

# 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

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

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

Received January 29, 2009

Stabilization of G-quadruplex structures formed from telomeric DNA, by means of G-quadruplex selective ligands, is a means of inhibiting the telomerase enzyme. This makes G-quadruplex an emerging target for cancer therapy. The objective of the current 3D QSAR study is to uncover structural requirements for acridine derivatives, which would eventually assist and complement the rational drug-design attempts. Various protonation strategies were investigated to check *in situ* protonation sites present on ligands when they bind to G-quadruplex, and predictive 3D-QSAR CoMFA and CoMSIA models have been developed. Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) studies were carried out on substituted acridines as telomerase inhibitors. Molecular models with good predictive power were derived using steric, electrostatic, hydrophobic, and H-bond donor fields of the compounds. The CoMSIA coefficient contour plots identified several key features explaining the wide range in activities. The present study not only offers a highly significant predictive CoMSIA model for trisubstituted acridine derivatives as telomerase inhibitors but also throws more light on the molecular structure of these compounds at physiological pH.

## INTRODUCTION

Telomeres are repetitive DNA sequences at the ends of linear chromosomes that protect the chromosomes from recombination, end fusion, and nucleus degradation.<sup>1,2</sup> In human cells, the telomeric DNA is typically composed of 5–15 kb of double stranded pairs of tandem repeats of the guanine-rich TTAGGG sequence with a single stranded 3'-end overhang which is necessary to ensure complete chromosomal DNA replication. With each cell division telomere shortens by 50–200 bp because synthesis of the lagging strand of DNA is unable to replicate the 3'-end overhang. When the telomere shortens to a critical length, normal cells stop growing and enter a state of senescence where end-to-end fusion and chromosomal instability leads to cell death. A cell can escape from this normal cycle and become immortal by stabilizing (capping) the length of its telomere.<sup>3–5</sup> This happens almost always under activation of the enzyme telomerase.<sup>6</sup>

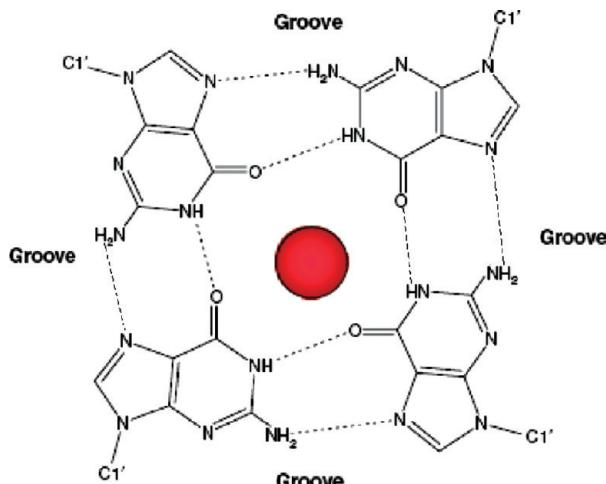
Telomerase, a ribonucleoprotein complex, is active in 80–90% of all human tumors but not in normal somatic cells.<sup>6</sup> Consequently, telomerase has become a high profile target for the development of novel anticancer agents.<sup>7–10</sup> There are two major components of human telomerase, the RNA protein hTR and the reverse transcriptase hTERT. The catalytic subunit of this enzyme, hTERT, adds repeats of TTAGGG to the ends of human chromosomes using the template on the RNA domain of hTR. One approach to inhibit the enzyme activity is to target the telomerase RNA

template with oligonucleotides that inhibit recognition of the 3'-end nucleotides of the telomere, the initial step in telomerase processing. One such oligonucleotide inhibitor to achieve this target is GRN163L, which is currently in clinical trials.<sup>11</sup>

In vertebrates, telomeres consist of tandem repeats of guanine (G) rich DNA sequences of 5'-TTAGGG-3', where the 3'-end extends to form a single-stranded overhang of 150–250 nt.<sup>12–14</sup> This overhang can form a higher order DNA G-quadruplex structure under appropriate conditions, although in cellular conditions binding of the single stranded binding protein hPOT1 normally inhibits quadruplex formation.<sup>15</sup> Formation of intermolecular quadruplex structures (guanine in reverse Hoogsteen base pair, Figure 1) halts the synthesis of further telomeric DNA repeats since the 3'-end can no longer be recognized by the hTR template of the telomerase complex.<sup>16,17</sup> Quadruplex formation can be induced by appropriate quadruplex binding ligands, which thus inhibit telomerase function.<sup>18</sup>

Acridines and anthraquinones with minimal substitutions have been reported previously<sup>19,20</sup> to stabilize G-quadruplex. Various other ligands that stabilize the G-quadruplex structures include cationic porphyrins,<sup>21</sup> perylenes,<sup>22</sup> ethidium derivatives,<sup>23</sup> fluoroquinoanthroxazines,<sup>24</sup> indoquinolines,<sup>25</sup> dibenzophenanthrolines,<sup>26</sup> bisquinacridines,<sup>27</sup> pentacyclic acridiniums,<sup>28,29</sup> telomestatin,<sup>30</sup> and the recently discovered bisquinolinium derivatives.<sup>31,32</sup> Structure-based drug design using human G-quadruplex telomeric repeat d [AG<sub>3</sub>(TTAG<sub>3</sub>)<sub>3</sub>] (Protein Data Bank 143D) is helpful in accelerating the inhibitor optimization process.<sup>33</sup> It was found that the improvement in telomerase inhibitory activity shown by

\* Corresponding author phone: +91 265 2434187; fax: +91 265 2018927; e-mail mryadav11@yahoo.com.



**Figure 1.** Structure of G-quadruplex, showing Hoogsteen base-pairing with a centrally placed metal ion.

the trisubstituted acridines compared well with disubstituted ones.<sup>33</sup> Based on this finding novel selective 3,6,9-trisubstituted acridines<sup>34–37</sup> have been reported as G-quadruplex-mediated telomerase inhibitors.

Although some of the 3,6,9-trisubstituted acridines<sup>34–37</sup> have been reported to possess potent telomerase inhibitory activity, still one has to grope in the dark regarding the molecular requirements for selectivity and potency in these derivatives. There exists a need for the development of newer more selective and potent telomerase inhibitors. A rational basis for designing of G-quadruplex-mediated telomerase inhibitors is the need of the hour.

Recent studies in our laboratory have focused on refining the structures using 3D-QSAR CoMFA and CoMSIA approaches. A successful 3D-QSAR model not only helps in better understanding of the structure–activity relationships of any class of compounds but also provides researchers an insight at the molecular level about lead compounds for further developments. CoMFA<sup>38</sup> and CoMSIA<sup>39</sup> are well accepted and widely used tools for performing 3D-QSAR analysis of any series of compounds. These techniques have been successfully applied in our laboratory for development of COX-II inhibitors,<sup>40</sup> farnesyltransferase inhibitors,<sup>41–43</sup> and TACE inhibitors.<sup>44,45</sup> CoMFA analysis involves the alignment of molecules in a structurally and pharmacologically reasonable manner on the basis of assumption that each compound exhibits biological activity via a common macromolecular target binding site. They can give critical information of the interaction between the ligand and the putative receptors. In these methods it is possible to predict the biological activity of molecules and represent the relationship between molecular properties (Steric, Electrostatic, Hydrophobic, H-bond donors, H-bond acceptors) and biological activity in the form of contour maps.

In the current study it was planned to identify the structural requirements in trisubstituted acridines for their G-quadruplex stabilizing telomerase inhibiting activity using 3D-QSAR techniques. CoMFA and CoMSIA, being the most widely used ones, were used as 3D-QSAR techniques for the chosen series of compounds.

Although the aim of the study was simply to find structural requirements for selective binding to telomerase of trisubstituted acridines and develop a 3D-model for the

purpose of optimization of the leads, certain interesting observations were made during the study concerning the structures of these molecules and their interactions with the G-quadruplex of the enzyme telomerase. All these findings are reported in this work, which is the first of its kind in terms of development of a 3D-model for telomerase inhibiting activity of trisubstituted acridines and throws fresh light on the molecular structures of these compounds in the physiological environment.

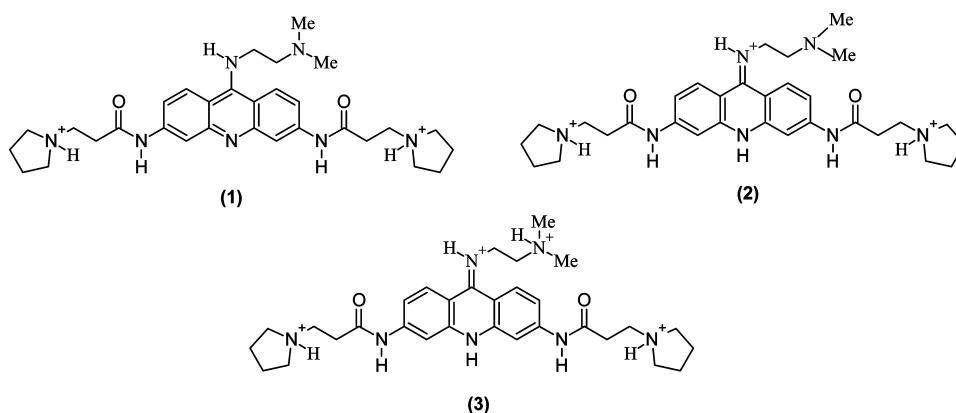
## MATERIALS AND METHODS

**Data Set.** In the present work, a total of 79 compounds with their telomerase inhibitory activity from the reported work<sup>19,20,29,34–37</sup> were selected. The reported compounds showed wide variations in their structures and potency profile. Considering a high deviation in the biological activity and structural variations among the compounds of the series it was considered as an ideal series for performing QSAR analysis. Biological data with negative logarithm of minimum inhibitory concentration (MIC) expressed in mol/liter was used as a dependent variable in the 3D-QSAR study, thus correlating the data linearly to the free energy change.

**Selection of Training and Test Sets.** In view of the finding that  $q^2$  appears to be a necessary but not a sufficient condition for a model to have high predictive power,<sup>46</sup> an emphasis has been given in the present study for validation of the developed model using an external test set. The whole set of 79 compounds was divided into training set (61) and test set (18) compounds. In the training set, most potent, moderately active and low 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 of the training set was chosen for the test set.

