International Journal of Research and Reviews in Pharmacy and Applied science

www.ijrrpas.com

A Quantitative Structure-Activity Relationship and Molecular Docking Study on a Series of Indole-5-carboxamides Acting as Anti-hepatitis C Virus Agents

Sakshi Gupta*
Department of Pharmaceutical Technology.

Neeraj Agarwal
Department of Biotechnology, Meerut Institute of Engineering and Technology, Meerut 250005,U.P., India

ABSTRACT

A quantitative structure-activity relationship (QSAR) and molecular docking study has been performed on a series of indole-5-carboxamides acting as hepatitis C virus inhibitors. In this case a significant correlation has been found between the anti-HCV potencies and some physicochemical properties of the compounds, indicating that HCV inhibition potency of this series of compounds is controlled by hydrophobicity, polarizability, 3D Wiener index of the molecules and the number of nitrogen atoms present in them and that the 6-membered heterocyclic ring present in the molecule has added advantage. The validity of the correlation has been judged by leave-one-out jackknife procedure and predicting the activity of some test compounds. Using the correlations obtained, some new compounds of high potency have been predicted in each series. A docking study using Molegro Virtual Docker has been performed on these predicted compounds to find their interactions with the receptor. It is observed that all the predicted compounds have better interaction energy and the docking score than the ligand complexed with the protein.

Keywords: Anti-hepatitis C virus agents, QSAR study, Indole-5-carboxamides, Docking.

INTRODUCTION

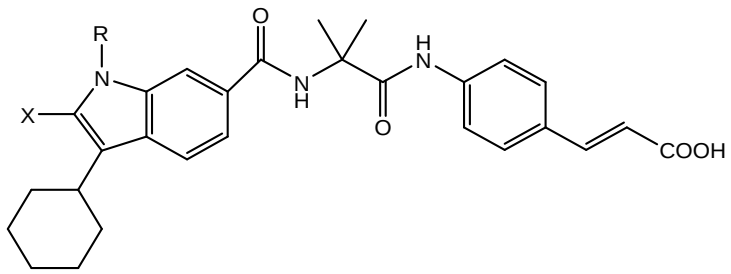
Hepatitis C virus (HCV) is a single, enveloped positive sense RNA virus belonging to family of *Flaviridae* and having the size between 55-65 nm. It is the major cause of blood borne diseases such as non-A, non-B hepatitis, and most commonly hepatitis C in humans affecting over 170 million people throughout the world. The length of HCV genome is approximately 9.6 kb with a polyprotein of about 3,000 amino acids encoded in it. The host and HCV-encoded proteases are responsible for its cleavage into 10 structural and non-structural components. [1-3]

HCV virus is classified into seven genotypes 1-7 based on genetic differences among different HCV species. Each genotype is further subdivided into various subtypes, denoted by lower case letters such as 1a, 1b etc. The genomic compositions of various subtypes of a particular genotype differ generally by 20-25%. 60% of all cases of HCV virus infection are caused by subtypes 1a and 1b which are found throughout the world. The genome of HCV virus is evolved rapidly due to the errors created during the replication of RNA dependent RNA polymerase. [4-5]

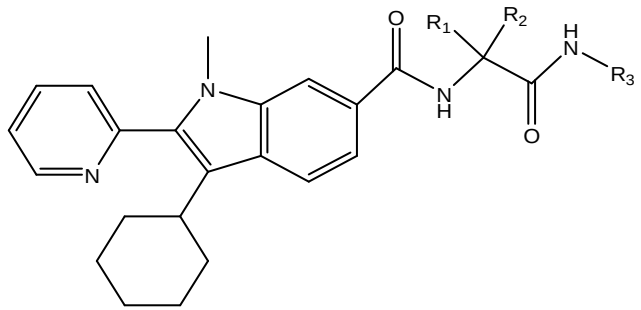
Chronic HCV infections are also responsible for many liver diseases like hepatitis, liver cirrhosis, liver fibrosis, hepatocellular carcinoma, and other forms of liver malfunctioning, which ultimately develop into liver cancer if not treated in time. The disease caused due to HCV virus is having widespread impact on people; therefore there is a substantial medical need to develop novel and potent anti-HCV agents so as to compliment the already existing therapies. The current therapy that is available until now includes interferon-α (INF-α) and ribavirin, but the overall chances of curing the disease in this case is only up to 50% and the undesirable side effects are also associated with it [6-9]. In the present communication, we have focused on developing the novel inhibitors of HCV NS5B RNA-dependent RNA polymerase with improved potency and lesser side effects with the help of quantitative structure-activity relationship (QSAR) studies.

MATERIALS and METHODS

We have taken two similar series of indole-5-carboxamide inhibitors (**1**, **2**) that were synthesized and evaluated for their anti-HCV activity by Beaulieu *et al.* in their two successive studies [10, 11]. A combine of these series is listed in Table **1** along with the anti-HCV activities of the compounds. This table also lists the physicochemical parameters of the compounds that were found to govern their potency. The physicochemical parameters used have been calculated by using Chemdraw 2004, Chemsketch version 11.0 and e-Dragon software [12]. Several parameters were calculated, but the parameters that were found to be important were calculated hydrophobicity (ClogP), polarizability (Pol), 3D Weiner index (W3D), and number of nitrogen atoms in the molecule (n_N). Other parameters calculated were not found to be significant. An indicator parameter 'I' has also been used that refers to the presence of a 6-membered heterocyclic ring in X-substituent. It has been given a value of 1 if X-substituent has a 6-membered heterocyclic ring otherwise its value is zero. In activity term log ($1/IC_{50}$), IC_{50} refers to molar concentration of the compound leading to 50% inhibition of HCV RNA replication.

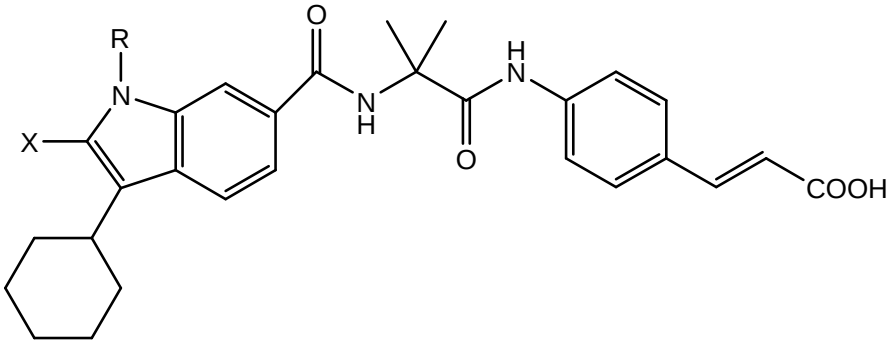


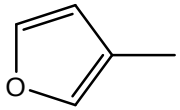
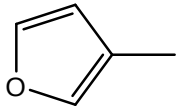
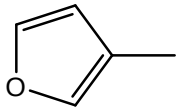
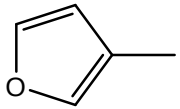
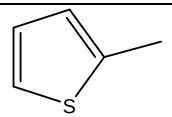
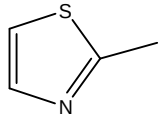
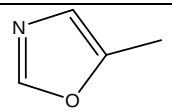
(**1**)

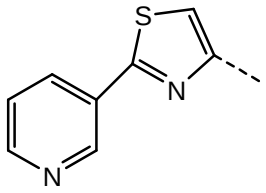
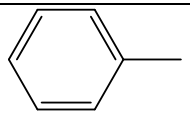
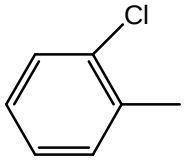
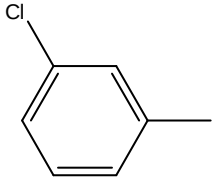
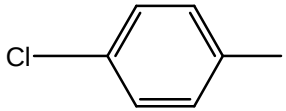
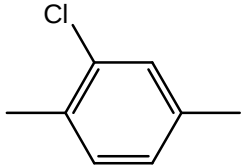
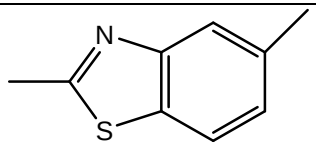


