAR models with statistical significance and high predictive abilities by using CoMSIA analysis were developed. The CoMSIA model generated by a combination of hydrophobic, hydrogen bond acceptor and hydrogen bond donor fields (HAD) showed good cor-

relative and predictive properties. The predictive power of the best CoMSIA model (HAD) has been assessed by the test set molecules. As for as prediction of the activities of the anthraquinones and acridones as telomerase inhibitors are concerned, the Conf1-based alignment model performed the best. Presence of hydrophobic fields in CoMSIA models suggest that telomerase inhibitory activity of anthraquinone and acridone derivatives is influenced primarily by an optimum value of lipophilicity. This study also revealed that apart from hydrophobic and H-bond donor fields, the highly predictive model also included hydrogen bond acceptor field for acridone derivatives.

It is clear that 3D-QSAR information obtained from this kind of analysis could provide important guidelines for the drug design process. As a consequence, the outcome of this study could be used as a guide for further development of selective and more potent telomerase inhibitors. The proposed best model could also be used to predict the activity of newly designed analogs, prior to synthesis.

Acknowledgments

We thank the All India Council for Technical Education (AICTE), New Delhi, India for financial support [File No. 8023/BOR/RID/RPS-148/2007-08]. V.P.Z. is thankful to AICTE, New Delhi, India for the award of a National Doctoral Fellowship [File No. 1-10/RID/NDF-PG/(34)2007-08].