**Computational Details.** A Silicon Graphics Fuel workstation with IRIX 6.5 operating system running SYBYL 7.0<sup>47</sup> was used for three-dimensional structure building and molecular modeling studies. Initial optimization of the structures was carried out using TRIPPOS force field with Gasteiger-Huckel charges, and repeated minimization was performed using steepest-descent and conjugate gradient methods until the root-mean-square (rms) deviation of 0.001 kcal/mol was achieved. Conformational energies were computed with electrostatic terms; the lowest energy structures finally minimized were used in superimposition. The partial atomic charges required for the electrostatic interactions were computed by the semiempirical molecular orbital method using Molecular Orbital PAckage (MOPAC)<sup>48</sup> with Austin Model 1 (AM1) Hamiltonian.<sup>49</sup>

**Docking Protocol.** Docking studies were performed with different protonation strategies (Figure 2) of the most active compound **12** in order to understand the cause of the significant difference in molecular structure with different protonation sites on binding interactions with G-quadruplex. The decision of right protonation strategy was taken on the basis of docking results. Docking studies were performed using the molecular dynamics simulated annealing algorithm-based docking program CDOCKER to generate ligand poses wherein a set of ligand conformations are generated using high temperature molecular dynamics. Random orientations of the conformations are produced by translating the center of the ligand to a

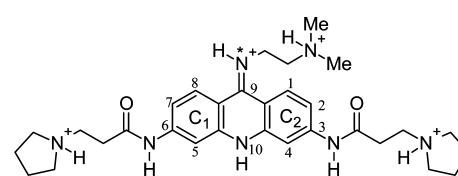


**Figure 2.** Different protonation strategies.

specified location within the receptor active site and performing a series of random rotations. A softened energy is calculated, and the orientation is kept if the energy is less than a specified threshold. This process continues until either the desired number of low-energy orientations is found or the maximum numbers of bad orientations have been tried. Each orientation is subjected to simulated annealing molecular dynamics. The target is heated up to a temperature of 700 K and then cooled to 300 K. A final minimization of the ligand in the rigid receptor using nonsoftened potential is performed. For each final pose, the CHARMM energy (interaction energy plus ligand strain) and the interaction energy alone are calculated. The poses are sorted by CHARMM energy, and the top scoring poses are retained where lower values indicate more favorable binding.<sup>50</sup> The X-ray crystal structure of the quadruplex-drug DNA complex (PDB 1L1H)<sup>51</sup> contained a cocrystallized ligand, **71**, in the active site. Potassium ions in the central channel between the planes of each G-quartet were preserved, and all water molecules were deleted. The CHARMM force field was assigned to a quadruplex, and the hydrogen atoms were automatically added. The structure of the drug–DNA complex after knocking out the drug **71** was used for docking studies. Docking of each protonation strategy of compound **12** into the active site was performed.

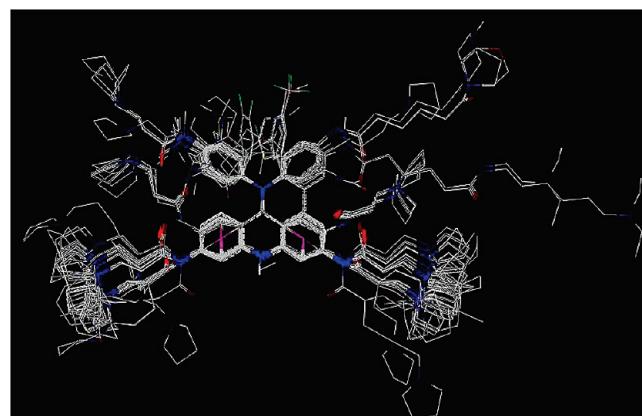
**Selection of Template and Molecular Alignment.** In the development of 3D-QSAR models, the choice of the template conformation is the most important factor to provide a reliable pharmacophore model. This renders the spatial alignment of molecules under study as one of the most sensitive and determining factors in obtaining a robust and meaningful model. The energy minimized docked conformation<sup>50</sup> of the most active compound **12** from the series was used as a template. The molecular alignment was carried out using the centroid and atom-based rms fit. These options use alignment of structures through defining the centroid and pairwise superimposition of these centroids and atoms which render all structures in the database in the same frame of reference as the template compound. The centroids and atoms considered for the alignment are marked with C1, C2, and an asterisk (\*), respectively, in Figure 3. The superimposition of all the compounds on template (compound **12**) is shown in Figure 4.

**CoMFA Studies.** To derive the CoMFA descriptor fields, a 3D cubic lattice with grid spacing of 2 Å in x, y, and z



**C<sub>1</sub>, C<sub>2</sub>** Centroids and \*Atom considered for alignment

**Figure 3.** Template used for the alignment.



**Figure 4.** Superimposition of compounds on template.

directions was created to encompass the aligned molecules. CoMFA descriptors were calculated using an  $sp^3$  carbon probe atom with a van der Walls 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. The steric and electrostatic energy values were truncated at a default value of 30 kcal/mol. The CoMFA steric and electrostatic fields generated were scaled by the CoMFA standard option available in SYBYL.

A partial least-squares (PLS) approach, an extension of multiple regression analysis, was used to derive 3D-QSAR, in which the CoMFA descriptors were used as independent variables and  $pEC_{50}$  values as dependent variables. Cross-validation analysis was performed for selecting the model which is most likely to have predictive values. The intensity of the cross-validation process is controlled by selecting the number of 10 groups. After the optimum number of components (ONC) was determined, a non-cross-validated analysis was performed. The  $r^2_{cv}$ , PRESS (the root mean predictive error sum of squares),  $r^2_{ncv}$ , F-value, and standard error of estimate

**Table 1.** Telomerase Inhibition Data for Substituted Acridine Compounds<sup>a</sup>

Sr. No	R	<sup>tel</sup> EC <sub>50</sub> (μM)	Actual pEC <sub>50</sub>	Predicted pEC <sub>50</sub>		
1a	-C <sub>6</sub> H4NMe <sub>2</sub> (p)	0.115	6.94	6.425		
2	-C <sub>6</sub> H4NH <sub>2</sub> (p)	0.074	7.13	7.233		
3	-CH <sub>2</sub> CH <sub>2</sub> CH <sub>2</sub> NHMe <sub>2</sub>	0.06	7.22	7.486		
4	-CH <sub>2</sub> CH <sub>2</sub> C <sub>5</sub> H <sub>10</sub> N (c)	0.05	7.30	7.250		
5	-C <sub>6</sub> H <sub>4</sub> NH <sub>2</sub> (m)	0.06	7.22	7.185		
6	-C <sub>6</sub> H <sub>4</sub> NH <sub>2</sub> (o)	0.02	7.70	7.309		
7	-C <sub>6</sub> H4NMe <sub>2</sub> (m)	0.1	7.00	7.044		
8	-C <sub>6</sub> H <sub>11</sub> (c)	0.09	7.05	7.007		
9	-CH <sub>2</sub> CH <sub>2</sub> OMe	0.14	6.85	7.058		
10	-C <sub>6</sub> H <sub>13</sub> (c)	0.21	6.68	6.939		
11	-C <sub>6</sub> H <sub>4</sub> COCH <sub>3</sub> (p)	0.04	7.40	7.112		
12	-CH <sub>2</sub> CH <sub>2</sub> NHMe <sub>2</sub>	0.018	7.74	7.515		
13a		0.018	7.74	7.125		
14	-CH <sub>2</sub> C <sub>5</sub> H <sub>4</sub> N (m) (c)	0.066	7.18	7.295		
15	-C <sub>6</sub> H <sub>4</sub> NHCOCH <sub>3</sub> (m)	0.1	7.00	7.249		
16	-C <sub>3</sub> H <sub>5</sub> (c)	0.05	7.30	7.055		
17	-C <sub>6</sub> H <sub>4</sub> F (p)	0.07	7.15	6.989		
18a	-C <sub>6</sub> H <sub>4</sub> SMe (o)	0.15	6.82	6.996		
19	-C <sub>6</sub> H <sub>4</sub> SMe (m)	0.1	7.00	6.996		
Sr. No	R	<sup>tel</sup> EC <sub>50</sub> (μM)	Actual pEC <sub>50</sub>	Predicted pEC <sub>50</sub>		
20	-C <sub>6</sub> H <sub>4</sub> NH <sub>2</sub> (p)	0.08	7.10	7.127		
21	-C <sub>6</sub> H <sub>4</sub> NMe <sub>2</sub> (p)	0.17	6.77	6.877		
22	-CH <sub>2</sub> CH <sub>2</sub> NHMe <sub>2</sub>	0.27	6.57	6.037		
23a	-C <sub>6</sub> H <sub>4</sub> NH <sub>2</sub> (m)	0.21	6.68	6.813		
24	-C <sub>6</sub> H <sub>4</sub> NH <sub>2</sub> (o)	0.11	6.96	7.169		
25*	-C <sub>6</sub> H <sub>5</sub>	1.33	5.88	-		
26a	-CH <sub>2</sub> CH <sub>2</sub> CH <sub>2</sub> NHMe <sub>2</sub>	0.08	7.10	7.045		
27	-C <sub>6</sub> H <sub>11</sub> (c)	0.21	6.68	6.713		
Sr. No	R	<sup>tel</sup> EC <sub>50</sub> (μM)	Actual pEC <sub>50</sub>	Predicted pEC <sub>50</sub>		
28	-C <sub>6</sub> H <sub>4</sub> OMe (p)	0.46	6.34	6.091		
29	-C <sub>6</sub> H <sub>4</sub> NH <sub>2</sub> (o)	0.17	6.77	6.237		
30	-C <sub>6</sub> H <sub>4</sub> NH <sub>2</sub> (m)	1.09	5.96	6.182		
31a	-C <sub>6</sub> H <sub>4</sub> NMe <sub>2</sub> (m)	0.6	6.22	6.059		
32	-C <sub>6</sub> H <sub>4</sub> NH <sub>2</sub> (p)	0.2	6.70	6.274		
33	-C <sub>6</sub> H <sub>4</sub> NMe <sub>2</sub> (p)	0.5	6.30	6.070		
34	-C <sub>6</sub> H <sub>5</sub>	1.29	5.89	6.096		
35	-C <sub>6</sub> H <sub>4</sub> OMe (m)	2.73	5.56	6.047		
36	-C <sub>6</sub> H <sub>4</sub> OH (o)	1.03	5.99	6.105		
37	-CH <sub>2</sub> CH <sub>2</sub> NHMe <sub>2</sub>	0.57	6.24	6.213		
Sr. No	R	n	<sup>tel</sup> EC <sub>50</sub> (μM)	Actual pEC <sub>50</sub>	Predicted pEC <sub>50</sub>	
38	-C <sub>6</sub> H <sub>4</sub> N(Me) <sub>2</sub> (p)	3	0.099	7.00	6.635	
39	-C <sub>6</sub> H <sub>4</sub> N(Me) <sub>2</sub> (p)	4	1.93	5.71	5.805	
40	-C <sub>6</sub> H <sub>4</sub> N(Me) <sub>2</sub> (p)	5	6.91	5.16	5.302	
41		3	0.326	6.49	6.796	
42		4	0.255	6.59	6.844	
43a		5	0.146	6.84	6.815	
Sr. No	N	<sup>tel</sup> EC <sub>50</sub> (μM)	Actual pEC <sub>50</sub>	Predicted pEC <sub>50</sub>		
44	1	0.167	6.78	6.497		
45	2	0.067	7.17	6.485		
46	3	0.117	6.93	6.547		
Sr. No	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	<sup>tel</sup> EC <sub>50</sub> (μM)	Actual pEC <sub>50</sub>	Predicted pEC <sub>50</sub>
47*	F	F	H	0.03	7.52	-
48a	OCH <sub>3</sub>	H	OCH <sub>3</sub>	0.35	6.46	6.289
49	CH <sub>3</sub>	H	H	1.00	6.00	6.174
50a	CF <sub>3</sub>	H	F	0.24	6.62	6.142
51	CF <sub>3</sub>	H	CF <sub>3</sub>	0.86	6.07	6.092
Sr. No	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	<sup>tel</sup> EC <sub>50</sub> (μM)	Actual pEC <sub>50</sub>	Predicted pEC <sub>50</sub>
52	F	F	H	0.88	6.06	6.462
53a	OCH <sub>3</sub>	H	OCH <sub>3</sub>	0.44	6.36	6.301
54	CH <sub>3</sub>	H	H	0.23	6.64	6.186
Sr. No	Structure	<sup>tel</sup> EC <sub>50</sub> (μM)	Actual pEC <sub>50</sub>	Predicted pEC <sub>50</sub>		
55		0.36	6.44	6.951		
56		0.39	6.41	6.202		
Sr. No	X	<sup>tel</sup> EC <sub>50</sub> (μM)	Actual pEC <sub>50</sub>	Predicted pEC <sub>50</sub>		
57		0.318	6.50	6.630		
58a		0.267	6.57	6.686		
59		0.165	6.78	6.672		
60		0.098	7.01	6.912		
61a		0.080	7.10	6.757		