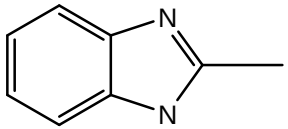
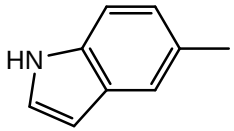
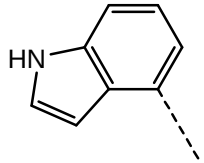
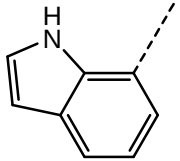
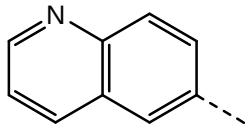
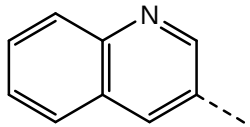
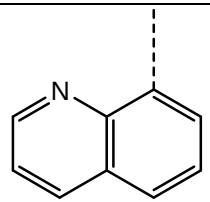
(**2**)

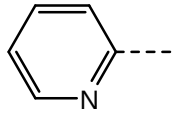
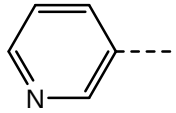
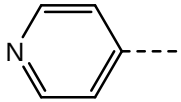
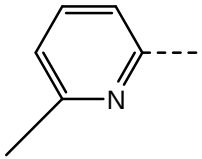
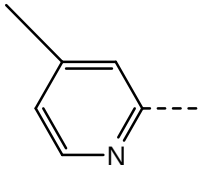
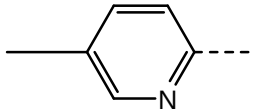
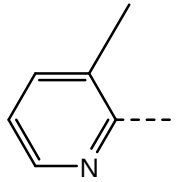
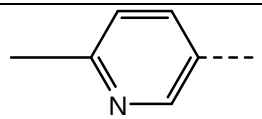
Table 1. Indole-5-carboxamide Derivatives and Their Anti-HCV Activity and Physicochemical Parameters

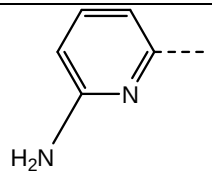
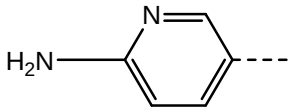
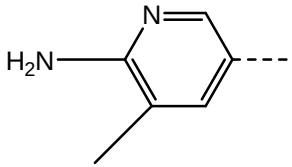
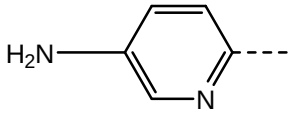
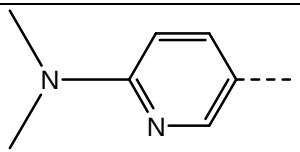
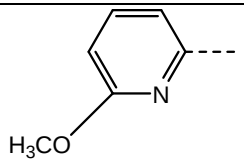
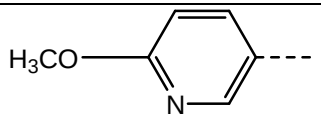
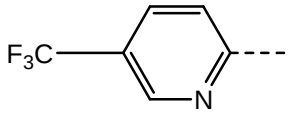


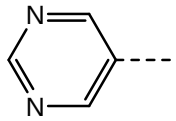
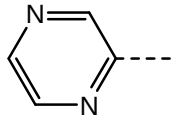
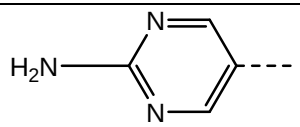
Comp.	R	X	ClogP	W3D	Pol	I	n_N	log(1/IC_{50})		
								Obsd[a]	Cald, Eq.(1)	Pred.(LOO)
1	H		6.35	13.14	61.61	0.00	3.00	6.76	6.95	7.02
2[b]	CH_3		6.84	14.05	62.13	0.00	3.00	6.97	6.71	-
3	C_2H_5		7.36	15.76	63.96	0.00	3.00	6.55	6.41	6.36
4[b]	*i*Bu		8.29	18.15	67.54	0.00	3.00	6.20	6.08	-
5	CH_3		7.52	14.62	64.67	0.00	3.00	6.74	6.72	7.01
6[c]	CH_3		6.23	14.22	64.06	0.00	4.00	7.34	6.77	-
7[b]	CH_3		5.57	14.08	60.86	0.00	4.00	7.47	6.58	-

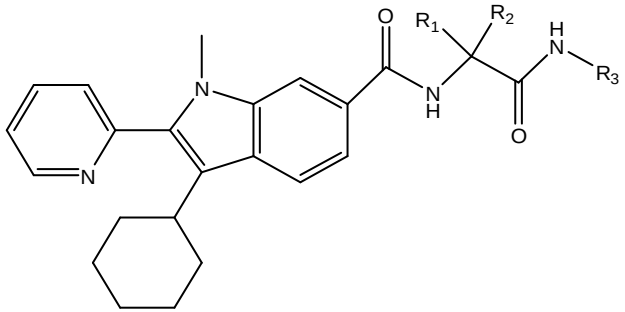
8	CH_3		6.92	19.11	73.39	1.00	5.00	6.57	6.97	7.05
9	CH_3		7.66	16.12	65.23	0.00	3.00	6.46	6.41	6.44
10	CH_3		8.13	15.99	67.06	0.00	3.00	6.38	6.55	6.58
11	CH_3		8.38	15.89	67.06	0.00	3.00	6.44	6.52	6.55
12[b]	CH_3		8.38	16.02	67.06	0.00	3.00	6.64	6.50	-
13	CH_3		8.88	17.73	68.81	0.00	3.00	6.52	6.19	6.13
14	CH_3		8.10	18.52	72.11	0.00	4.00	6.16	6.30	6.31

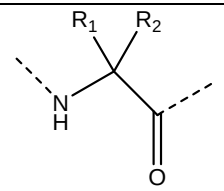
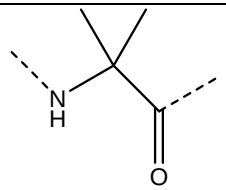
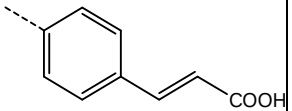
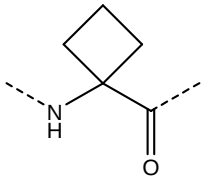
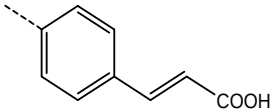
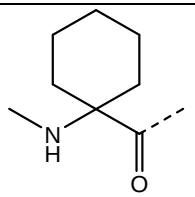
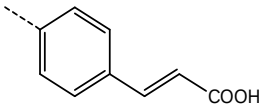
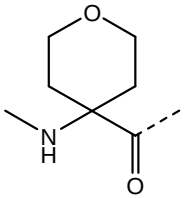
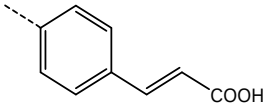
15[b]	CH_3		5.73	17.04	68.47	0.00	5.00	6.01	6.48	-
16	CH_3		7.65	17.74	69.09	0.00	4.00	6.38	6.23	6.23
17	CH_3		7.65	17.34	69.09	0.00	4.00	6.30	6.33	6.34
18	CH_3		7.65	17.23	69.09	0.00	4.00	6.36	6.35	6.20
19[b]	CH_3		7.51	18.48	70.91	0.00	4.00	6.40	6.30	-
20	CH_3		5.40	18.36	70.91	0.00	4.00	6.54	6.75	6.77
21	CH_3		7.71	17.68	70.91	0.00	4.00	6.57	6.44	6.45

