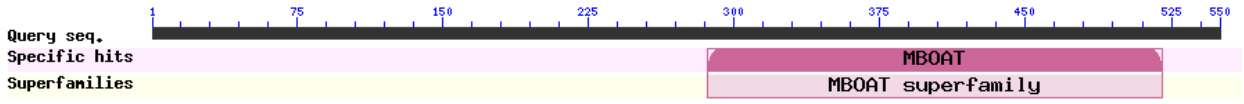
4. RESULTS AND DISCUSSION

Sterol O-acyltransferase (SOAT) (Acyl-CoA cholesterol acyltransferase, Acyl-CoA cholesterin acyltransferase ACAT) is a member of membrane bound O-acyl transferase (MBOAT) family. It is involved in the conversion of cholesterol to the cholesteryl esters. All members of MBOAT family have several transmembrane domains and majority of them possess two conserved residues which may be the active-site residues, a conserved histidine (His) embedded in a hydrophobic stretch of residues and an asparagine (Asn) or histidine within a more hydrophilic region some 30-50 residues upstream. There are two important isoforms of SOAT i.e. SOAT-1 and SOAT-2 which consist of 550 and 522 amino acids respectively.[16,152] Till date the 3D structures of both of these isoforms have not been reported. As there is no homologous protein present for the conserved part of SOAT, it is very difficult to generate the 3D structure of SOAT. So, it was planned to generate a 3D structure of the conserved part of SOAT using *de novo* protein modeling. With the help of NCBI Blast the MBOAT conserved region (amino acids 286 to 520) of the isoform SOAT-1 has been identified which represents the functional unit of the target protein (**Figure 4**).

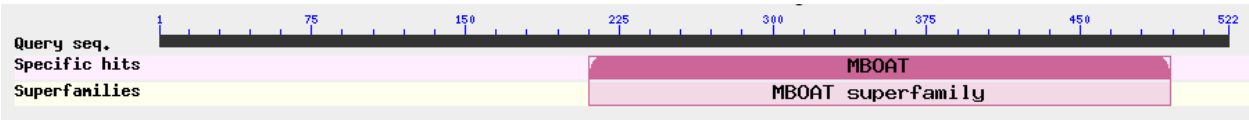


>protein
SSTVPIPTVNQYLYFLFAPTLIYRDSYPRNPTVRWGYVAMKFAQVFGCFFYVYYIFERLCAPLFRNIKQEPFSA
RVLVLCVFNSILPGVLILFLTFFAFLHCWLNAFAEMLRFGDRMFYKDWWNSTSYSNYYRTWNVVVHDWLYYYAY
KDFLWFFSKRFKSAAMLAVFAVSAVVHEYALAVCLSFFYPVLFVLFMFFGMAFNFIVNDSRKKPIWNVLMWTSL
FLGNGVLLCFYSQ

Figure 4: MBOAT sequence for ACAT-1 (amino acids 286 to 520).

The I-TASSER (Iterative Threading ASSEmbly Refinement)[157-159] is a hierarchical method for prediction of the protein structure and function. In this method the structural templates are initially identified from the PDB by multiple threading approach LOMETS (Local Meta-Threading-Server) which is a locally installed PDB library for meta-threading approach having multiple threading programs in I-TASSER server; then the full-length atomic models are developed by iterative template fragment assembly simulations; at the end the function insights

of the target are derived by threading the 3D models through protein function database BioLiP (which is high-quality, biologically relevant ligand-protein binding interactions database). The 3D-structure of ACAT-1 using the MBOAT sequence obtained from NCBI blast technique was developed by using this highly predictive web based server. As the isoforms 1 and 2 are having almost 55% overall sequence identity and more than 60% MBOAT sequence identity, the 3D-structure of ACAT-2 by using MBOAT sequence (**Figure 5**) was developed by homology modeling using the developed ACAT-1 model as template structure in MODELLER. As the probable active site residue His-460 for ACAT-1 and His-434 for ACAT-2 are reported, both of them were used to identify the active sites in both the developed receptor models. Further, the docking study of the most active ligand (**1**) was performed on the predicted active site and discussed here after the discussion on the pharmacophore and atom based 3D-QSAR models. The pharmacophore and 3D-QSAR models were developed on 32 urea derivatives developed as ACAT inhibitors in our laboratory. For the development of pharmacophore and 3D-QSAR models, PHASE module from Schrodinger was implemented in this study.