**Table 1.** Continued

Sr. No.	R	$^{tel}EC_{50}$ ( $\mu M$ )	Actual pEC <sub>50</sub>	Predicted pEC <sub>50</sub>
62		5.8	5.24	5.506
63		8.2	5.09	5.482
64		2.7	5.57	5.409
65		2.6	5.58	5.385
66		1.35	5.87	5.472
67a		4.4	5.36	5.420
68		5.4	5.27	5.337
69a		4.1	5.39	5.336

<sup>a</sup> c = cyclic, o = ortho, m = meta, p = para, \* = outlier.

Sr. No.	R	$^{tel}EC_{50}$ ( $\mu M$ )	Actual pEC <sub>50</sub>	Predicted pEC <sub>50</sub>
70		8.0	5.10	5.364
71		5.2	5.28	5.379
72		2.8	5.55	5.424
73a		3.1	5.51	5.509

c = cyclic, o = ortho, m = meta, p = para, \*outlier

(SEE) values were computed according to the definition in the SYBYL. The cross-validated coefficient was calculated using eq 1

$$r^2_{cv} = 1 - \frac{\sum (Y_{Predicted} - Y_{Observed})^2}{\sum (Y_{Observed} - Y_{Mean})^2} \quad (1)$$

where  $Y_{Predicted}$ ,  $Y_{Observed}$ , and  $Y_{Mean}$  are predicted, observed, and mean values of the target property ( $pEC_{50}$ ), respectively.  $\sum (Y_{Predicted} - Y_{Observed})^2$  are the predictive residual sum of square (PRESS). To further assess the robustness and statistical confidence of the derived model bootstrapping analysis was performed.

The predictive correlation coefficient ( $r^2_{pred}$ ), based on the test set molecules, was calculated using eq 2

$$r^2_{pred} = (SD-PRESS)/SD \quad (2)$$

where SD is the sum of the squared deviations between the biological activities of the test set and the mean activity of the training set molecules, and PRESS is the sum of the squared deviations between predicted and actual activities for every compound in the test set. The activity of the test set was predicted by the CoMFA model using the predict command. CoMFA coefficient maps were generated by interpolation of the pairwise products between the PLS coefficients and the standard deviations of the corresponding CoMFA descriptor values.

**CoMSIA Studies.** Comparative molecular similarity index 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 eq 3

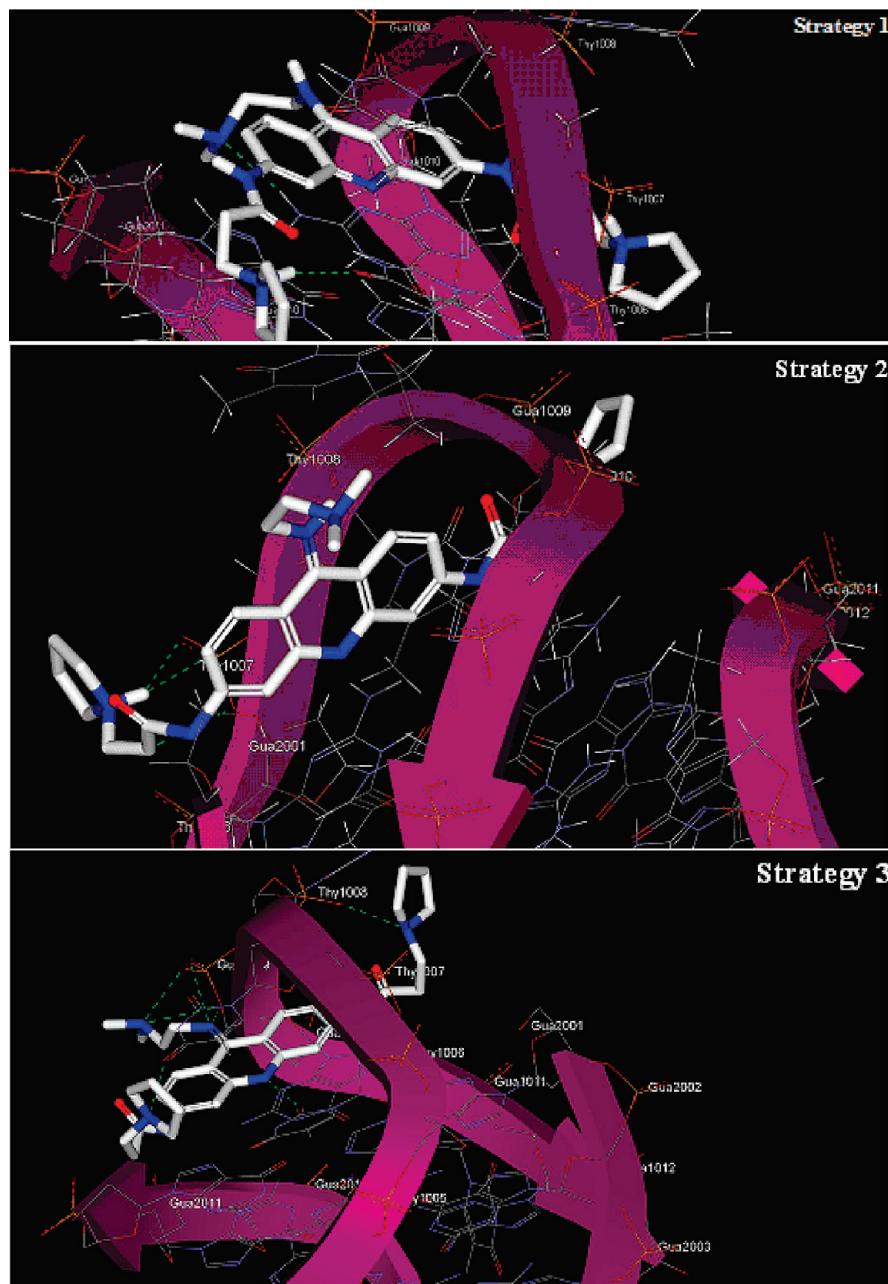
$$A_{F,K}^q(j) = - \sum_{i=1}^n \omega_{probe,k} \omega_{ik} e^{-\alpha r_{iq}^2} \quad (3)$$

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

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

## RESULTS AND DISCUSSION

CoMFA analysis initially was performed on the reported compounds of the training set as such. The model obtained thus, although observed to be on the edge of statistically acceptable limits, offered seriously erroneous contour diagrams wherein the central nitrogen of the acridine ring was shown to have H-bond acceptor region as negatively charged electrostatic contour. The central part of the G-quadruplex quartet (Figure 1) is already rich in electron density converting it into an electronegative ion channel that runs along the central region of a DNA G-quadruplex. For inhibition of the enzyme telomerase, the inhibitor has to stack in such a way that the nitrogen atom, of the acridine ring, overlapped with



**Figure 5.** Interacting docking models for three different strategies.

the central electronegative channel. This was not possible with this model as there would be repulsion between the two electron rich species. 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. So, various protonation strategies were applied to the molecules of the series. Nitrogen atoms of the amides and those that were directly connected to phenyl ring were spared from this protonation strategy as these were not supposed to be protonated at physiological pH due to their lower basicity.

**Docking.** To elucidate the interaction mechanism and right protonation strategy of G-quadruplex ligands, compound **12** the most potent ligand in protonated forms with three different protonation strategies (Figure 2) were independently docked with the G-quadruplex. Interestingly, results of docking studies supported our protonation ‘strategy 3’ as its model showed favorable binding interactions with G-

quadruplex DNA as shown in Figure 5. It was observed that a side chain interacted better with the particular nucleotide Thy 1007 and Thy 1008 present in the X-ray crystal structure of G-quadruplex. The particular conformational geometry adopted by the side chains on the two phenyl rings brings aliphatic substituents present on the ninth position of acridine ring significantly deeper into the cavity of the DNA groove. This leads to better interaction with the surface of the groove when the terminal nitrogen atom is in the protonated state, resulting in low interaction energy compared to other strategies (Table 2). The docking experiment showed that the terminal protonated nitrogen atom and the 9-ammonium proton of compound **12** (strategy 3) can form a strong hydrogen bond with Gua 1009, whereas such interactions were not observed in strategies 1 and 2. In strategy 2, though the nitrogen atom present on the C-9 position is in a protonated form, it is not a sufficient structural requirement to produce favorable binding interactions. However, proton-

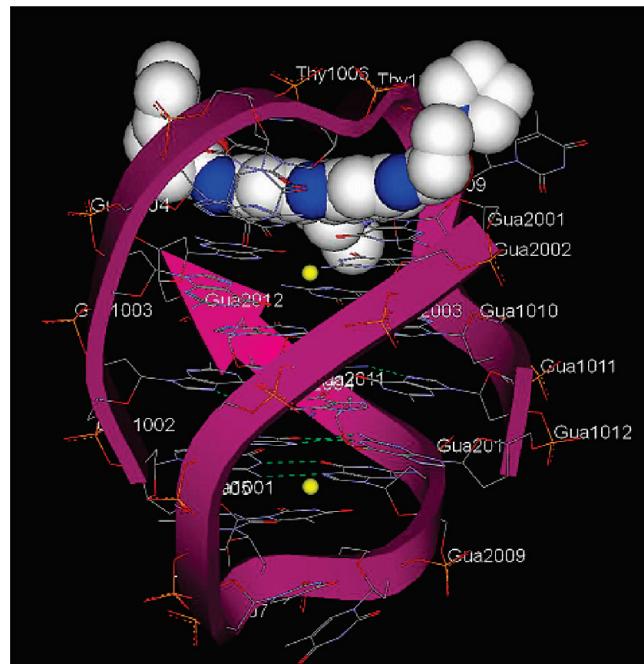
**Table 2.** Interaction Energies and CDOCKER Energy (Score) for Compound **12** Using the Three Different Strategies with Quadruplex, in kcal mol<sup>-1</sup>

strategy	interaction energy	CDOCKER energy
1	-92.549	-45.767
2	-88.629	-50.717
3	-104.321	-58.000

ation of a terminal nitrogen atom along with nitrogen present on the C-9 position (strategy 3) resulted in an ideal binding conformation which offered favorable binding interactions of proton on the N-10 nitrogen of acridine with the electronegative oxygen of Gua 1009 (Figure 5).