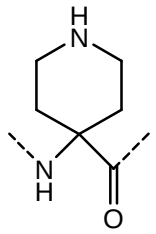
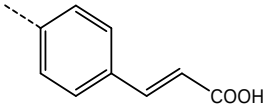
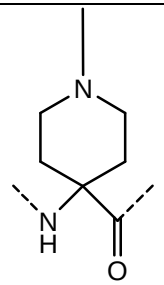
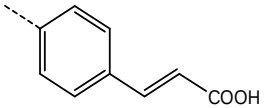
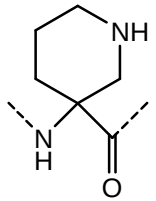
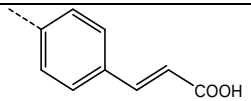
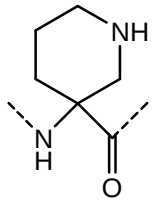
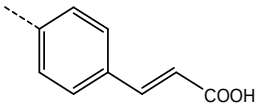
22	CH_3		6.39	15.58	64.62	1.00	4.00	7.44	7.14	7.15
23	CH_3		6.18	15.57	64.62	1.00	4.00	6.96	7.19	7.17
24	CH_3		6.18	15.42	64.62	1.00	4.00	6.96	7.22	7.24
25	CH_3		6.89	17.06	66.37	1.00	4.00	7.21	6.89	6.84
26[c]	CH_3		6.89	17.10	66.37	1.00	4.00	6.34	6.89	-
27	CH_3		6.89	17.34	66.37	1.00	4.00	7.15	6.83	6.81
28[b]	CH_3		6.59	16.86	66.37	1.00	4.00	6.00	7.01	-
29	CH_3		6.68	17.08	66.37	1.00	4.00	7.08	6.93	6.91

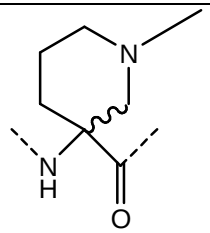
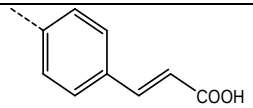
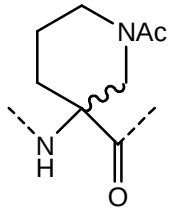
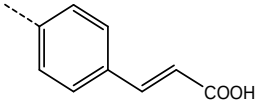
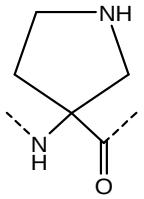
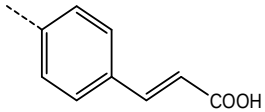
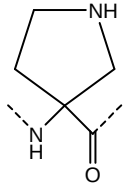
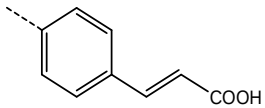
30	CH_3		6.07	16.72	65.37	1.00	5.00	7.12	6.79	6.80
31	CH_3		5.86	16.62	65.37	1.00	5.00	6.75	6.85	6.85
32	CH_3		6.30	18.12	67.13	1.00	5.00	6.20	6.62	6.68
33[b]	CH_3		6.07	16.78	65.37	1.00	5.00	7.74	6.78	-
34	CH_3		6.89	20.38	69.69	1.00	5.00	6.06	6.26	6.31
35	CH_3		7.21	17.81	66.92	1.00	4.00	7.10	6.72	6.67
36	CH_3		7.00	17.75	66.92	1.00	4.00	6.50	6.77	6.78
37	CH_3		7.42	17.36	66.50	1.00	4.00	6.80	6.73	6.74

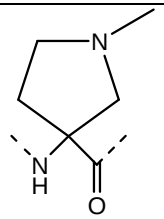
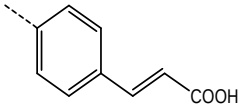
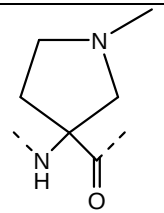
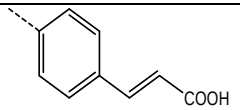
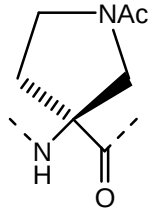
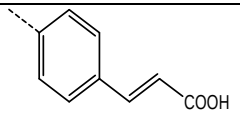
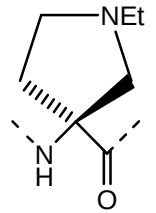
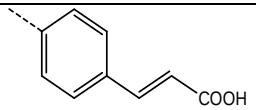
38	CH_3		5.22	14.93	64.00	1.00	5.00	7.39	7.22	7.24
39[c]	CH_3		5.43	15.24	64.00	1.00	5.00	7.70	7.11	-
40	CH_3		5.21	16.20	64.76	1.00	6.00	6.82	6.78	6.78

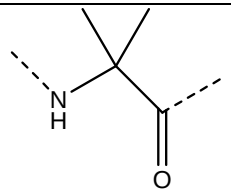
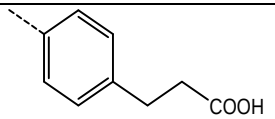
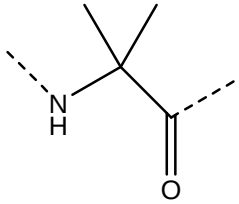
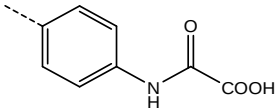
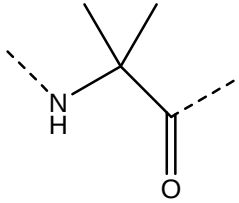
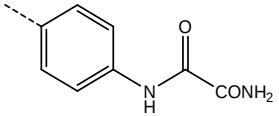
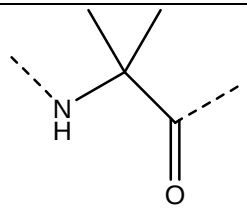
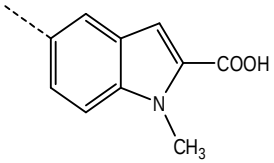


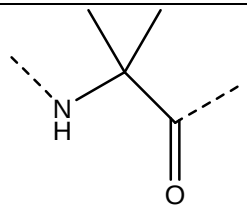
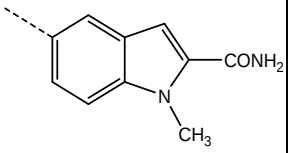
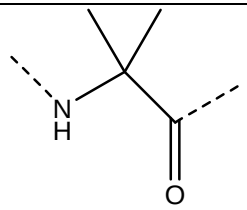
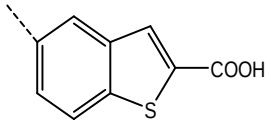
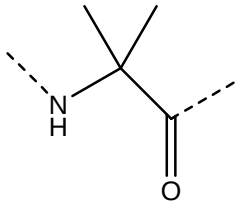
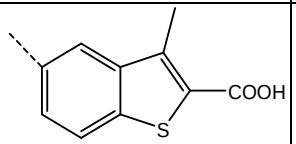
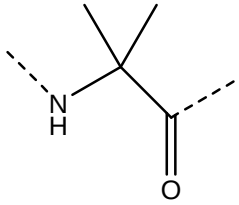
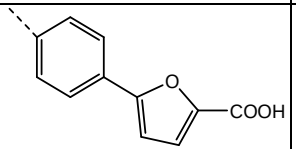
Compd		R_3	ClogP	W3D	Pol	I	n_N	log(1/IC_{50})		
								Obs.	Cald (Eq.1)	Pred. (LOO)
41[b]			6.39	15.77	64.62	1.00	4.00	7.44	7.10	-
42			6.59	15.34	65.58	1.00	4.00	7.28	7.26	7.27
43[b]			7.70	17.29	69.23	1.00	4.00	6.71	7.01	-
44			5.30	17.24	67.96	1.00	4.00	7.17	7.36	7.39