>protein
VLCALPVHVAVEHQLPPASRCVLVFEQVRFLMKSYSFLREAVPGTLRARRGEGIQAPSFSSYLYFLFCPTLIYR
ETYPRTPYVRWNYVAKNFAQALGCVLYACFILGRLCVPVFANMSREPFSTRALVLSILHATLPGIFMLLLIFFA
FLHCWLNAFAEMLRFGDRMFYRDWWNSTSFSNYYRTWNVVVHDWLYSYVYQDGLRLLGARARGVAMLGVFLVSA
VAHEYIFCFVLGFFYPVMLILFLVIGGMLNFMMHDQRTGPAWNVLMWTMLFLGQGIQVSLYCQ

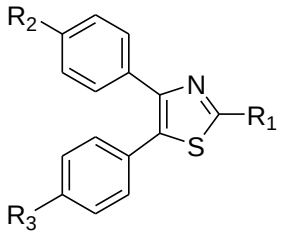
Figure 5: MBOAT sequence for ACAT-2 (amino acids 210 to 495).

A set of 32 ACAT inhibitors were selected and activity in IC_{50} (μM) were used for the development of the pharmacophore model. The $1/\log IC_{50}$ values (pIC_{50}) for ACAT inhibition was used as biological activity parameter in this study (**Table 1**).

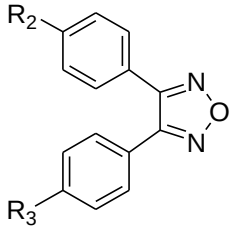
PHASE[160] employs the tree-based partitioning algorithm and thoroughly detects spatial arrangements of functional groups that are common and necessary for the biological activity for a set of high-affinity ligands. Activity data and a given hypothesis in combination create a 3D-QSAR model which is able to recognize significant aspects of molecular structures that govern

the activity. The developed pharmacophore model was additionally studied with an atom based 3-D QSAR model. In atom-based 3D-QSAR, a molecule is treated as a set of overlapping van der Waals spheres. Here van der Waals models of the aligned training set molecules were positioned in a regular grid of cubes, with each cube allotted zero or more 'bits' to account for the diverse types of atoms in the training set that occupy the cube. This illustration gives rise to binary-valued occupation patterns that have been used as independent variables to create partial least-squares (PLS) 3D-QSAR models. The data set with 32 molecules was divided into 26 molecules in the training set and 6 into the test set, and the 3D-QSAR models were developed. The developed models were evaluated by cross validated R^2,Q^2, SD, RMSE and Pearson-*R*.

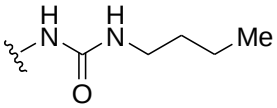
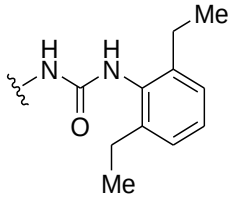
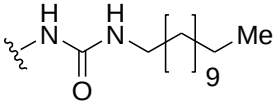
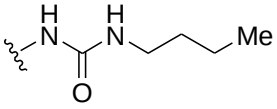
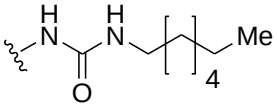
Table 1: Structures of ACAT inhibitors used for model development with actual and predicted activities.