The docked view of compound **12** with a solvent accessible surface area is shown in Figure 6, where the white color of the ligand complements the surface area of the binding site, the blue region indicates nitrogen atoms in the ligands, and potassium ions are shown in yellow balls. Thus, strategy 3 was considered ideal for further development of predictive 3D-QSAR CoMFA and CoMSIA models.

**CoMFA Analysis.** After alignment of all training set compounds, Analysis-A (Table 3) yielded a correlation with an  $r^2_{cv}$  of 0.540 (4 optimum number of components), a conventional  $r^2$  of 0.778, an *F*-value of 49.054, and an external predictivity with an  $r^2_{pred}$  of 0.660. The steric and electrostatic contributions were 67.50% and 32.50%, respectively. In order to increase the predictive power of the model further experiments were performed. Based on the results of QSAR studies from partial least-squares (PLS) analysis, 2 molecules (compounds **25** and **47**) from the training set with high residual values (Figure 7) were omitted. Compounds **34** and **52** with similar functional groups at the C9 position as that of compounds **25** and **47**, respectively, are comparatively well predicted. Hence, it is most likely that outliers might be a result of an experimental error that could happen when analyzing a large data set.<sup>52</sup> After removing



**Figure 6.** Docked view of compound **12** shown with solvent accessible surface with nitrogen atom in blue and potassium ions in the central channel of the G-quartet in yellow color.

**Table 3.** PLS Statistical Results of CoMFA (Strategy 3)

parameter	CoMFA	
	Analysis-A	Analysis-B
$r^2_{cv}^a$	0.540	0.552
SEP <sup>b</sup>	0.477	0.474
ONC <sup>c</sup>	4	3
$r^2_{ncv}^d$	0.778	0.783
SEE <sup>e</sup>	0.332	0.330
<i>F</i> value	49.054	67.184
Prob $r^2 = 0$	0	0
$r^2_{pred}$	0.660	0.752
$r^2_{bs}^f$	0.917	0.919
SD <sub>bs</sub> <sup>g</sup>	0.018	0.019
steric contribution	67.50%	62.90%
electrostatic contribution	32.50%	37.10%

<sup>a</sup> Cross-validated  $r^2$ . <sup>b</sup> Standard error of prediction. <sup>c</sup> Optimum number of components. <sup>d</sup> Non-cross-validated  $r^2$ . <sup>e</sup> Standard error of estimate. <sup>f</sup> Boot-strapped  $r^2$ . <sup>g</sup> Bootstrapping standard deviation.

outliers, CoMFA results obtained using the training set of 60 compounds showed a better confidence level of higher statistical significance (Analysis-B Table 3). In 3D-QSAR studies  $r^2_{cv}$  of 0.3 is considered statistically significant.<sup>53</sup> In view of it, the models having  $r^2_{cv} > 0.5$  can be considered to be much better and statistically significant. Analysis-B showed an improved cross-validated  $r^2$  of 0.552 (3 optimum number of components), a conventional  $r^2$  of 0.783, an *F*-value of 67.184, and a predictive  $r^2$  of 0.752. Histograms of residual values of training and test sets are shown in Figure 8. To further assess the robustness of the model, bootstrapping analysis (100 runs) was performed and an  $r^2_{bs}$  of 0.919 ( $SD_{bs} = 0.019$ ) was obtained, further establishing the strength of the model. The steric and electrostatic contributions were found to be 62.90% and 37.10%, respectively. Data set and alignment of CoMFA was further used for CoMSIA analysis.

**CoMSIA Analysis.** CoMSIA is similar to CoMFA but uses a Gaussian function rather than Coulombic and Lennard-Jones potentials to assess the contribution from different fields. CoMSIA was performed using steric, electrostatic, hydrophobic, hydrogen bond donor, and hydrogen bond acceptor fields. 3D-QSAR models were generated using the above fields in different combinations, and the results of study are summarized in Table 4. CoMSIA models showed higher correlation and high predictive properties. In most of the models, hydrophobic field was a common factor indicating the importance of lipophilicity for the present series of molecules. We found that the CoMSIA descriptors such as steric, electrostatic, hydrophobic, and hydrogen bond donor fields played a significant role in the prediction of biological activity. An excellent value of 0.890 for  $r^2$  prediction and 0.606 for  $r^2$  cross-validation with 3 optimum numbers of components were obtained for this model. A good  $r^2_{ncv}$  0.812 was observed for internal prediction of the model. The contribution of steric, electrostatic, hydrophobic, and hydrogen bond donor fields of this model were 12.2%, 24.9%, 23.7%, and 39.2%, respectively. Incorporation of hydrogen bond acceptor field descriptor produced no significant change in internal predictivity. However, a substantial decrease in  $r^2$  prediction from 0.890 to 0.644 was observed which was noteworthy. Thus, incorporation of the hydrogen bond acceptor descriptor to steric, electrostatic, hydrophobic, and hydrogen bond donors in different combination gave statistically poor models. Exclusion of the steric descriptor from

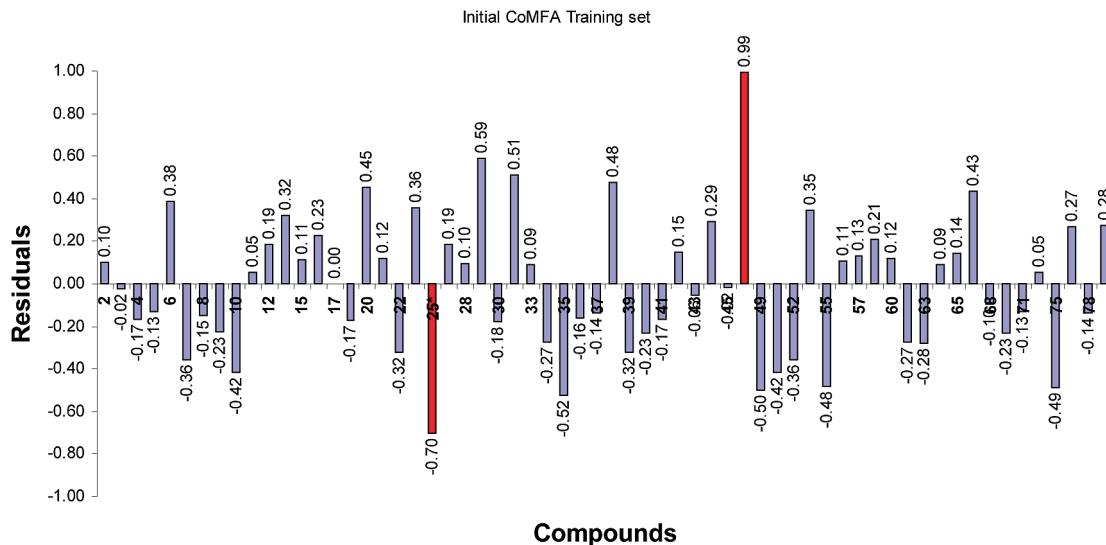


Figure 7. Histogram of CoMFA residual value for initial training set (Analysis-A).

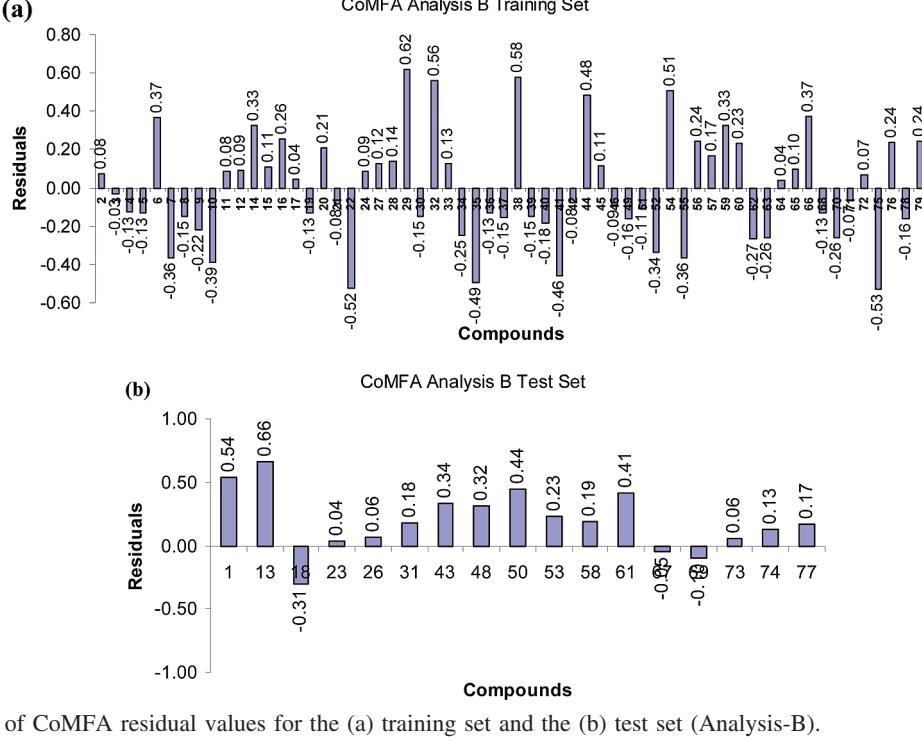


Figure 8. Histogram of CoMFA residual values for the (a) training set and the (b) test set (Analysis-B).

the highly predictive model (SEHD) resulted in slightly improved internal predictivity. However, a little decrease in  $r^2$  prediction from 0.890 to 0.870 was observed. Above findings indicate that steric, electrostatic, hydrophobic, and hydrogen bond donor fields can be more favorable for G-quadruplex stabilization, and they could make an important contribution to the binding energy of association. Overall, positively charged ligands with hydrogen bonding ability and large lipophilic patches could offer good binding properties for G-quadruplex indicating the importance of hydrophobic fields and H-bond donor functional groups for the biological activity. The graphs of actual vs predicted activities for training and test set molecules from the best CoMSIA model (Steric, Electrostatic, Hydrophobic, and Hydrogen bond donor descriptors) are shown in Figure 9a,b, and the histogram of residual values for training and test set molecules are shown in Figure 10a,b.