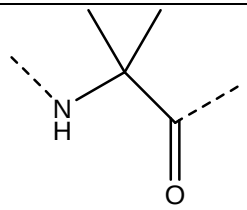
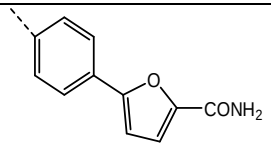
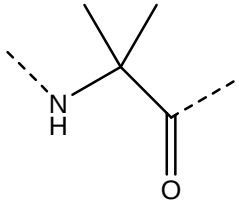
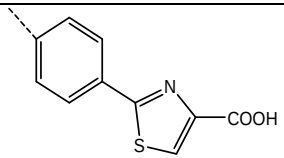
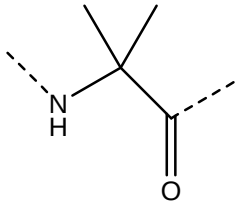
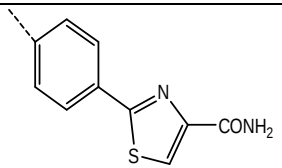
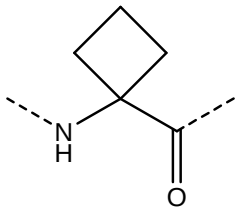
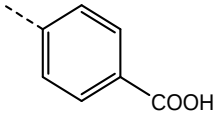
45			3.01	17.02	68.62	1.00	5.00	7.89	7.71	7.65
46[c]			3.46	19.24	70.73	1.00	5.00	8.00	7.35	-
47			3.80	17.77	68.62	1.00	5.00	7.40	7.37	7.38
48[c]			3.80	17.77	68.62	1.00	5.00	7.96	7.38	-

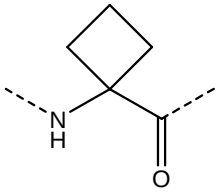
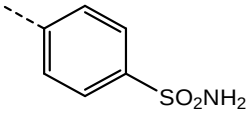
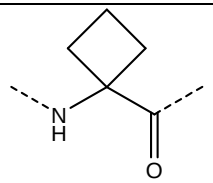
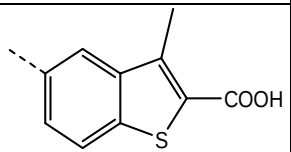
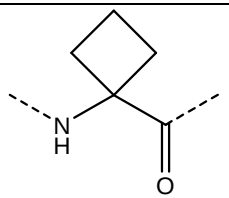
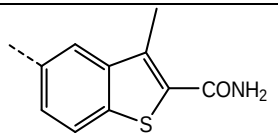
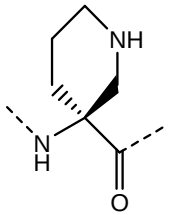
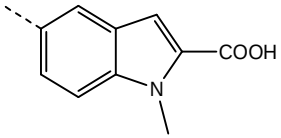
49			4.35	18.70	70.73	1.00	5.00	7.37	7.29	7.30
50			5.54	19.66	72.89	1.00	5.00	6.64	7.07	7.11
51			3.24	16.46	66.79	1.00	5.00	7.26	7.58	7.64
52			3.24	16.46	66.79	1.00	5.00	7.64	7.58	7.58

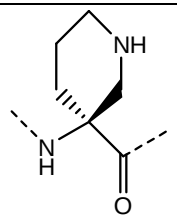
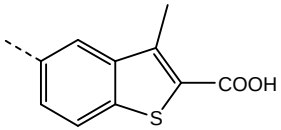
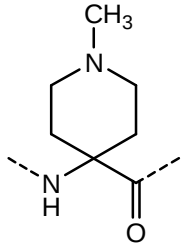
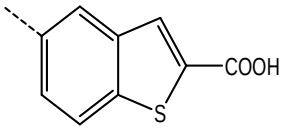
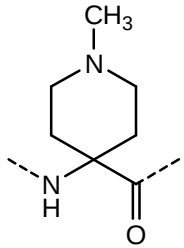
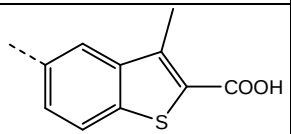
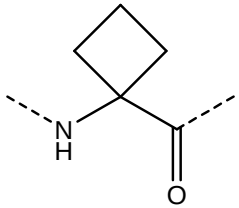
53[b]			3.79	17.65	68.90	1.00	5.00	7.18	7.44	-
54[b]			3.79	17.65	68.90	1.00	5.00	7.68	7.44	-
55			4.98	18.75	71.07	1.00	5.00	7.62	7.18	7.14
56			4.32	18.69	70.73	1.00	5.00	7.64	7.29	7.25

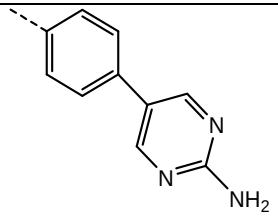
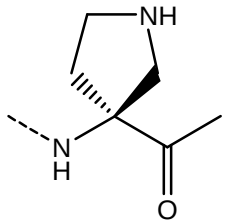
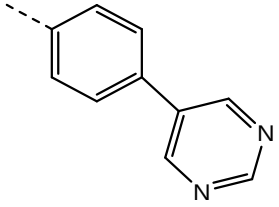
57			6.05	16.13	64.62	1.00	4.00	6.88	7.08	7.09
58			4.07	15.29	64.34	1.00	5.00	7.47	7.41	7.38
59			6.84	15.46	66.93	1.00	6.00	6.91	6.86	6.88
60			3.43	16.52	64.76	1.00	5.00	7.12	7.30	7.32

61			5.97	17.54	67.35	1.00	6.00	6.78	6.61	6.56
62			7.42	14.80	66.70	1.00	4.00	7.64	7.35	7.32
63			7.62	16.37	68.45	1.00	4.00	7.45	7.14	7.12
64			7.37	17.07	67.81	1.00	4.00	6.82	6.96	6.94

65			6.14	17.44	68.23	1.00	5.00	6.48	6.93	6.97
66[b]			7.47	16.79	69.74	1.00	5.00	6.81	6.99	-
67			6.42	17.45	70.16	1.00	6.00	6.50	6.85	6.92
68			6.53	13.15	61.92	1.00	4.00	7.03	7.37	7.45

69[b]			4.97	14.73	64.05	1.00	5.00	6.50	7.33	-
70			7.81	16.98	69.42	1.00	4.00	6.91	7.07	7.08
71			6.59	17.28	69.84	1.00	5.00	7.06	7.06	7.06
72			4.25	18.38	70.93	1.00	6.00	7.36	7.17	7.18

73			5.03	18.26	72.46	1.00	5.00	7.45	7.44	7.46
74			4.49	18.55	72.81	1.00	5.00	7.77	7.53	7.47
75			4.69	18.82	74.57	1.00	5.00	7.57	7.62	7.64
76[b]			6.13	17.15	68.90	1.00	7.00	6.13	6.61	-

										
77			5.07	16.51	69.35	1.00	7.00	7.19	7.01	6.94

[a]For compounds 1-40 taken from ref [10] and for 41-77 from ref [11]. [b]Used for test set. [c]Not included in the derivation of Eq. (1).

RESULTS & DISCUSSION

QSAR Results

There were in total 77 compounds in Table 1. We have divided them into two subsets, the training set containing 62 compounds and the test set containing 15 compounds. The test set is selected randomly by keeping in mind the wide structural diversity and span in the activity data. The compounds of test set are indicated in bold and with superscript 'b' in the table. When a multiple regression analysis was performed on the training set, it revealed the following correlation.

$$\log(1/IC_{50}) = -0.203(\pm0.060)\text{ClogP} - 0.231(\pm0.086)\text{W3D} + 0.113(\pm0.044)\text{Pol} + 0.657(\pm0.215)\text{I} - 0.238(\pm0.115)n_{\mathbf{N}} + 5.035(\pm1.937)$$

$$n = 57,\ r = 0.879,\ r^2_{cv} = 0.685,\ r^2_{pred.} = 0.612,\ s = 0.25,\ F_{5,51} = 28.70(3.40) \qquad (1)$$

In this equation, n is the number of data points, r is the correlation coefficient, r^2_{cv} is the square of cross-validated correlation coefficient obtained from leave-one-out (LOO) jackknife procedure, r^2_{pred} is the square of correlation coefficient showing the predictive ability of the correlation, s is the standard deviation, data within the parentheses with ± sign are 95% confidence intervals, and the F is F-ratio between the variances of calculated and observed activities. Equation 1 represents a highly significant correlation between the HCV inhibition potency of the compounds and the physicochemical parameters. In this equation, the positive coefficient of Pol suggests that the activity of the compounds will increase as its polarizability increases. However, the negative coefficients of ClogP and W3D indicate that the increase in the hydrophobic property of the molecules and in their topology measured by Wiener index will not be conducive to the activity. Thus, it seems that the polarizability of the molecule is the most favorable factor for the potency of the compound and thus dependence of the activity on polarizability leads to suggest that there is strong electrostatic interaction between the compounds and the receptor.