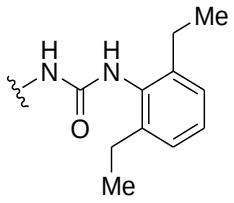
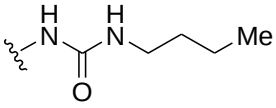
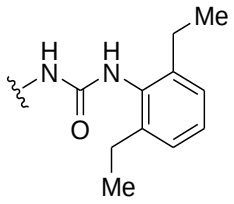
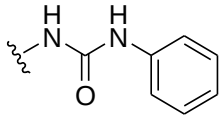
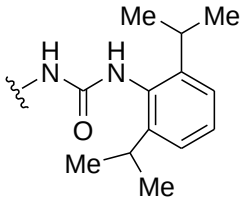
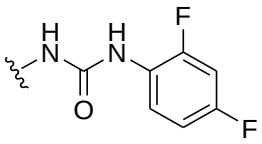
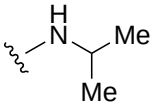
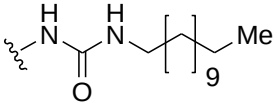
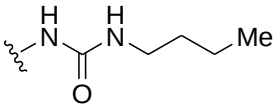
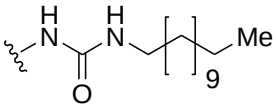
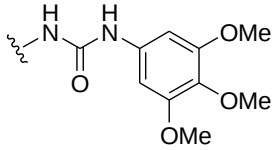
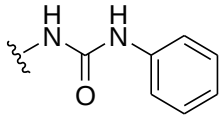


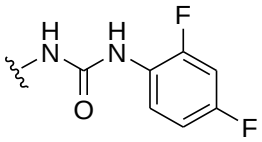
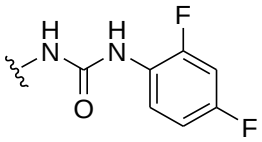
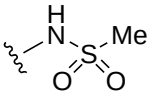
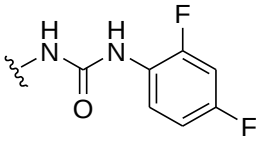
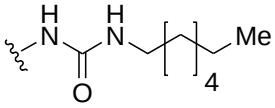
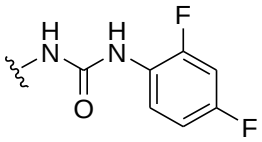
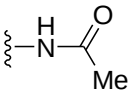
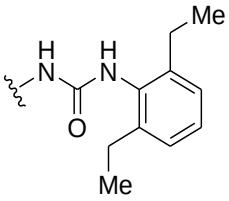
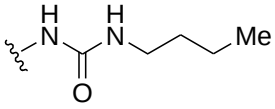
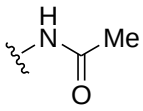
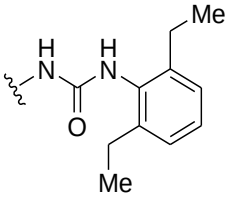
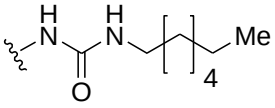
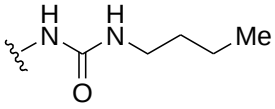
1-**7**, **9**-**12**, **15**-**18**, **20**-**30** and **32**

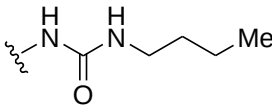
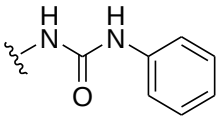
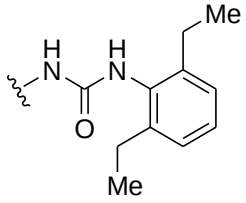
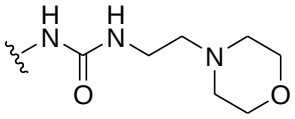
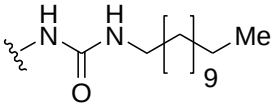
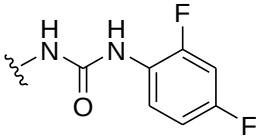


8, **13**, **14**, **19** and **31**

Compd	R_1	R_2	R_3	*p*IC_{50}		Residuals
				Actual	Predicted	
1	-Me	-Me		5.614	5.35	0.264
2	-Me	-Cl		5.393	5.41	-0.017
3	-Me	-F		5.244	5.09	0.154
4*	-Me	-Cl		5.202	5.37	-0.168
5	-Me	-Cl		5.105	5.08	0.025

6	-Me	-F		5.075	5.06	0.015
7	-Me	-F		5.018	5.25	-0.232
8*	---	-F		5.011	5.09	-0.079
9		-OMe	-Cl	4.972	4.99	-0.018
10	-Me	-F		4.92	4.97	-0.05
11		-OMe		4.916	4.97	-0.054
12	-Me	-Me		4.914	5.14	-0.226
13	---	-F		4.908	4.86	0.048
14	---	-F		4.905	4.89	0.015
15*	-Me	-F		4.89	4.72	0.17
16		-H	-H	4.751	4.72	0.031