References

[1] E.H. Blackburn, Telomeres: no end in sight, Cell 77 (1997) 621–623.
[2] E.H. Blackburn, J.W. Szostak, The molecular structure of centromers and telomeres, Annu. Rev. Biochem. 53 (1984) 163–194.
[3] E.H. Blackburn, Structure and function of telomeres, Nature 350 (1991) 569–573.
[4] W.C. Hahn, C.M. Counter, A.S. Lundberg, R.L. Beijersbergen, M.W. Brooks, R.A. Weinberg, Creation of human tumour cells with defined genetic elements, Nature 400 (1999) 464–468.
[5] N. Kim, M. Piatyszck, K. Prowse, C. Harley, M. West, P.L. Ho, G.M. Coviello, W.E. Wright, J.W. Shay, Specific association of human telomerase activity with immortal cells and cancer, Science 266 (1994) 2011–2015.
[6] F. Cuenca, O. Greciano, M. Gunaratnam, S. Haider, D. Munnur, R. Nanjunda, W.D. Wilson, S. Neidle, Tri- and tetra-substituted naphthalene diimides as potent G-quadruplex ligands, Bioorg. Med. Chem. Lett. 18 (2008) 1668–1673.
[7] K. Masutomi, E.Y. Yu, S. Khurts, I. Ben-Porath, J.L. Currier, G.B. Metz, M.W. Brooks, S. Kaneko, S. Murakami, J.A. DeCaprio, R.A. Weinberg, S.A. Stewart, W.C. Hahn, Telomerase maintains telomere structure in normal human cells, Cell 114 (2003) 241–253.
[8] J.L. Mergny, J.F. Riou, P. Mailliet, M.P. Teulade-Fichou, E. Gilson, Natural and pharmacological regulation of telomerase, Nucleic Acids Res. 30 (2002) 839–865.
[9] S.M. Gowan, J.R. Harrison, L. Patterson, M. Valneti, M.A. Read, S. Neidle, L.R. Kelland, A G-quadruplex interactive potent small-molecule inhibitor of telomerase exhibiting in vitro and in vivo antitumor activity, Mol. Pharmacol. 61 (2002) 1154–1162.
[10] T.R. Cech, Life at the end of the chromosome: telomeres and telomerase, Angew. Chem. Int. Ed. 39 (2000) 34–43.
[11] D. Yang, K. Okamoto, Structural insights into G-quadruplexes: towards new anticancer drugs, Future Med. Chem. 2 (2010) 619–646.
[12] R. Harrison, S. Gowan, L. Kelland, S. Neidle, Human telomerase inhibition by substituted acridine derivatives, Bioorg. Med. Chem. Lett. 9 (1999) 2463–2468.
[13] M.A. Read, A. Wood, R. Harrison, S. Gowan, L. Kelland, H. Dosanjh, S. Neidle, Molecular modeling studies on G-quadruplex complexes of telomerase inhibitors: structure activity relationships, J. Med. Chem. 42 (1999) 4538–4546.
[14] N. Campbell, G. Parkinson, A. Reszka, S. Neidle, Structural basis of DNA quadruplex recognition by an acridine drug, J. Am. Chem. Soc. 130 (2008) 6722–6724.
[15] N. Campbell, M. Patel, A. Tofa, R. Ghosh, G. Parkinson, S. Neidle, Selectivity in ligand recognition of G-quadruplex loops, Biochemistry 48 (2009) 1675–1680.
[16] J.R. Harrison, A.P. Reszka, S.M. Haider, B. Romagnoli, J. Morrell, M.A. Read, S.M. Gowan, C.M. Incles, L.R. Kelland, S. Neidle, Evaluation of by disubstituted acridone derivatives as telomerase inhibitors: the importance of G-quadruplex binding, Bioorg. Med. Chem. Lett. 14 (2004) 5845–5849.
[17] M.A. Read, A.A. Wood, J.R. Harrison, S.M. Gowan, L.R. Kelland, H.S. Dosanjh, S. Niedle, Molecular modeling studies on G-quadruplex complexes of telomerase inhibitors: structure–activity relationship, J. Med. Chem. 42 (1999) 4538–4546.
[18] P.J. Perry, S.M. Gowan, A.P. Reszka, P. Polucci, T.C. Jenkins, L.R. Kelland, S. Neidle, 1,4- and 2,6-disubstituted amidoanthracene-9,10-dione derivatives as inhibitors of human telomerase, J. Med. Chem. 41 (1998) 3253–3260.
[19] P.J. Perry, A.P. Reszka, A. Wood,.A. Read, M.A. Gowan, S.M. Dosanjh, H.S. Trent, J.O. Jenkins, T.C. Kelland, L.R. Neidle, S. Human telomerase inhibition by regioisomeric disubstituted amidoanthracene-9,10-diones, J. Med. Chem. 41 (1998) 4873–4884.
[20] G. Zagotto, C. Sissi, S. Moro, D. Ben, G. Parkinson, K. Fox, S. Neidle, M. Palumbo, Amide bond direction modulates G-quadruplex recognition and telomerase inhibition by 2,6 and 2,7 bis-substituted anthracenedione derivatives, Bioorg. Med. Chem. 16 (2008) 354–361.
[21] F. Cuenca, M.J.B. Moore, K. Johnson, B. Guyen, A. Cian, Design, synthesis and evaluation of 4,5-disubstituted acridone ligands with high G-quadruplex affinity and selectivity, together with low toxicity to normal cell, Bioorg. Med. Chem. Lett. 19 (2009) 5109–5113.