A positive coefficient of the indicator variable 'I' also indicated that the X-substituent being or containing a 6-membered heterocyclic ring will also be favorable to the activity. The negative coefficient of n_N indicates that the more no. of nitrogen atoms may have negative effect on the potency of compound. The lone pair of nitrogen might experience some electronic repulsion with active sites of the receptor.

All the variables used in Eq. (1) were found to have no mutual correlation (**Table 2**) and each one of them was found to be statistically quite significant as an appreciable drop in the overall significance in the successive equation was observed when each one of them was dropped one by one (Eqs. 2-5).

Table 2. Correlation Matrix Showing Mutual Correlations among the Variables Used in Eq. (1)

	ClogP	W3D	Pol	I	n_N
ClogP	1.000	-0.019	-0.034	0.211	0.424
W3D		1.000	-0.823	-0.193	0.042
Pol			1.000	0.286	-0.265
I				1.000	-0.488
n_N					1.000

$$\log(1/IC_{50}) = -0.151(\pm0.062)\text{ClogP} - 0.224(\pm0.098)\text{W3D} + 0.089(\pm0.049)\text{Pol} + 0.441(\pm0.215)\text{I} + 5.322(\pm2.211)$$

$$n = 57,\ r = 0.836,\ s = 0.28,\ F_{4,52} = 24.11(3.70) \quad (2)$$

$$\log(1/IC_{50}) = -0.218(\pm0.060)\text{ClogP} - 0.184(\pm0.109)\text{W3D} + 0.070(\pm0.054)\text{Pol}$$

$$+ 6.666(\pm 2.408)$$

$$n = 57,\ r = 0.767,\ s = 0.32,\ F_{3,53} = 20.35(4.17) \quad (3)$$

$$\log(1/IC_{50}) = -0.217(\pm 0.063)\text{ClogP} - 0.066(\pm 0.063)\text{W3D} + 9.416(\pm 1.184)$$

$$n = 57,\ r = 0.721,\ s = 0.34,\ F_{2,54} = 24.57(5.01) \quad (4)$$

$$\log(1/IC_{50}) = -0.209(\pm 0.065)\text{ClogP} + 8.247(\pm 0.409)$$

$$n = 57,\ r = 0.679,\ s = 0.35,\ F_{1,55} = 42.11(7.12) \quad (5)$$

Thus the correlation expressed by Eq. (1) seems to be highly significant and its internal and external validations can be judged by r^2_{cv} and r^2_{pred} values, which are 0.68 and 0.61, respectively.

The r^2_{cv} is calculated as follows.

$$r^2_{cv} = 1 - [\sum_i (Y_{i,obsd} - Y_{i,pred})^2 / \sum_i (Y_{i,obsd} - Y_{av,obsd})^2] \quad (6)$$

where $Y_{i,obsd}$ and $Y_{i,pred}$ are the observed and predicted (from LOO) activity values of compound *i*, respectively, and $Y_{av,obsd}$ the average of the observed activities of all compounds used in the correlation. Similarly, the r^2_{pred} is calculated as

$$r^2_{pred} = 1 - [\sum_i (Y_{i,obsd} - Y_{i,pred})^2 / \sum_i (Y_{i,obsd} - Y_{av,obsd})^2] \quad (7)$$

where $Y_{i,obsd}$ is the observed activity of compound *i* in the test set and $Y_{i,pred}$ is its activity predicted from Eq. (1). $Y_{av,obsd}$ is the same as in Eq. (6).

The activity values predicted from Eq. (1) for the test set compounds are given (in bold) in Table 1. A comparison shows that these predicted values are in fairly good agreement with the corresponding observed ones. In the training set also the predicted values of compounds are found to be in good agreement with their observed ones. All these observations can be better visualized in the graph drawn between the predicted and observed activities for training as well as test sets (**Fig. 1**).

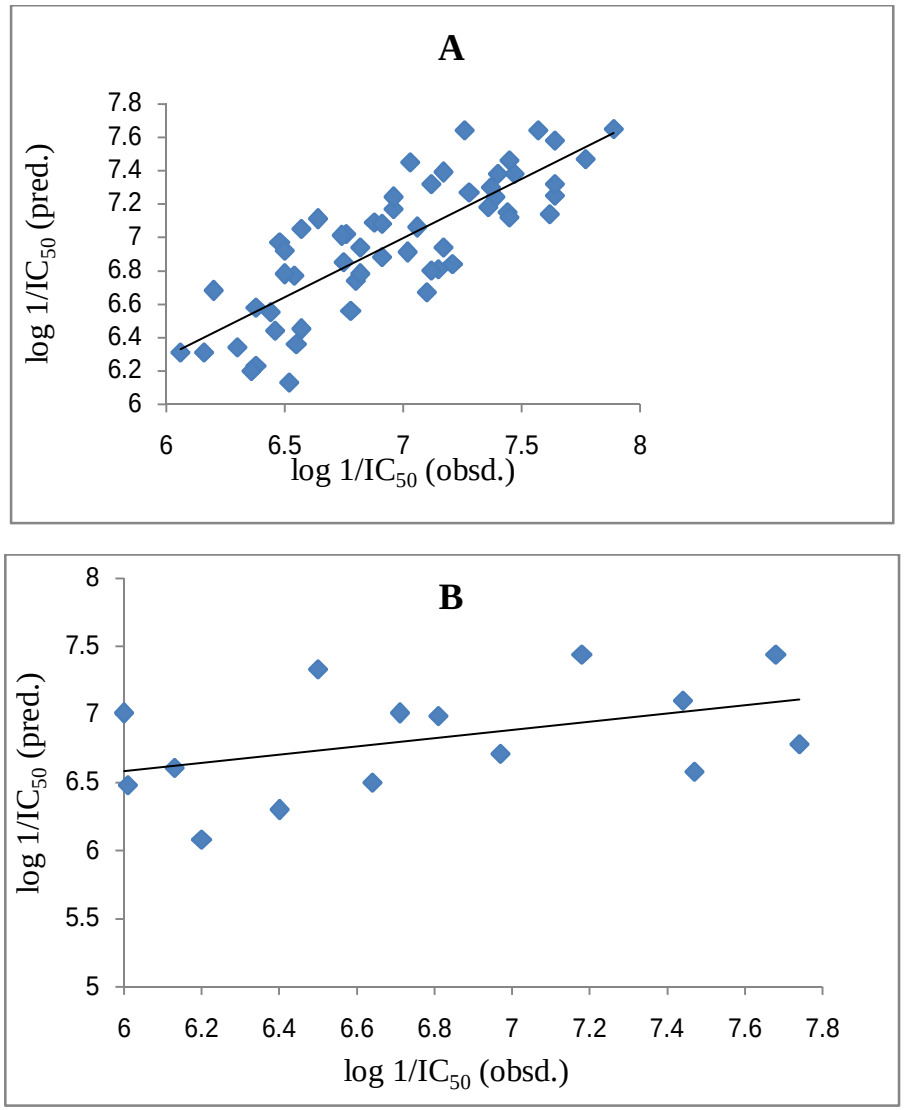
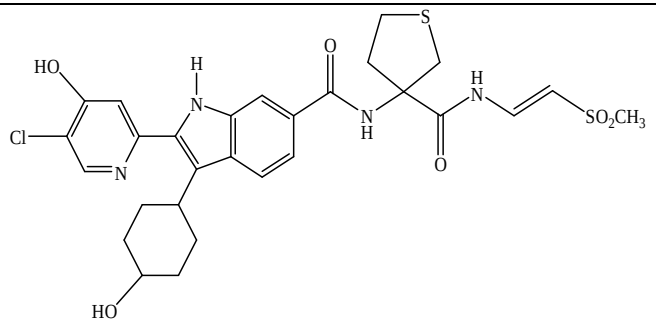