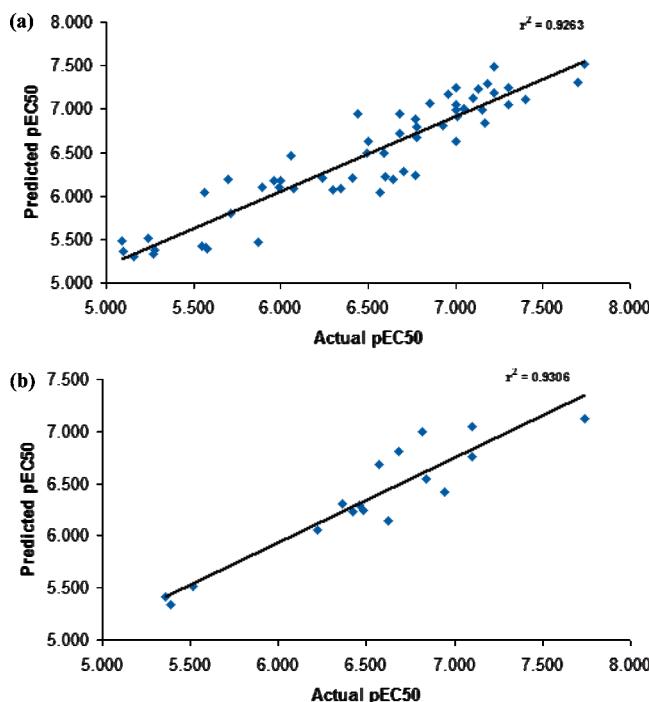
**Model Validation.** *a) Internal Validation.* In model validation we examined the internal predictive power of the models and their ability to reproduce biological activities of the compounds for the training set. The computed telomerase inhibitory activity from the CoMFA and CoMSIA showed a good correlation with experimental telomerase inhibitory activity.

*b) External Validation.* External validation of QSAR models was done to verify the excellent statistical parameters that were obtained and also to investigate whether the activity of substituted acridines from external data could be predicted well with this model. A large data set of compounds containing 17 molecules, as a test set (Table 1), was selected from the data set of 77 compounds for the validation experiments. The ultimate test for predictability of QSAR analysis in the drug design process is to predict the biological activity of new compounds that have not been included in

**Table 4.** Summary of CoMSIA Analysis<sup>a</sup>

	HAD	SED	SEA	SHE	EHD	SEDA	<b>SEHD</b>	SEHA	SEHDA
$r^2_{cv}$	0.574	0.568	0.532	0.607	0.585	0.552	<b>0.606</b>	0.562	0.613
ONC	3	3	2	3	3	3	<b>3</b>	3	3
SEP	0.466	0.469	0.484	0.448	0.456	0.478	<b>0.444</b>	0.473	0.444
$r^2_{ncv}$	0.806	0.821	0.658	0.794	0.826	0.816	<b>0.812</b>	0.827	0.832
SEE	0.314	0.302	0.414	0.324	0.295	0.306	<b>0.300</b>	0.297	0.274
F-value	76.289	84.011	53.768	70.806	88.437	81.515	<b>80.440</b>	87.479	102.024
P $r^2=0$	0	0	0	0	0	0	0	0	0
Contribution (fraction)									
S	-	0.164	0.159	0.195	-	0.148	<b>0.122</b>	0.154	0.109
E	-	0.329	0.322	0.406	0.284	0.253	<b>0.249</b>	0.264	0.191
H	0.323	-	-	0.399	0.286	-	<b>0.237</b>	0.280	0.218
D	0.395	0.507	-	-	0.430	0.347	<b>0.392</b>	-	0.267
A	0.282	-	0.519	-	-	0.252	-	0.303	0.215
$r^2_{pred}$	0.567	0.823	-	0.895	0.870	0.514	<b>0.890</b>	0.644	0.643
$r^2_{bs}$	0.929	0.917	0.865	0.918	0.890	0.915	<b>0.908</b>	0.910	0.923
SD <sub>bs</sub>	0.016	0.018	0.029	0.019	0.020	0.032	<b>0.016</b>	0.017	0.015

<sup>a</sup> ONC = optimum number of components, SEP = standard error of prediction, SEE = standard error of estimate, S = steric, E = electrostatic, D = H-bond donor, A = H-bond acceptor, and H = hydrophobic.



**Figure 9.** Graph of actual vs predicted activities for training and test set molecules from the best predictive CoMSIA model: (a) training set and (b) test set.

the training set. The values of  $r^2_{pred}$  were calculated for the test set and gave the best results for CoMFA and CoMSIA with the values 0.752 and 0.890, respectively. Thus, the CoMSIA model displays higher predictivity both in regular cross-validation and in the prediction of the test compounds.

c) *Fischer Statistics (F-Test)*. Fischer statistics (*F*) is the ratio between explained and unexplained variance for a given number of degrees of freedom. *F*-test is a 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 equation is significant. The *F* values for the CoMFA and CoMSIA models were 67.184 and 80.440, respectively, [CoMFA/CoMSIA  $F_{05}$  (3, 54) = 2.84 (Tab)]

at 95% confidence level, which suggests that these models are statistically significant with CoMSIA having an upper hand over CoMFA.

**Visualization of Contour Maps.** CoMFA steric and electrostatic contour maps are not much different than CoMSIA, hence these are not discussed. CoMSIA steric fields are shown in Figure 11. The green contours represent regions of high steric tolerance (80% contribution), while the yellow contours represent regions of low steric bulk tolerance (20% contribution). Prominent green contours present in the vicinity of the ninth position of the acridine ring and small contours present surrounding the pyrrolidine rings indicate that generally steric bulk is favored at these sites. The good inhibitory potency of compounds **12** ( $^{tel}EC_{50}$  0.08) and **13** ( $^{tel}EC_{50}$  0.08) is due to orientation of the C-9 substituted bulkier group toward the sterically favored regions. While in the case of compounds **63** ( $^{tel}EC_{50}$  8.2), **68** ( $^{tel}EC_{50}$  5.4), **70** ( $^{tel}EC_{50}$  8.0), **71** ( $^{tel}EC_{50}$  5.2), and **73** ( $^{tel}EC_{50}$  3.1) there was an absence of a bulkier group in this region. The cycloheptane ring present at the C9 position of compound **10** ( $^{tel}EC_{50}$  0.21) was oriented away from the prominent green contour, and the C3 side chain was directed toward a sterically unfavorable yellow region that may lead to steric clashes in these regions. This makes compound **10** comparatively less active among the 3,6,9-substituted acridine derivatives. 2,6,9-Substituted acridine compounds (**21** to **27**) show less activity than that of 3,6,9-substituted acridine as the side chain substituted at the C2 position directs toward sterically unfavorable regions. Compound **20** ( $^{tel}EC_{50}$  0.08) showed good inhibitory potency as a sterically favored space was occupied by the ninth substituted aryl ring. In the case of the 2,7,9-substituted acridine side chain at the C-7 position of acridines, it directs toward yellow contours which restrict the steric substitution. This is in line with SAR indicating lower activities for compounds **30** ( $^{tel}EC_{50}$  1.09), **34** ( $^{tel}EC_{50}$  1.29), **35** ( $^{tel}EC_{50}$  2.73), and **36** ( $^{tel}EC_{50}$  1.03). The case of compound **38** ( $^{tel}EC_{50}$  0.099) with three carbons in C-3, C-6 substituted side chains is more favorable as these side chains are perfectly entrenched in sterically favored regions. Incorporation of an additional -CH<sub>2</sub> group in the carbon spacer chain directs the pyrrolidine ring of either side into disfavorable yellow contours, e.g. compounds **39** ( $^{tel}EC_{50}$

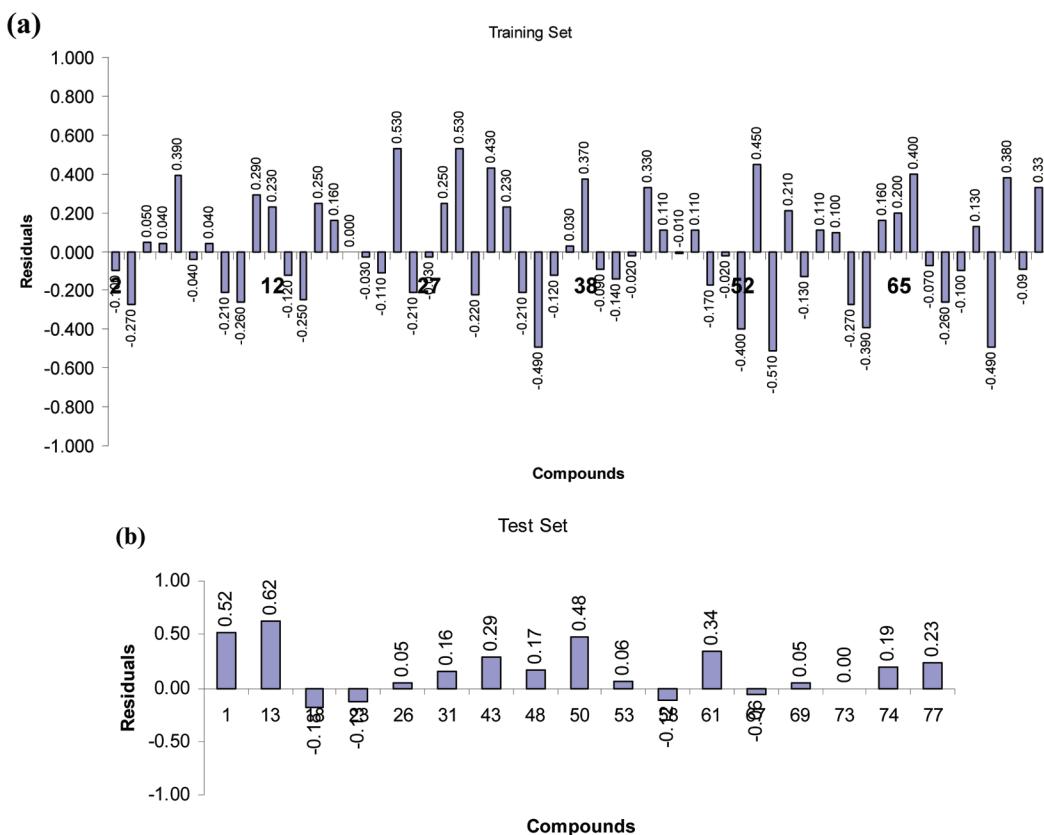


Figure 10. Histogram of CoMSIA residual values for the training set (a) and the (b) test set.