17		-F	-Cl	4.735	4.69	0.045
18		-OMe		4.721	4.67	0.051
19	---	-F		4.697	4.66	0.037
20*		-OMe	-Cl	4.672	4.52	0.152
21		-OMe		4.496	4.52	-0.024
22*	-Me		-Me	4.412	4.3	0.112
23		-OMe		4.384	4.42	-0.036
24	-Me		-OMe	4.36	4.28	0.08
25	-Me	-F		4.329	4.4	-0.071
26		-OMe	-Cl	4.201	4.27	-0.069

27		-Me	-Cl	4.027	3.95	0.077
28*	-Me		-OMe	4.023	4.08	-0.057
29		-Me	$-NH_2$	3.986	3.99	-0.004
30	-Me	-Cl		3.927	3.84	0.087
31	---	-Me		3.909	3.9	0.009
32	-Me		-F	3.708	3.86	-0.152

*test set compounds.

By means of a tree-based partition algorithm, with a prerequisite that all five active compounds should match all the features, five featured common pharmacophore hypotheses were developed from different variants. These hypotheses generated by PHASE were assessed on the basis of 'Survival' and 'Survival-inactive' scores. The best model was found to be linked with the five-point hypotheses having a hydrogen bond acceptor group, a hydrogen bond donor group, a hydrophobic group and two aromatic rings (ADHRR). The inter-site distances and angles among the pharmacophoric features aligned on the best fit and the most active compound (**1**) are depicted in **Figure 6**. The data set of 32 compounds was divided into two subsets with 26 in training set and 6 in test set for the development of a 3D-QSAR model. For the distribution of dataset into training and test set, structural diversity over the full range of activity was considered and the model was developed. The graph for actual verses predicted activity of training set and test set is shown in **Figure 7**. To evaluate the predictive ability of the model, the test set was used. The model was developed using PLS with five factors. Based on the R^2,Q^2, SD, RMSE and

Pearson-*R* values, the developed 3D-QSAR model was found to be accurately developed and these results are shown in **Table 2**.

Table 2: Statistical results for the common pharmacophore feature and 3D-QSAR model generated by PLS.

Training set	Test set
PLS = 5	
n = 26	n(test) = 06
R^2 = 0.9551	Q^2 = 0.8946
SD = 0.1170	RMSE = 0.1276
F = 85.1	Pearson-*R* = 0.9612
P = 8.84e-013	