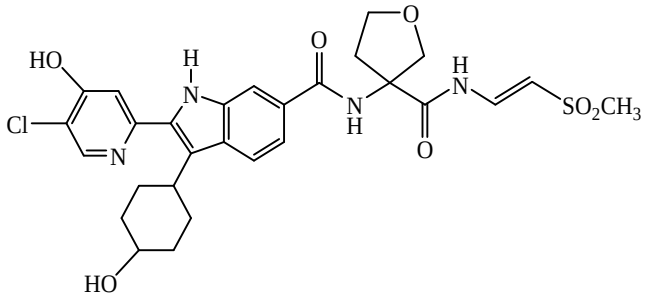
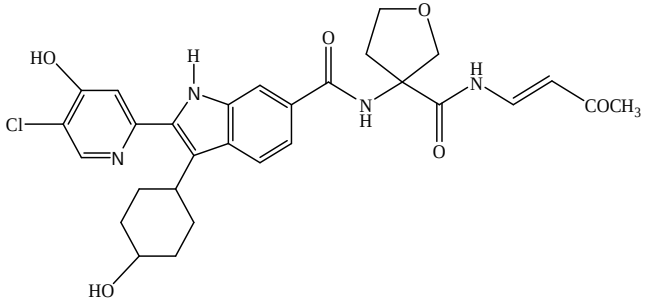
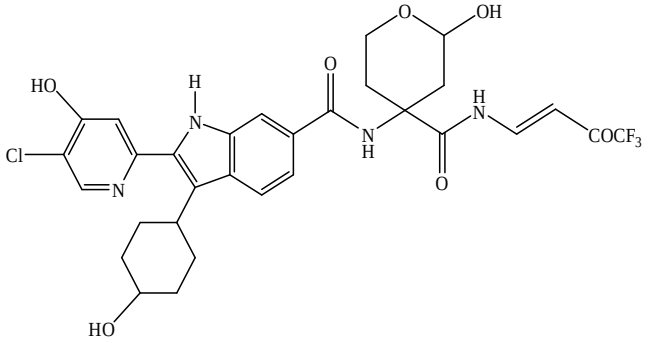
Fig. (1). Plots between observed and predicted activities for indole-5-carboxamides: A, for training set; B, for test set.

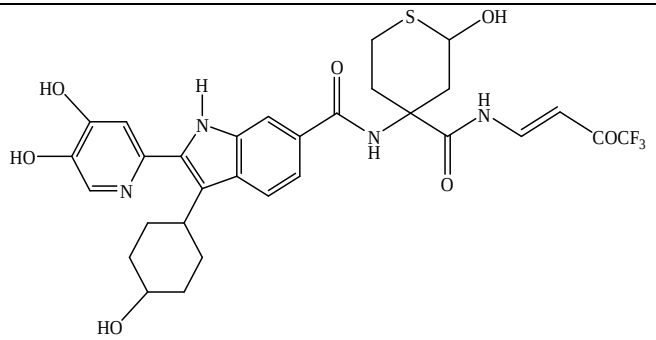
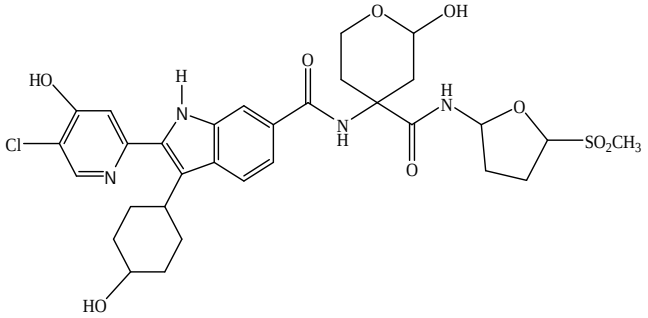
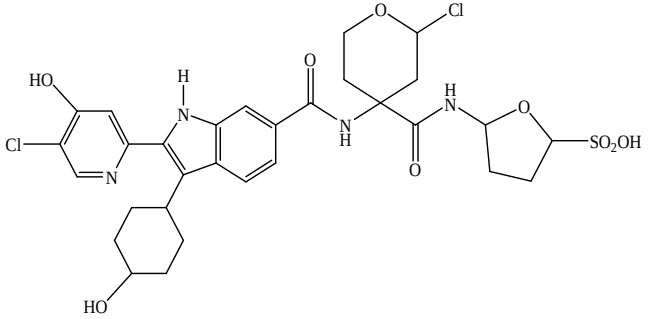
In deriving Eq. (1), however, certain compounds, namely 6, 26, 39, 46 and 48, were not included as they exhibited aberrant behavior. No reasons seem to be obvious to explain this behavior of these compounds.

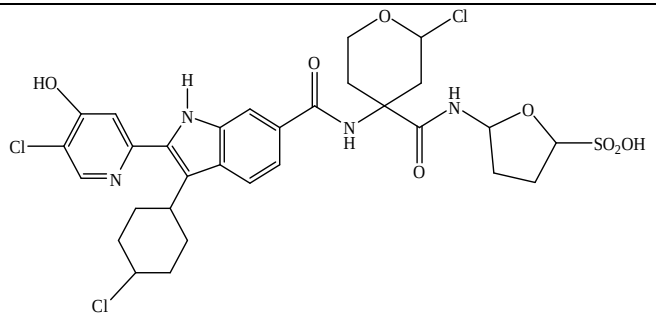
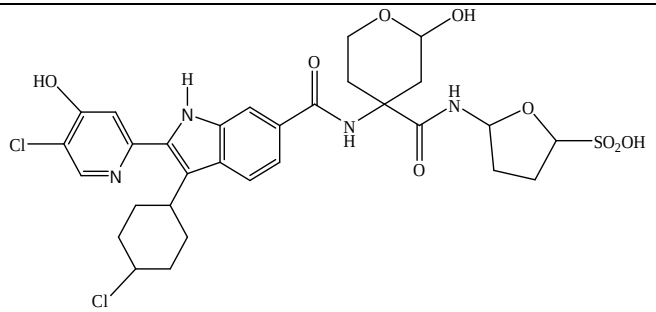
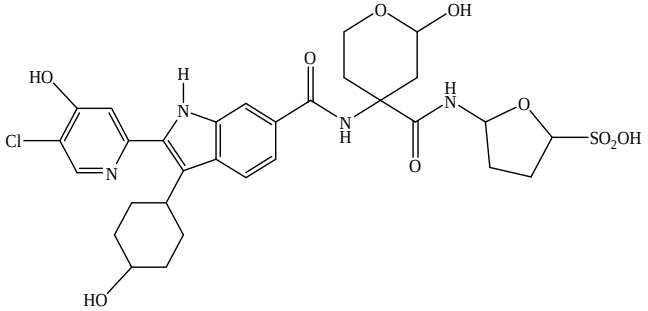
Using Eq. (1), we have predicted some new compounds as shown in **Table 3**. As can be seen, the predicted activities of these compounds are higher than any compound in the existing series (**Table 1**).