[22] V.P. Zambre, P.R. Murumkar, R. Giridhar, M.R. Yadav, Structural investigations of acridine derivatives by CoMFA and CoMSIA reveal novel insight into their structures toward DNA G-quadruplex mediated telomerase inhibition and offer a highly predictive 3D model for substituted acridines, J. Chem. Inf. Model. 49 (2009) 1298–1311.
[23] P.R. Murumkar, S. Dasgupta, V.P. Zambre, R. Giridhar, M.R. Yadav, Development of predictive 3D- QSAR CoMFA and CoMSIA models for β-aminohydroxamic acid-derived TACE inhibitors, Chem. Biol. Drug Des. 73 (2009) 97–107.
[24] P.R. Murumkar, R. Giridhar, M.R. Yadav, 3D-quantitative structure–activity relationship studies on benzothiadiazepine hydroxamates as inhibitors of tumor necrosis factor-alpha converting enzyme, Chem. Biol. Drug Des. 71 (2008) 363–373.
[25] D.S. Puntambekar, R. Giridhar, M.R. Yadav, 3D-QSAR studies of farnesyltransferase inhibitors: a comparative molecular field analysis approach, Bioorg. Med. Chem. Lett. 16 (2006) 1821–1827.
[26] D.S. Puntambekar, R. Giridhar, M.R. Yadav, Understanding the anti-tumor activity of novel tricyclicpiperazinyl derivatives as farnesyltransferase inhibitors using CoMFA and CoMSIA, Eur. J. Med. Chem. 41 (2006) 1279–1292.
[27] D.S. Puntambekar, R. Giridhar, M.R. Yadav, Insight into the structural requirements of farnesyltransferase inhibitors as potential anti-tumor agents based on 3D-QSAR CoMFA and CoMSIA models, Eur. J. Med. Chem. 43 (2008) 142–154.
[28] D.S. Puntambekar, R. Giridhar, M.R. Yadav, 3D-QSAR CoMFA/CoMSIA studies on 5-aryl 2-,2-dialkyl–4 phenyl–3 (2*H*)–furanone derivatives, as selective COX-2 inhibitors, Acta Pharm. 56 (2006) 157–174.
[29] G. Klebe, U. Abraham, T. Meitzner, Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity, J. Med. Chem. 37 (1994) 4130–4146.
[30] R. Cramer, D. Patterson, J. Bunce, Comparative molecular field analysis (CoMFA) Effect of shape on binding of steroids to carrier proteins, J. Am. Chem. Soc. 110 (1988) 5959–5967.
[31] SYBYL 7.0, Tripos Inc., 1699 South Hanley Road, St. Louis, Missouri 63144, USA.
[32] S.M. Haider, G.N. Parkinson, S. Neidle, Structure of a G-quadruplex-ligand complex, J. Mol. Biol. 326 (2003) 117–125.
[33] J. Stewart, MOPAC: a semi impirical molecular orbital program, J. Comput. Aided Mol. Des. 4 (1990) 1–105.
[34] M. Dewar, E. Zoebisch, E. Healy, J. Stewart, Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model, J. Am. Chem. Soc. 107 (1985) 3902–3909.
[35] CompuDrug, Latest Upgrades, (accessed February 10, 2010), http://www.compudrug.com.
[36] T.A. Halgren, Maximally diagonal force constants in dependent angle-bending co-ordinates. II. Implications for the design of empirical force fields, J. Am. Chem. Soc. 112 (1990) 4710–4723.
[37] B.L. Podlogar, D.M. Fergusson, QSAR and CoMFA: a perspective on the practical application to drug discovery, Drug Des. Discov. 17 (2000) 4–12.
[38] C.D. Selassie, S. Kapur, R.P. Verma, M. Rosario, Cellular opoptosis and cytotoxicity of phenolic compounds: a quantitative structure–activity relationship study, J. Med. Chem. 48 (2005) 7234–7242.
[39] A. Tropsha, P. Gramatica, V.K. Gombar, The importance of being earnest: validation id the absolute essential for successful application and interpretation of QSAR model, QSAR Comb. Sci. 22 (2003) 69–77.
[40] S. Wold, L. Eriksson, in: H. van de waterbeemd (Ed.), Chemometric Methods in Molecular Design, Wiley-VCH, Weinheim, 1995, pp. 309–318.
[41] M.F. Melzig, G.D. Tran, K. Henke, C.D. Selassie, R.P. Verma, Inhibition of neutrophil elastase and thrombin activity by caffeic acid esters, Pharmazie 60 (2005) 869–873.
[42] L. Stahle, S. Wold, Multivariate data analysis and experimental design in biomedical research, Prog. Med. Chem. 25 (1988) 291–338.
[43] R.D. Cramer, J.D. Bunce, D.E. Patterson, I.E. Frank, Crossvalidation, bootstrapping, and partial least squares compared with multiple regression in conventional QSAR studies, Quant. Struct. Act. Relat. 7 (1988) 18–25.
[44] M. Baroni, G. Costantino, G. Crucian, D. Riganelli, R. Valigi, S. Clementi, Generating optimal linear PLS estimations (GOLPE): An advanced chemometric tool for handling 3D QSAR problems, Quant. Struct. Act. Relat. 12 (1993) 9–20.
[45] J. Shao, Linear model selection by cross-validation, J. Am. Stat. Assoc. 88 (1993) 486–494.
[46] A. Golbraikh, A. Tropsha, Beware of q^2!, J. Mol. Graph. Model. 20 (2002) 269–276.