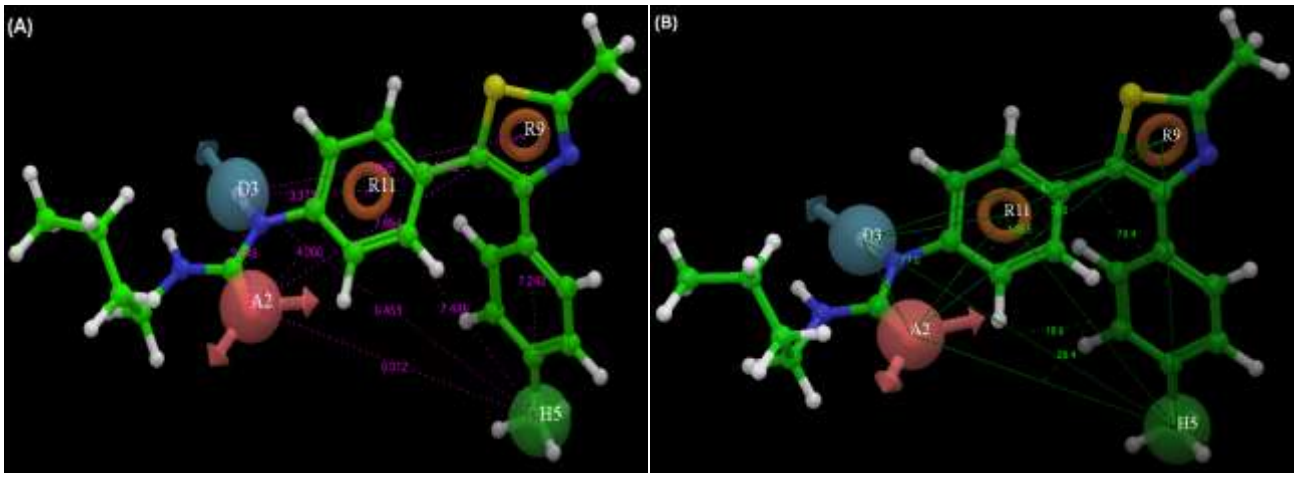


Figure 6: The best pharmacophore model ADHRR aligned on the best fit and the most active compound (**1**) with **(A)** intersite distance in Å; **(B)** angles between pharmacophore features.

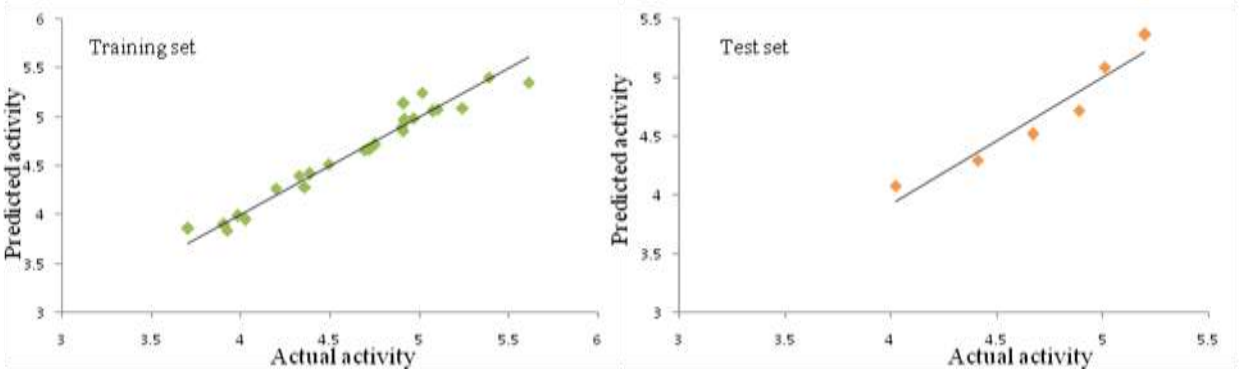


Figure 7: Scatter plot of the predicted pIC_{50} against the experimental pIC_{50} for training and test set compounds.

The volume occlusion maps obtained from 3D-QSAR recognized the important features necessary for the ligand-receptor interactions. The occlusion maps are represented by blue and red colors indicating favorable and unfavorable interactions respectively. The occlusion maps (donor, hydrophobic and electronegative) with the most active compound (**1**) for the developed 3D-QSAR model are shown in **Figure 8**. As observed in **Figure 8A** for the most active compound **1**, the blue regions are observed near the NH group of urea linker, explaining its high activity because of hydrogen bond donating capability. Whereas, lesser activity of compounds (**30** and **32**) were explained very well as the NH groups were overlapping the non-favorable i.e. red region.