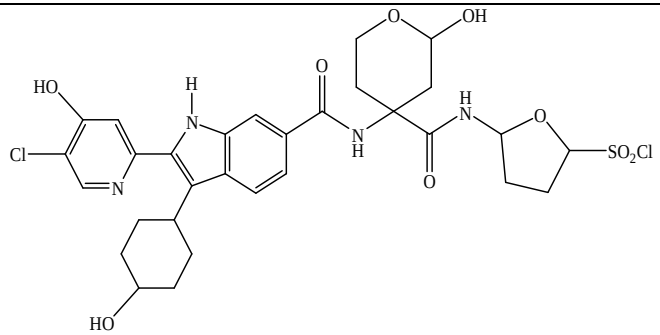
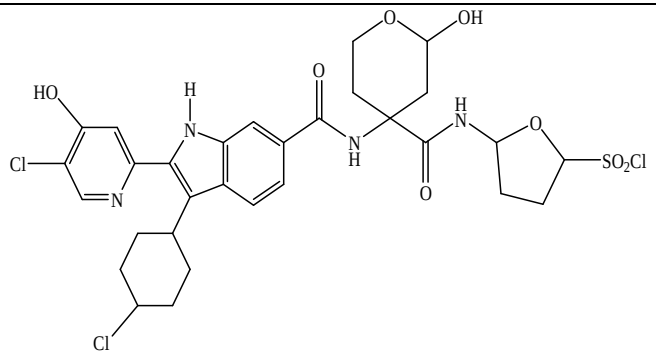
Table 3. Some Proposed Compounds belonging to the Series of Indole-5-Carboxamides and Their Predicted Activities

S.No	Compound	ClogP	W3D	Pol.	I	n_N	log ($1/IC_{50}$)
1.		2.17	12.83	62.73	1.00	4.00	8.43

2.		1.49	12.87	60.24	1.00	4.00	8.28
3.		2.26	12.18	58.63	1.00	4.00	8.10
4.		1.58	13.85	61.21	1.00	4.00	8.14

5.		2.09	14.28	62.39	1.00	4.00	8.07
6.		0.15	17.77	66.25	1.00	4.00	8.09
7.		0.85	16.16	66.33	1.00	4.00	8.33

8.		2.76	15.71	67.64	1.00	4.00	8.19
9.		1.20	16.07	66.33	1.00	4.00	8.28
10.		-0.719	16.53	65.02	1.00	4.00	8.42

11.		0.02	15.98	66.33	1.00	4.00	8.54
12.		1.94	15.48	67.64	1.00	4.00	8.41

Docking Results

Molecular docking is a computational technique for exploration of the possible binding modes of a substrate or inhibitor in a given enzyme or receptor to give the optimal interactions [13]. To perform a docking, the first requirement is to have 3D structure of the receptor or protein of interest which can be determined by X-ray crystallography or NMR spectroscopy. This protein structure and a 3D database of potential ligands serve as input to a docking program. The success of a docking program depends on two components, viz., the search algorithm and the scoring function.

Docking Simulations

Molegro Virtual Docker (MVD) [14, 15] (trial version) was used for flexible ligand docking wherein the software makes use of differential evolution algorithm [16]. Fast and accurate identification of potential binding modes during the search process is made by the use of predicted cavities. The scoring function makes use of piecewise linear

potential (PLP) [17]. The scoring function takes into account hydrogen bonding terms along with their directionality, ligand-protein interaction energy, and intramolecular interaction energy of the ligand. For enhanced docking accuracy, the highest ranked poses are yet again reranked [18].

Validation of Docking Method

For the present studies, we have selected the ligand 1BI_601 [B] (PDB code: 4GMC). The X-ray diffraction structure of HCV NS5B polymerase in complex with a thumb inhibitor (4GMC) refined at 2.70 Å resolution is considered for the docking studies. The ligand was extracted from the complex (4GMC) and redocked using flexible docking simulations into the original structure of HCV NS5B Polymerase. The protein complex of HCV NS5B polymerase with a thumb inhibitor (4GMC) was imported from the Protein Data Bank [18]. The scoring function MolDock Grid with 0.30 Å resolution along with an algorithm MolDock optimizer was used for docking. The following parameters were fixed: number of runs = 10, population size = 50, max iterations = 2000, termination scheme = variance based. Docking results have been shown for all the predicted compounds in **Table 4.** The hydrogen-bond, hydrophobic and electrostatic interactions are shown in **Figs (2-4)** for one of the predicted compounds that have the highest number of H-bonds (Compd. 10).

CONCLUSION: The HCV inhibition potency of indole-5-carboxamides is found to be controlled by hydrophobicity, polarizability, 3D Wiener index of the molecules and the number of nitrogen atoms present in them. Of all these parameters, the positive effect has been found to be produced by only polarizability, suggesting that drug-receptor interaction will involve predominantly only electrostatic interaction. Correlation also suggested through the use of one indicatior parameter that the 6-membered heterocyclic ring present in the molecule will have added advantage. Based on QSAR equation, some new compounds with higher activity have been predicted. A docking study has shown that these predicted compounds have better energies of interactions with the enzyme and docking scores than the original ligand complexed with the enzyme

Table 4: Docking Results of Predicted Compounds Relative to Reference Compound Complexed with Protein

Predicted compound	Overall Interaction Energy (kJ/mol)	H -bond Energy (kJ/mol)	No of H- bonds	H-bonds Ligands-Protein	H-bond Length(Å)	Mole Dock Score	Internal Energy of Pose (kJ/mol)
Ref. Compd	-119.444	0.000	0	–	–	-146.167	-26.723
1	-134.073	-2.442	2	O (31) – Glu (70)	3.52	-157.083	-0.857
				O (38) – Glu (70)	3.17		
2	-133.101	-5.127	3	O (31) – Glu (70)	3.16	-149.216	4.708
				O (38) – Glu (70)	3.52		
				O (39) – Tyr (296)	2.16		
3	-139.325	-8.584	5	O (31) - Ala (185)	2.73	-157.718	4.7656
				O (31) - Lys (74)	3.18		
				O (31) - Thr (77)	3.03		
				O (38) - Asp (66)	3.32		
				N (23) - Ala (73)	3.54		
4	-142.632	-2.500	1	O (34) – Asp (66)	3.04	-149.201	16.347
5	-140.045	-8.961	4	O (31) - Lys (74)	2.88	-152.811	9.244
				O (31) - Thr (77)	2.81		
				O (31) - Ala (185)	3.04		
				O (41) - Thr (69)	3.31		

6	-142.918	-6.570	4	O (44) – Lys (69)	2.87	-160.112	5.941
				O (45) – Lys (69)	2.87		
				O (36) - Glu (70)	3.09		
				O (43) - Lys (74)	3.14		
7	-148.810	-4.779	3	O (43) - Glu (70)	3.49	-164.601	5.897
				O (36) - Glu (70)	3.25		
				O (46) - Glu (70)	2.97		
8	-140.689	-2.831	2	O (36) - Glu (70)	3.53	-161.693	0.870
				O (46) - Glu (70)	3.11		
9	-143.385	-11.887	5	O (36) - Ala (185)	3.09	-162.718	3.050
				O (36) - Lys (74)	2.88		
				O (36) - Thr (77)	2.80		
				O (46) - Asp (66)	2.60		
				O (46) - Glu (70)	3.07		
10	-145.315	-16.701	7	O (36) - Ala (185)	2.96	-164.078	2.219
				O (36) - Lys (74)	3.05		
				O (36) - Thr (77)	2.84		
				O (46) - Ala (185)	2.95		
				O (46) - Lys (74)	2.98		
				O (46) - Thr (77)	2.90		

				O (44) - Lys (74)	3.26		
11	-137.195	-11.841	6	O (36) - Ala (185)	3.27	-169.792	-1.655
				O (36) - Lys (74)	2.93		
				O (36) - Thr (77)	2.85		
				O (44) - Lys (153)	3.00		
				O (45) - Lys (152)	3.11		
				O (43) - Asp (66)	3.33		
12	-142.169	-9.254	4	O (44) - Thr (77)	2.97	-163.495	0.206
				O (45) - Lys (69)	2.64		
				O (42) - Asp (66)	2.60		
				O (36) - Lys (69)	2.87		