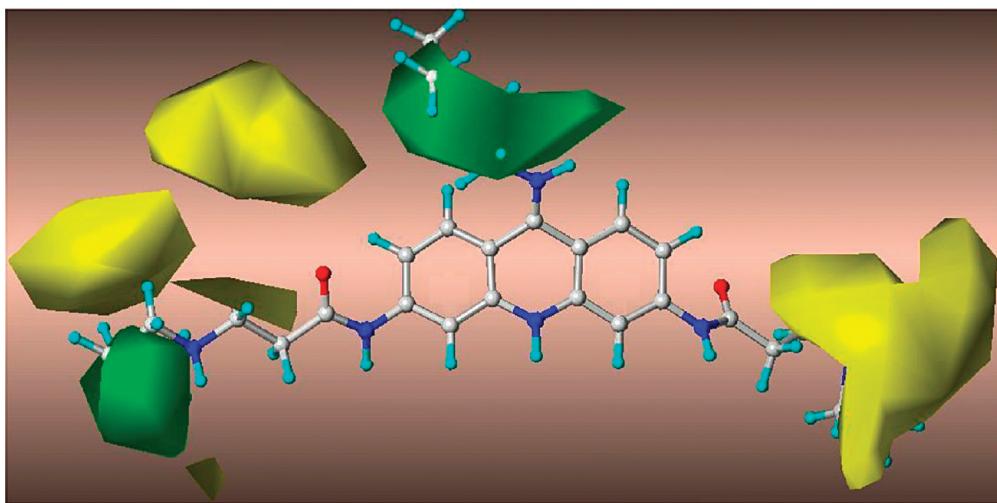
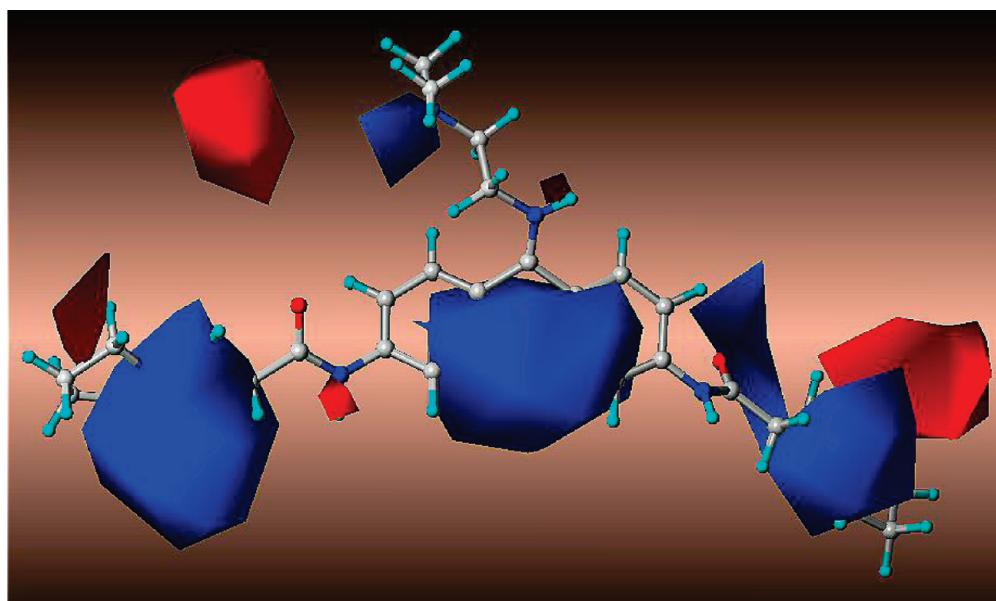


Figure 11. CoMFA stdev\*coeff steric contour maps. The most active compound **12** is displayed in the background. Green and yellow polyhedra indicate regions where more steric bulk or less steric bulk, respectively, will enhance the activity.

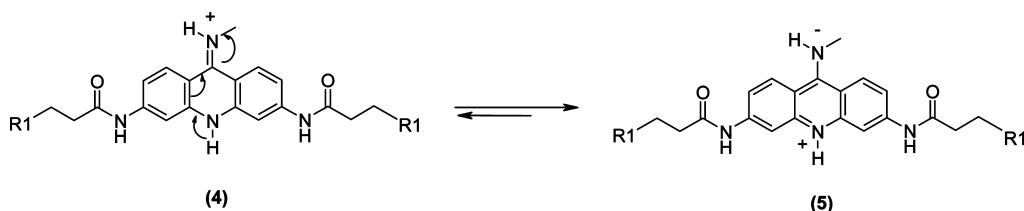
1.93) and **40** ( ${}^{\text{tel}}\text{EC}_{50}$  6.91). Compounds **49** ( ${}^{\text{tel}}\text{EC}_{50}$  1.00), **51** ( ${}^{\text{tel}}\text{EC}_{50}$  0.86), and **52** ( ${}^{\text{tel}}\text{EC}_{50}$  0.88) are relatively less active because the benzylamino substituent changes its conformation and directs away from the sterically favorable green contours. Also, the 3,6-substituted side chain changes its orientation to sterically unfavorable yellow contours. This could be one of the reasons why these compounds showed lower activity than anilino substituted acridines. Among the polycyclic acridines, compound **75** ( ${}^{\text{tel}}\text{EC}_{50}$  2.0) showed lower activity because of the protrusion of the  $-\text{CH}_2\text{CH}_3$  group away from the sterically favorable green contours.

CoMSIA electrostatic fields are shown in Figure 12. The increase in positive charge is favored in the blue regions, while the increase in negative charge is favored in the red

regions. The electrostatic contours of CoMSIA show blue regions surrounding the central ring nitrogen atom of acridine indicating that incorporation of electropositive substituents would enhance the activity. Blue contours in the vicinity of pyrrolidine rings of both of the side chains substituted at C-3 and C-6 of the acridine ring have been observed. A blue favorable electrostatic region is observed above the acridine ring in the model. This particularly explains the importance of the acridine chromophore for  $\pi-\pi$  or van der Waals interactions with G-quadruplex. These findings could lead to new generation of acridine analogues by introducing more electropositive atoms at these positions. Though, the molecular structure of acridine derivatives was built with a general structure (**4**) (Figure 13), the compounds predomi-



**Figure 12.** CoMFA stdev\*coeff electrostatic contour maps. The most active compound **12** is displayed in the background. Blue contours indicate regions where electropositive groups increase activity, whereas red contours indicate regions where electronegative groups increase activity.



**Figure 13.** Predominant structure of acridine.

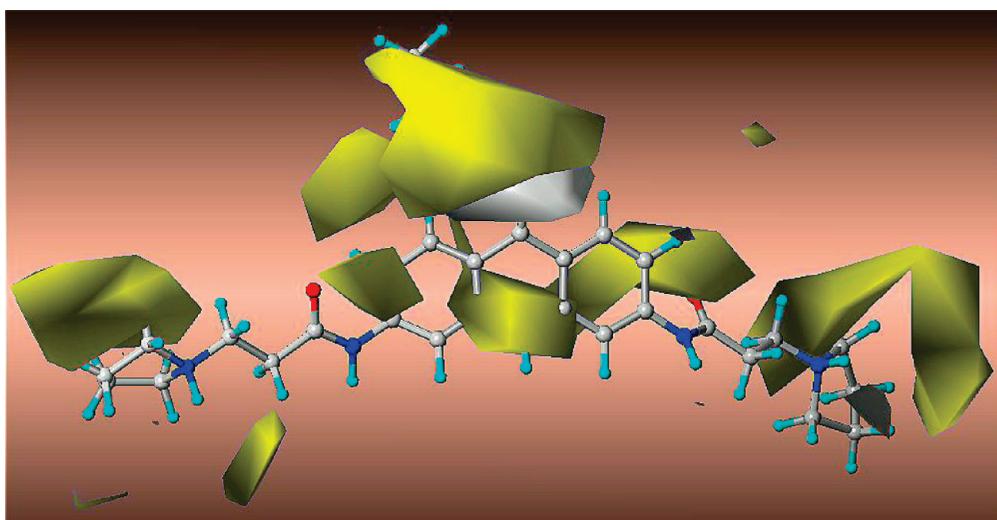
nantly exist in the structural form (**5**) as shown in Figure 13. A big blue contour surrounding N-10 and a red contour at the nitrogen attached to C-9 is indicative of a positively charged nitrogen (N-10) and a lone pair of electrons on a nitrogen attached to C-9, respectively. This is a very important finding which shows that N-10 is much more basic than nitrogen attached to C-9 and gets protonated faster. Aromaticity of all the three rings of acridine is also preserved, making structure **5** to be much more stable than structure **4** (Figure 13).

Compound **12** ( ${}^{tel}EC_{50}$  0.018) was found to be the most active among the series due to the reason that the terminal nitrogen atom on ninth substituents, pyrrolidine ring nitrogen atoms on both the C3 and C6 side chains, and the central ring nitrogen atom of the acridine are perfectly entrenched into positively charged favored blue contours. Among the 3,6,9-trisubstituted acridines compound **9** ( ${}^{tel}EC_{50}$  0.14) showed poor activity probably due to the fact that the terminal -OCH<sub>3</sub> group is embedded in an electropositive blue region. The influence of the positively charged favored blue region was seen in compound **20** ( ${}^{tel}EC_{50}$  0.08), being the most potent among 2,6,9-trisubstituted acridines. Compound **20** showed better telomerase inhibitory activity than compound **22** ( ${}^{tel}EC_{50}$  0.20) among the 2,6,9-trisubstituted acridines due to the reason that the p-NH<sub>2</sub> group (which might not get protonated due to its lesser basicity) on the phenyl ring is oriented in the blue region. However, the terminal nitrogen on the C-9 substituted aliphatic chain in compound **22** is oriented away from the positively charged favored region.