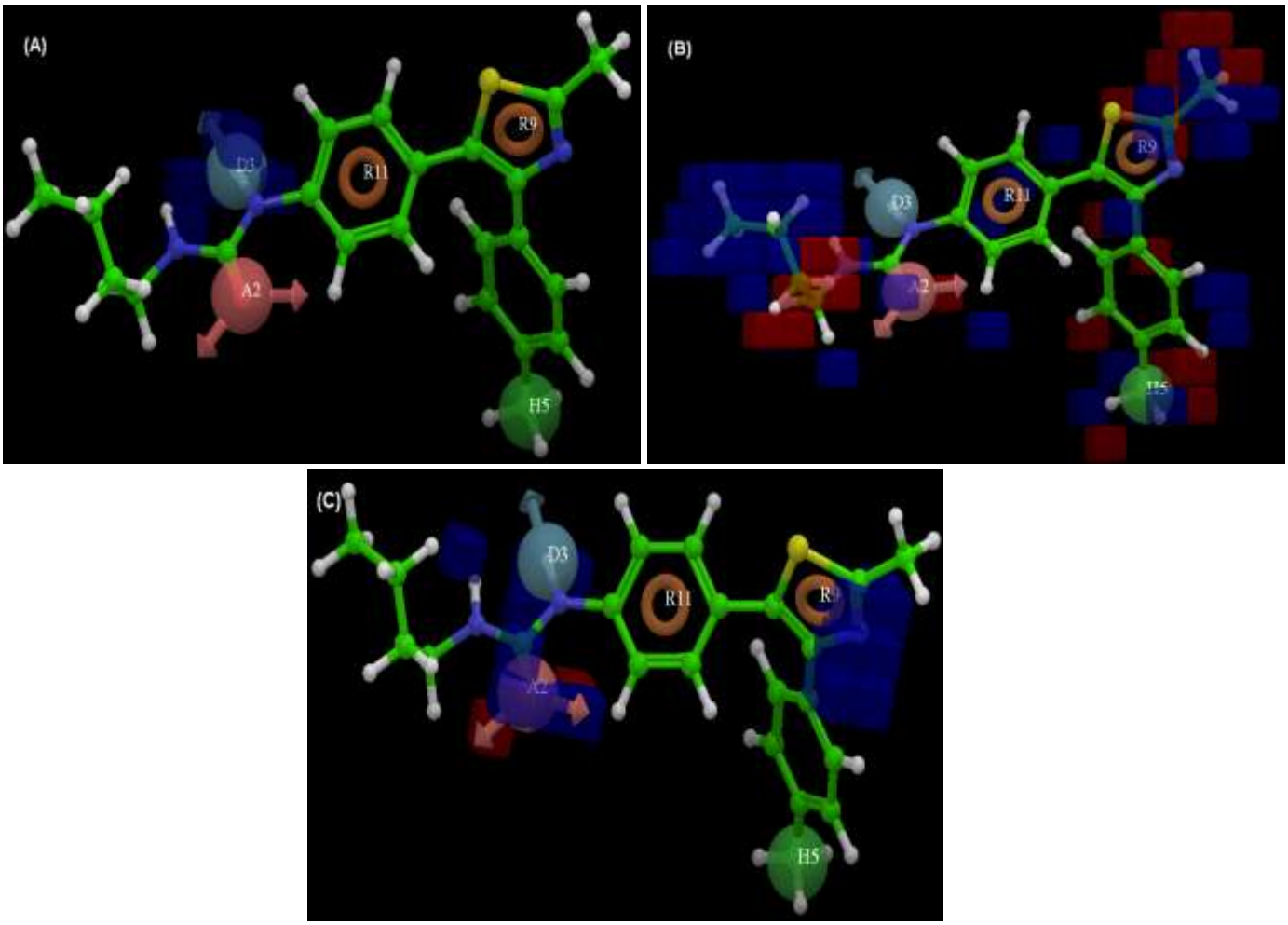


Figure 8: 3D-QSAR visualization for **1**; H-bond donor (**A**), Hydrophobic (**B**) and Electron withdrawing features (**C**).

Figure 8B shows the hydrophobic volume occlusion map. Here the high activity of **1** was explained pretty well, as the methyl, alkyl and aromatic groups were in blue region explaining the possible hydrophobic interactions, while the urea group was in the red region explaining hydrophilic interactions. Lower activity of compound (**28**) was also explained as most of the hydrophobic groups of the ligand were oriented to the red regions. **Figure 8c** illustrates the electronegative maps; the high activity of **1** was validated by the presence of carbonyl and thiazole ring N in the blue regions.