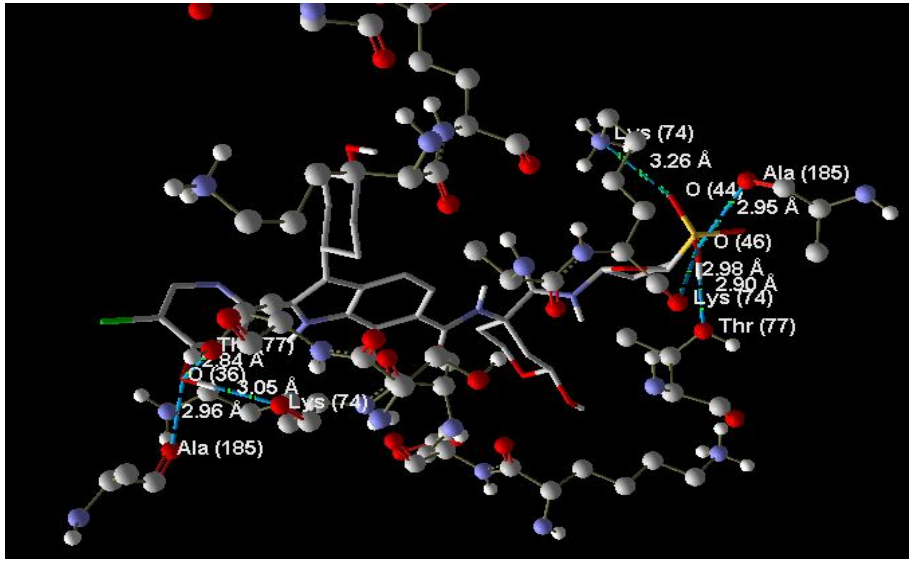


Fig.2: A model showing hydrogen bond interactions of predicted compound 10 with the enzyme HCV NS5B Polymerase. Compound 10 is one of the compounds that have the highest number of H-bonds

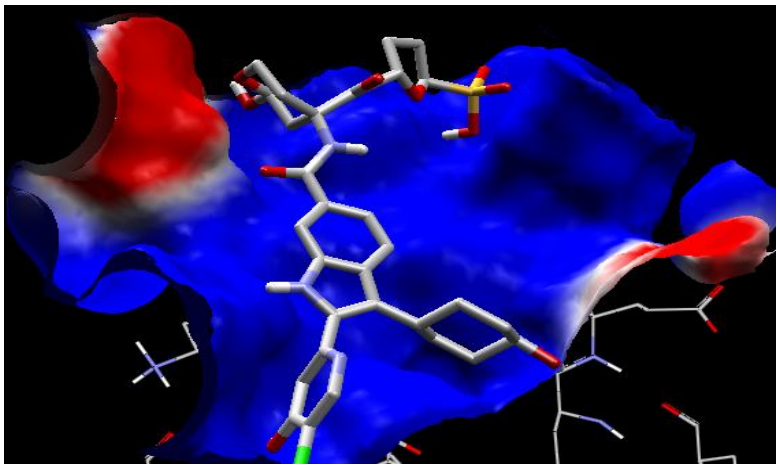


Fig. 3 The model showing electrostatic interactions of predicted compound 10 with the enzyme

HCV NS5B Polymerase.The blue surface shows strong electrostatic zone and red one showing the low electrostatic zone

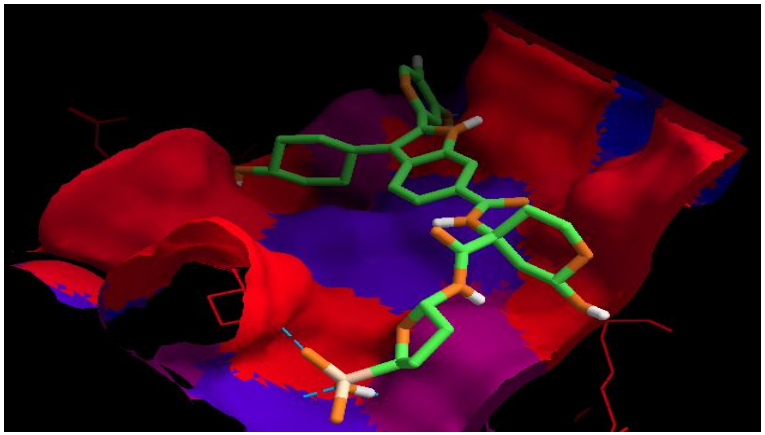


Fig. 4 The model showing hydrophobic interactions of predicted compound 10 with the enzyme HCV NS5B Polymerase. The red surface shows strong hydrophobic zone and blue one the low hydrophobic zone

REFERENCES:

1. Jeffrey, A. P.; Nugent, R; Gross, R. J. Inhibitors of HCV NS5B polymerase. Part 2: Evaluation of the northern region of (2Z)-2-benzoylamino-3-(4-phenoxy-phenyl)-acrylic acid. *Bioorg. Med. Chem. Lett.* 2005, *15*, 2812–2818.
2. Ishida, T.; Suzuki, T.; Hirashima, S. Benzimidazole inhibitors of hepatitis C virus NS5B polymerase: Identification of 2-[(4-diarylmethoxy) phenyl]-benzimidazole. *Bioorg. Med. Chem. Lett* 2006, *16*, 1859–1863.
3. Chinnaswamy, S.; Cai, H.; Kao, C. An update on small molecule inhibitors of the HCV NS5B polymerase: effects on RNA synthesis *in vitro* and in cultured cells, and potential resistance in viral quasispecies. *Virus Adapt. Treat.* 2010, *2*, 73–89.
4. Simmonds, P.; Holmes, E.C.; Cha, T.A.; Chan, S.W.; McOmish, F.; Irvine, B.; Beall, E.; Yap, P.L.; Kolberg, J.; Urdea, M.S. Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. *J. Gen. Virol.* 1993, *74*, 2391-2399.
5. Xavier, F.; Labrador, L. Hepatitis C virus NS3/4A protease inhibitors. *Recent Patents on Anti-Infect. Drug Discov*. 2008, *3*, 157-167.
6. Zuoa, G.; Li, Z.; Chena, L.; Xua, X. Activity of compounds from Chinese herbal medicine *Rhodiola kirilowii* (Regel) maxim against HCV NS3 serine protease. *Antiviral Res.* 2007, *76*, 86–92.
7. Kwong, C.D.; Clark, J.L.; Fowler, A.T.; Geng, F. Novel substituted pyrimidines as HCV replication (replicase) inhibitors. *Bioorg. Med. Chem. Lett.* 2012, *22*, 1160–1164.

8. Li, X.; Zhang, Y.K.; Liu, Y. Synthesis and evaluation of novel amino cyclic boronates as inhibitors of HCV NS3 protease. *Bioorg. Med. Chem. Lett.* 2010, *20*, 3550–3556
9. Wei, W.; Cai, C.; Kota, S. New small molecule inhibitors of hepatitis C virus. *Bioorg. Med. Chem. Lett.* 2009, *19*, 6926–6930.
10. Beaulieu, P.L.; Gillard, J.; Jolicoeur, E. From benzimidazole to indole-5-carboxamide thumb pocket I inhibitors of HCV NS5B polymerase. Part 1: Indole C-2 SAR and discovery of diamide derivatives with nanomolar potency in cell-based subgenomic replicons. *Bioorg. Med. Chem. Lett.* 2011, *21*, 3658-3663.
11. Beaulieu, P.L.; Chabot, C.; Duan, J. Indole 5-carboxamide thumb pocket I inhibitors of HCV NS5B polymerase with nanomolar potency in cell-based subgenomic replicons (part 2): Central amino acid linker and right-hand-side SAR studies. *Bioorg. Med. Chem. Lett.* 2011, 21, 3664–3670.
12. http://www.vcclab.org/lab/edragon/
13. Gupta, S.P.; QSAR and Molecular Modeling. Anamaya Publishers, New Delhi, 2011.
14. http:// www.molegro.com (free trial version).
15. Thomsen, R.; Christen, M.H. MolDock. A new technique for high accuracy molecuar docking. *J. Med. Chem.* 2006, 49, 3315-3321
16. Yang, J.M.; Chen, C.C. A generic evolutionary method for molecular Docking. *Proteins* 2004, 55, 288-304.
17. Thomsen, R. Flexible ligand docking using evolutionary algorithms: investigating these effects of variation operators and local search hybrids. *Biosystems* 2003, 72, 57-73.
18. www.rcsb.org.