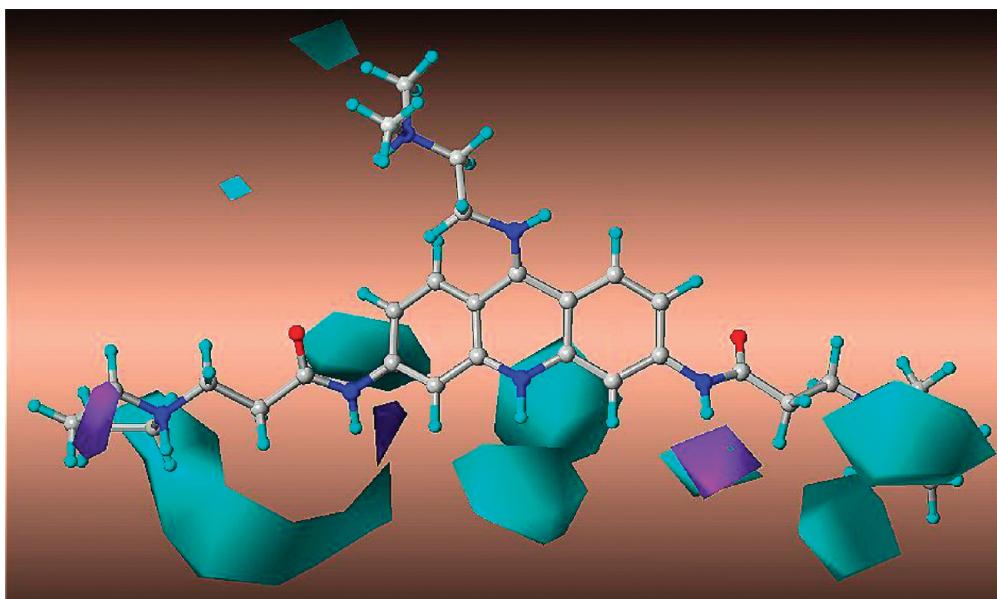
Compounds **63** to **72** exhibit comparatively low activity owing to the absence of electropositive side chains at the C-9 position of the acridine ring. Compounds **34** ( ${}^{tel}EC_{50}$  1.29) and **35** ( ${}^{tel}EC_{50}$  2.73) exhibit low activity as an electropositive pyrrolidine nitrogen present on the C-7 substituted side chain is embedded in electronegative red contours. In addition to this in the case of compound **35** the -OCH<sub>3</sub> group present on the phenyl ring is also oriented toward electropositive blue contours. This could be the reason for poorer activity of compound **35** than compound **34**. Electropositive hydrogen atoms of the C-3 and C-6 substituted side chain in compound **40** ( ${}^{tel}EC_{50}$  6.91) embedded in negatively charged favored red regions could be the reason for its lower activity.

Hydrophobic maps shown in Figure 14 indicate that the lipophilic favorable yellow region is found surrounding the aromatic acridine ring. This suggests that lipophilicity of this portion of the molecule is an important factor for the activity. The white region under yellow contour at the ninth position of the acridine ring indicates that hydrophilic atoms in this zone favor the inhibitory activity.

Compounds with aniline and benzylaniline substituents at the C-9 position of acridine show good activity as an aromatic ring fits well into the yellow region which is favorable for hydrophobic groups. In compounds **62** to **73**, absence of lipophilic substituents at the C-9 position of acridine could be one of the reasons for their minimal range of activity. The lipophilic ethyl group of compound **75** ( ${}^{tel}EC_{50}$  2.00) embedded in an unfavorable white region is responsible for its low activity.



**Figure 14.** CoMSIA stdev\*coeff hydrophobic contour maps. The most active compound **12** is displayed in the background. Yellow contours indicate regions where hydrophobic groups increase activity, white contours indicate regions where hydrophobic groups decrease activity.

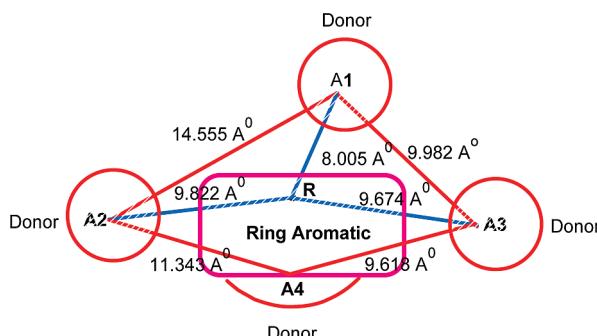


**Figure 15.** CoMSIA stdev\*coeff hydrogen bond donor contour maps. The most active compound **12** is displayed in the background. Cyan contours indicate regions where the H-bond donor group increases activity, whereas purple contours indicate regions where the H-bond donor group decreases activity.

The graphical interpretation of the field's contributions of the H-bond donor properties is shown in Figure 15. In principle H-bond donor contours should highlight areas beyond the ligands where putative hydrogen partners in the target could form H-bonds to influence binding affinity. A cyan colored hydrogen bond donor region surrounds the C-9 position terminal nitrogen atom, pyrrolidine ring nitrogen atoms, and the central ring nitrogen atom of the acridine ring stating a favorable region for H-bond donor fields. Purple disfavorable regions are found in the vicinity of the pyrrolidine ring substituted on C-6 and away from the amido alkyl chain substituted at C-3 of the acridine ring. Compound **20** ( $^{tel}EC_{50}$  0.08) which contains a *p*-NH<sub>2</sub> group on the phenyl ring oriented toward the cyan region has a high potency among 2,6,9-trisubstituted acridines. Compound **34** ( $^{tel}EC_{50}$  1.29) and compound **35** ( $^{tel}EC_{50}$  2.73) show lower activity as the C-2 and C-7 substituted side chains are oriented away from the favorable cyan region.

The role of hydrogen bond donor contributions can be clearly explained by analyzing the compounds **39** and **40**. Pyrrolidine ring nitrogen atoms present on the C-3 and C-6 positions of acridine in compound **39** ( $^{tel}EC_{50}$  1.93) and compound **40** ( $^{tel}EC_{50}$  6.91) cross the hydrogen bond donor favorable cyan region due to an increase in side chain length. This could be the reason for their low activities. Compounds **62** to **73** exhibit lower activity due to an unsubstituted C-9 position which plays a key role in hydrophobic and hydrogen-bond donors fields.

**Pharmacophore Model.** Using the results obtained by the best predictive CoMSIA model, a putative pharmacophore model has been proposed. To propose the pharmacophore model, the distance of contour maps surrounding the H-bond donor groups was measured. In the proposed preliminary pharmacophore model, pharmacophoric distances using the G-quadruplex stabilizing ligands resulted in distances of 14.55 Å (A1-A2), 9.982 Å (A1-A3), 18.840 Å (A2-A3),



**Figure 16.** Pharmacophore model from the CoMSIA study.

11.343 Å (A2-A4), 9.618 Å (A3-A4), 8.005 Å (R-A1), 9.822 Å (R-A2), and 9.674 Å (R-A3). The 5 point pharmacophore model along with the interfeature distances are shown in Figure 16. The key structural components for G-quadruplex stabilization are bulky hydrophobic and electropositive C-9 substitution, cationic sites, and a tricyclic aromatic proton donor pharmacophore which may be essential for perfect orientation of side chains into quadruplex grooves.

## CONCLUSION

Although there is a high degree of flexibility in compounds in the training set, 3D-QSAR models with statistical significance and good predictive abilities by using CoMFA and CoMSIA were obtained. A high bootstrapped  $r^2$  value and a small standard deviation indicate that a similar relationship exists in all compounds. In the CoMFA model both the steric and electrostatic parameters significantly contribute toward the activity. The CoMSIA model generated by a combination of steric, electrostatic, hydrophobic, and hydrogen bond donor fields showed good correlative and predictive properties. The predictive power of both the derived models was assessed by the test set molecules.

After a detailed study of CoMFA and CoMSIA, it could be concluded that (1) functionality attached to the ninth position of the acridine ring having some steric bulk with a positive charge may lead to compounds with higher activity. Bulky electron withdrawing groups along with high hydrophobicity may lead to improved telomerase inhibitory activity. (2) The number of protonated nitrogen atoms in the molecule brings charge transfer electrostatic interactions with higher chemical stability in the biochemical system and crucially govern the telomerase inhibitory activity. (3) Maintaining the bulkiness of the molecule and providing specific conformation to the molecule can also achieve an optimum G-quadruplex stabilization thereby increasing telomerase inhibition. (4) Hydrogen bonds are considered to make an important discriminatory contribution in the binding of ligands to G-quadruplex; this forms the basis for a strategy to design new G-quadruplex-binding ligands. Overall, the present 3D-QSAR study investigates the indispensable structural features of substituted acridines which can be exploited for structural modifications in order to achieve improved telomerase inhibitory activity. It is clear that 3D-QSAR information obtained from this kind of analysis provides 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. These proposed models can also

be used to predict the activity of newly designed analogs, prior to synthesis. It is hoped that these models could help in design and optimization of better telomerase inhibitors which are promising anticancer agents.

## ACKNOWLEDGMENT

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 AND NOTES

- (1) Blackburn, E. Telomeres-no end in sight. *Cell* **1994**, *77*, 621–623.
- (2) Zakiyan, V. Telomeres-beginning to understand the end. *Science* **1995**, *270*, 1601–1607.
- (3) Delange, T. Activation of telomerase in a human tumor. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 2882–2885.
- (4) Bryan, T.; Cech, T. Telomerase and the maintenance of chromosome ends. *Curr. Opin. Cell. Biol.* **1999**, *11*, 318–324.
- (5) Harley, C.; Villeponteau, B. Telomeres and telomerase in aging and cancer. *Curr. Opin. Genet. Dev.* **1995**, *5*, 249–255.
- (6) Kim, N.; Piatszczk, M.; Prowse, K.; Harley, C.; West, M.; Ho, P. L. C.; et al. Specific association of human telomerase activity with immortal cells and cancer. *Science* **1994**, *266*, 2011–2015.
- (7) Mergny, J.; Mailliet, P.; Lavelle, F.; Riou, J.; Laoui, A.; Helene, C. The development of telomerase inhibitors: G-quartet approach. *Anti-Cancer Drug Des.* **1999**, *14*, 327–339.
- (8) Perry, P.; Jenkins, T. C. Recent advances in the development of the telomerase inhibitors for the treatment of cancer. *Expert Opin. Invest. Drugs* **1999**, *8*, 1981–2008.
- (9) Raymond, E.; Sorid, J. C.; Izbicica, E.; Baussin, F.; Hurley, L.; Von Hoff, D. DNA G-quadruplexes, telomere-specific proteins and telomere associated enzymes as potential target for new anticancer drugs. *Invest. New Drugs* **2000**, *18*, 123–137.
- (10) Neidle, S.; Parkinson, G. Telomere maintenance as a target for anticancer drug discovery. *Nat. Rev. Drug Discovery* **2002**, *1*, 383–393.
- (11) Gellert, G.; Dikman, Z.; Wright, W.; Gryaznov, S.; Shay, J. Effect of novel telomerase inhibitor, GRN163L, in human breast cancer. *Breast Cancer Res. Treat.* **2006**, *96*, 73–81.
- (12) Blackburn, E. Switching and signaling at the telomere. *Cell* **2001**, *106*, 661–73.
- (13) Wright, W.; Tesmer, V.; Huffman, K.; Levene, S.; Shay, J. Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes Dev.* **1997**, *11*, 2801–2809.
- (14) Sfeir, A.; Chai, W.; Shay, J.; Wright, W. Telomere-end processing: the terminal nucleotides of human chromosomes. *Mol. Cell* **2005**, *18*, 131–138.
- (15) Zaung, A.; Podell, E.; Cech, T. Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 10864–10869.
- (16) Zahler, A.; Williamson, J.; Cech, T. Inhibition of telomerase by G-quartet DNA structures. *Nature (London)* **1991**, *350*, 718–720.
- (17) Oganessian, L.; Moon, I. K.; Bryan, T.; Jarstfer, M. Extension of G-quadruplex DNA by ciliate telomerase. *EMBO J.* **2006**, *25*, 1148–1159.
- (18) Sun, D.; Thompson, B.; Cathers, B.; Salazar, M.; Kerwin, S.; Trent, J.; Jenkins, T.; Neidle, S.; Hurley, L. Inhibition of human telomerase by a G-quadruplex interactive compound. *J. Med. Chem.* **1997**, *40*, 2113–2116.
- (19) Harrison, R.; Gowan, S.; Klland, L.; Neidle, S. Human telomerase inhibition by substituted acridine derivatives. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2463–2468.
- (20) Read, M. A.; Wood, A.; Harrison, R.; Gowan, S.; Klland, L.; Dosanjh, H.; Neidle, S. Molecular modeling studies on G-quadruplex complexes of telomerase inhibitors: Structure activity relationships. *J. Med. Chem.* **1999**, *42*, 4538–4546.
- (21) Han, H.; Hurley, L.; Salazar, M. A DNA polymerase stop assay for G-quadruplex-interactive compounds. *Nucleic Acid Res.* **1999**, *27*, 537–542.
- (22) Rossetti, L.; Franceschin, M.; Blanco, A.; Ortaggi, G.; Savino, M. Perylene diimides with different side chains are selective in inducing different G-quadruplex DNA structures and in inhibiting telomerase. *Bioorg. Med. Chem. Lett.* **2008**, *12*, 2527–2533.