Results of the docking studies further validated the features of the generated pharmacophore model of ACAT inhibitors that are responsible for their activity. The binding modes of the most active compound (**1**) with ACAT-1 and ACAT-2 are shown in **Figure 9** and **Figure 10** respectively. The active compound (**1**) into the proposed active site of both the isoforms show almost similar affinity. The docking score for compound (**1**) with ACAT-1 was -6.15, whereas with ACAT-2 it was observed to be -6.67. In ACAT-1 the most active ligand compound **(1**) interacts mostly by van der Waal or hydrophobic interactions. The thiazole ring forms π-π stacking with His460, whereas the two aromatic rings show hydrophobic interactions with Asn421, Phe453 and Trp499. In addition to these interactions the butyl chain fits into the hydrophobic pocket of Leu304, Ile377, Thr380, Phe381 and Tyr417. While in case of ACAT-2 the proposed active site residue His343 showed strong π-π interaction with *p*-methyl phenyl ring attached to the thiazole ring. Further, the thiazole and the two aromatic rings attached to the thiazole were observed to be stabilized by hydrophobic interactions with Val406, Asp409, Phe427, Ile456, Leu460, Met463, Gln467 and Trp473. The butyl chain was strongly stabilized by hydrophobic part of Leu275, Ile354, Tyr391, Trp394 and Val398.

This pharmacophore based 3D-QSAR study and the docking study based on the proposed 3D models of ACAT-1 and ACAT-2 could be fruitfully utilized for the development of potent novel ACAT inhibitors.

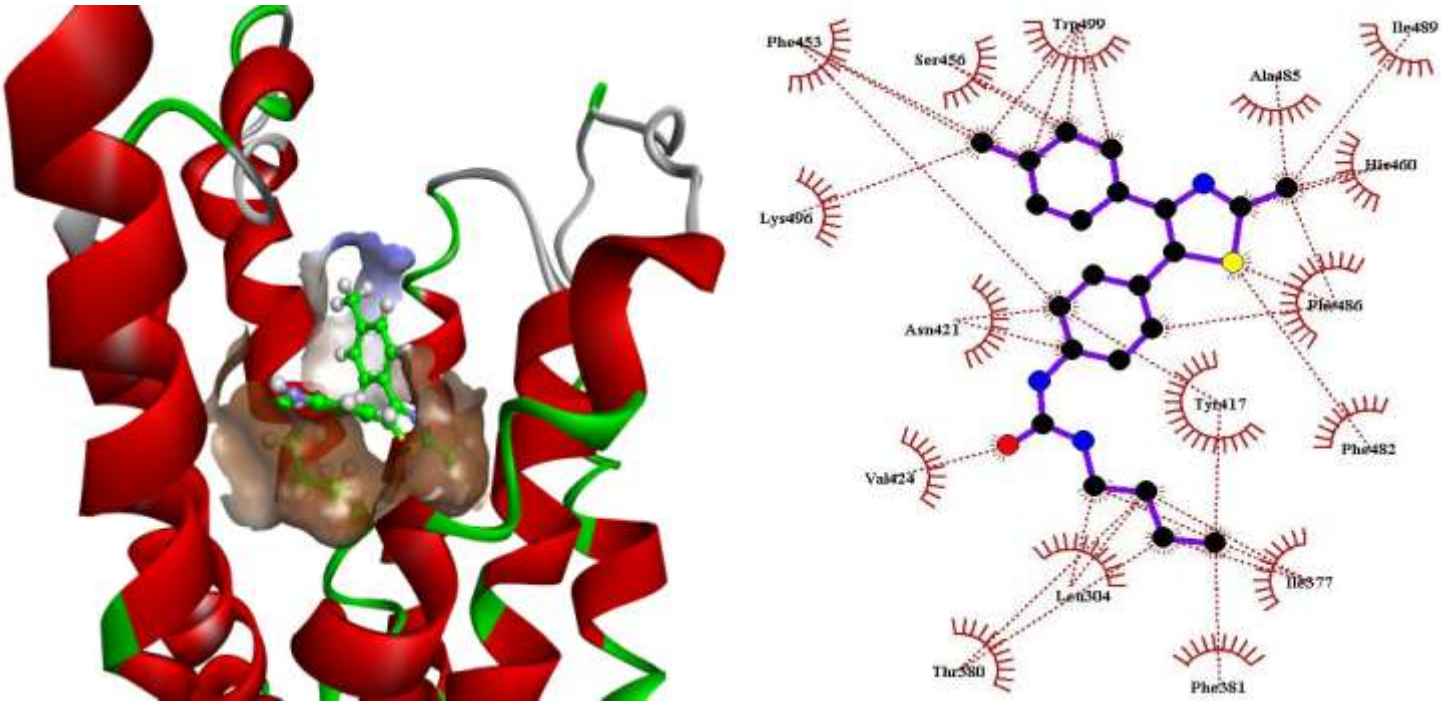


Figure 9: Docking of the most active compound (**1**) in the active site of the proposed ACAT-1 model.

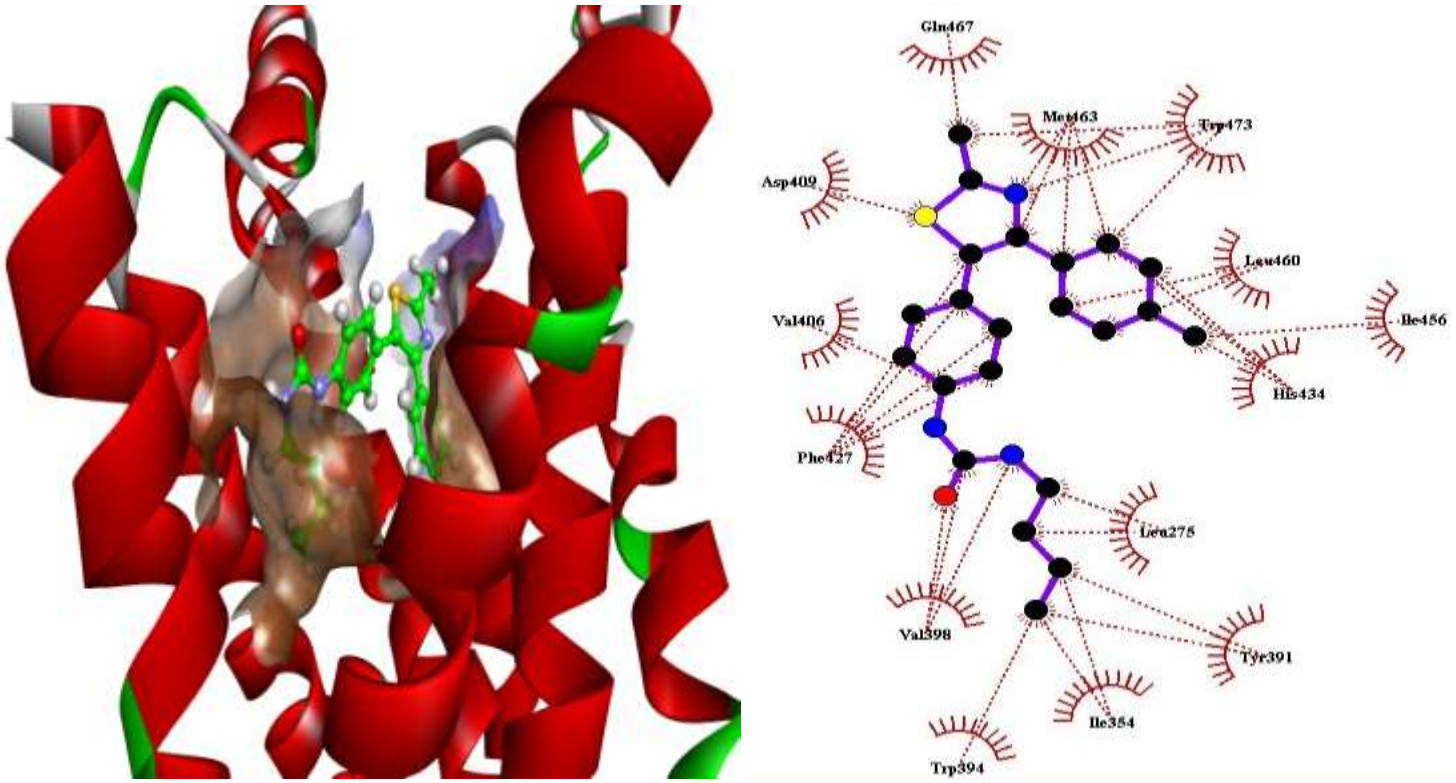


Figure 10: Docking of the most active compound (**1**) in the active site the of proposed ACAT-2 model.