- (23) Koeppe, F.; Riou, J.; Laoui, A.; Mailliet, P.; Arimondo, P.; Labit, D.; Petitgenet, O.; Helene, C.; Mergny, J. Ethidium derivatives bind to G-quartets, inhibit telomerase and act as fluorescent probes for quadruplexes. *Nucleic Acid Res.* **2001**, *29*, 1087–1096.
- (24) Kim, M.; Duan, W.; Gleason-Guzman, M.; Hurley, L. Design, synthesis, and biological evaluation of a series of fluoroquinoanthroazines with contrasting dual mechanisms of action against topoisomerase II and G-quadruplexes. *J. Med. Chem.* **2003**, *46*, 571–583.
- (25) Caprio, V.; Guyen, B.; Opoku-Boahen, Y.; Mann, J.; Gowan, S.; Kelland, L.; Read, M.; Neidle, S. A novel inhibitor of human telomerase derived from 10H-indolo[3,2-b]quinoline. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2063–2066.
- (26) Mergny, J.; Lacroix, L.; Teulade-Fichou, M.; Hounslow, C.; Guittat, L.; Hoarau, M.; Arimondo, P.; Vigneron, J.; Lehn, J.; Riou, J.; Garesti, T.; Helene, C. Telomerase inhibitors based on quadruplex ligands selected by a fluorescence assay. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 3062–3067.
- (27) Teulade-Fichou, M.; Carrasco, C.; Guittat, L.; Bailly, C.; Alberti, P.; Mergny, J.; David, A.; Lehn, J.; Wilson, W. Selective recognition of G-quadruplex telomeric DNA by a bis(quinacridine) macrocycle. *J. Am. Chem. Soc.* **2003**, *125*, 4732–4740.
- (28) Gowan, S.; Heald, R.; Stevens, M.; Kelland, L. Potent inhibition of telomerase by small-molecule pentacyclic acridines capable of interacting with G-quadruplexes. *Mol. Pharmacol.* **2001**, *60*, 981–988.
- (29) Heald, R.; Modi, C.; Cookson, J.; Hutchinson, I.; Laughton, C.; Gowan, S.; Kelland, L.; Stevens, M. Antitumor polycyclic acridines. 8 (1) synthesis and telomerase-inhibitory activity of methylated pentacyclic acridinium salts. *J. Med. Chem.* **2002**, *45*, 590–597.
- (30) Shin-ya, K.; Wierbla, K.; Matsuo, K.; Ohtani, T.; Yamada, Y.; Furihata, K.; Hayakawa, Y.; Seto, H. Telomestatin, a novel telomerase inhibitor from Streptomyces anulatus. *J. Am. Chem. Soc.* **2001**, *123*, 1262–1263.
- (31) De Cian, A.; Delemos, E.; Mergny, J. L.; Teulade-Fichou, M.; Monchoud, D. Highly efficient G-quadruplex recognition by bisquinoxolinium compounds. *J. Am. Chem. Soc.* **2007**, *129*, 1856–1857.
- (32) Pennarun, G.; Granotier, C.; Gauthier, L.; Gomez, D.; Hoffschir, F.; Mandine, E.; Riou, J.; Mergny, J.; Mailliet, P.; Boussin, F. Apoptosis related to telomere instability and cell cycle alterations in human glioma cells treated by new highly selective G-quadruplex ligands. *Oncogene* **2005**, *24*, 2917–2928.
- (33) Read, M.; Harrison, R.; Ramagnoli, B.; Tanious, F.; Gowan, S.; Reszka, A. P.; Wilson, W.; Kelland, L.; Neidle, S. Structure based design of selective and potent G-quadruplex mediated telomerase inhibitors. *Proc. Natl. Acad. Sci.* **2001**, *98*, 4844–4849.
- (34) Harrison, R.; Cuesta, J.; Chessari, G.; Read, M.; Basra, S.; Reszka, A.; Morrell, J.; Gowan, S.; Incles, C.; Tanious, F.; Wilson, W.; Kelland, L.; Neidle, S. Trisubstituted acridine derivatives as potent and selective telomerase inhibitors. *J. Med. Chem.* **2003**, *46*, 4463–4476.
- (35) Schultes, C.; Guyen, B.; Cuesta, J.; Neidle, S. Synthesis, biophysical and biological evaluation of 3,6-bis-amidoacridines with extended 9-anilino substituents as potent G-quadruplex binding telomerase inhibitors. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4347–4351.
- (36) Moore, M.; Schultes, C.; Cuesta, J.; Cuenga, F.; Gunaratnam, M.; Tanious, F.; Wilson, W.; Neidle, S. Trisubstituted acridines as G-quadruplex telomere targeting agent. Effect of extension of the 3, 6- and 9-sidechains on quadruplex binding, telomerase activity and cell proliferation. *J. Med. Chem.* **2006**, *49*, 582–599.
- (37) Martins, C.; Gunaratnam, M.; Stuart, J.; Makwana, V.; Greciano, O.; Reszka, A.; Kelland, L.; Neidle, S. Structure based design of benzylamini-acridine compounds as G-quadruplex DNA telomere targeting agents. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 2293–2298.
- (38) Cramer, R., III; Patterson, D.; Bunce, J. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.
- (39) Klebe, G.; Abraham, U.; Meitzner, T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J. Med. Chem.* **1994**, *37*, 4130–4146.
- (40) Puntambekar, D.; Giridhar, R.; Yadav, M. 3D-QSAR CoMFA/CoMSIA studies on 5-aryl-2,2-dialkyl-4-phenyl-3(2H)-furanone derivatives, as selective COX-2 inhibitors. *Acta Pharm.* **2006**, *56*, 157–174.
- (41) Puntambekar, D.; Giridhar, R.; Yadav, M. Understanding the antitumor activity of novel tricyclicpiperazinyl derivatives as farnesyltransferase inhibitors using CoMFA and CoMSIA. *Eur. J. Med. Chem.* **2006**, *41*, 1279–1292.
- (42) Puntambekar, D.; Giridhar, R.; Yadav, M. 3D-QSAR studies of farnesyltransferase inhibitors: a comparative molecular field analysis approach. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1821–1827.
- (43) Puntambekar, D.; Giridhar, R.; Yadav, M. Insights into the structural requirements of farnesyltransferase inhibitors as potential anti-tumor agents based on 3D-QSAR CoMFA and CoMSIA models. *Eur. J. Med. Chem.* **2008**, *43*, 142–154.
- (44) Murumkar, P.; Giridhar, R.; Yadav, M. 3D-quantitative structure-activity relationship studies on benzothiadiazepine hydroxamates as inhibitors of tumor necrosis factor-alpha converting enzyme. *Chem. Biol. Drug Des.* **2008**, *71*, 363–373.
- (45) Murumkar, P.; Dasgupta, S.; Zambre, V.; Giridhar, R.; Yadav, M. Development of Predictive 3D-QSAR CoMFA and CoMSIA Models for  $\beta$ -aminohydroxamic acid-derived TACE inhibitors. *Chem. Biol. Drug Des.* **2009**, *73*, 97–107.
- (46) Golbraikh, A.; Tropsha, A. Beware of q2. *J. Mol. Graphics Modell.* **2002**, *20*, 269–276.
- (47) SYBYL 7.0; Tripos Inc., 1699 South Hanley Road, St. Louis, Missouri 63144, U.S.A.
- (48) Stewart, J. MOPAC: a semiempirical molecular orbital program. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1–105.
- (49) Dewar, M.; Zoebisch, E.; Healy, E.; Stewart, J. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- (50) Wu, G.; Robertson, D.; Brooks, C., III; Vieth, M. Detailed analysis of Grid-based molecular docking: A case study of CDOCKER-A CHARMM-based MD docking algorithm. *J. Comput. Chem.* **2003**, *24*, 1549–1562.
- (51) Haider, S.; Parkinson, G.; Neidle, S. Structure of G-quadruplex-ligand complex. *J. Mol. Biol.* **2003**, *326*, 117–125.
- (52) Devillers, J.; Lipnick, R. In *Practical Applications of Quantitative-Structure Activity Relationships (QSAR) in Environmental chemistry and Toxicology*, 1st ed.; Karcher, W., Devillers, J., Eds.; Kluwer: Dordrecht, 1990; Vol. 1, pp 129–143.
- (53) Bohm, M.; Sturzebecher, J.; Klebe, G. Three-dimensional quantitative structure-activity relationship analyses using comparative molecular field analysis and comparative molecular similarity indices analysis to elucidate selectivity differences of inhibitors binding to trypsin, thrombin, and factor Xa. *J. Med. Chem.* **1999**, *42*, 458–477.

CI900036W