ScienSage

Journal of Advanced Scientific Research

Available online through http://www.sciensage.info

ISSN
0976-9595
Research Article

REGRESSION ANALYSIS AND DOCKING STUDIES OF COUMARIN BASED COMPOUNDS AS ANTI-CANCER AGENTS

Bhagat G. K.[1], Patle M.R.*[2]

[1]*Jagat Arts Comm. and IHP Sc. College, Goregaon, Maharashtra, India*

[2]*D. B. Science College, Gondia, Uttar Pradesh, India*

**Corresponding author: manojpatle14@gmail.com*

ABSTRACT

In present study, QSAR studies of Coumarin based compounds as anticancer agents is discussed and aimed at finding the equation for Coumarin based molecules as anticancer agent. At the beginning, the molecules whose biological properties are known are well-thought-out as a known set for regression analysis model building purpose. The descriptors were calculated for known set of Coumarin based compounds. Novel substituted Coumarin based molecules were designed improved and their descriptors are calculated. Moreover, the regression analysis model is used to determine the biological activities of these new molecules. On a systematic study, it is possible to enhance the activities of the molecules by using regression and docking methodology.

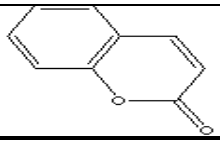
Keywords: QSAR, Benzopyrones, Coumarin, Molecular descriptors, Docking, Anti-cancer

1. INTRODUCTION

The premier health organization World Health Organization (WHO) has prepared Global burden diseases report under the instruction of World Bank in the year 1990 to understand the issue of risk level of diseases in developing and under developed countries. According to this report [1], there would be the likelihood of health crisis in the year 2020. Hence it is the urgent need to concentrate research on the most attributable diseases such as cardiovascular diseases, unipolar depression, cerebrovascular disease, Chronic Lung disease, and Lower respiratory track diseases, HIV and Tuberculosis and Cancer [2].

There are many areas which also needs tremendous research like problems due to road accidents, violence, self- inflicted injuries, etc. This report insists on need to have advanced tools to fight against the challenges due to the Cancer diseases [1]. Failing will result in pandemonium in the civilized world and will result in the uncontrolled spread of epidemic and violence; hence the society has to reinvent the wheel of progress [2].

Table 1: The basic details of Coumarin Molecules

Common Name	IUPAC Name	Molecular Formula & M.Wt.	Structural Formula
Coumarin	1,2-Benzopyrone and 2-Oxo-1,2-benzopyran	$C_9H_6O_2$ M.Wt. = 146.15	

2. EXPERIMENTAL

Present study deals with the QSAR study of Coumarin based series of compounds. The main aim of the experiment is to evaluate Multiple Regression Analysis for the series of newly designed Coumarin based compounds for understanding the properties of molecules participates in anti-cancer activities.

The stepwise experimental procedure is as follows for both types of molecules.

2.1. Identification of Known Compound Set

In QSAR methodology the first step is the Identification of the Coumarin based molecule series with known anti-cancer activities from literatures. This set of compounds is treated as Training Set for the design of QSAR model.

The known set of Coumarin based molecule is procured from the work of Zuping Xia et.al. [3] and this set of molecules is treated as Training Set**.**

2.2. Designing of Molecules

CHEMDRAW [4, 5] software is used to design both 2D and 3D structures. While designing molecules virtually, precautions are being taken for 3D as it tends to lock in "Local Minima". Full care is being taken to see that all molecules are reaching to their "Global Minima" and hence the designed molecules are verified for their 3D structures.

Total sixteen molecules are designed and verified for their 3D structures. Semi-empirical QM/MM2 method is used to design 3D molecules with ready-made computer based tools. The final structure of the molecule is having very less total energy and having minimum strain.

The various physio-chemical properties are derived using the advanced computer model for each known molecules and these properties are listed in table 3 and 4 for Coumarin (Training Set) and Coumarin (Test Set) based molecules respectively.

2.3. QSAR using stepwise multiple regression analysis

Stepwise multiple regression analysis performed by NCSS [6] software for both series *i.e.* for Coumarin based molecule series. Coumarin based molecule series is named as "Training Set' I"

In the preset study, the multiple regression analysis (stepwise) was performed for both the series by treating biological activity as dependent variable and physio-chemical properties as independent variables. The detail methodology and results obtained are described in following section.

In this methodology' initially 35 variables (properties) of each molecule are supplied to the system. The properties whose co-efficient has negligible values are removed and again the regression is performed. The data related to different parameters is given in table 2, 3, 4 and 5.

Table 2: Structure, IUPAC name and molecular formula of coumarin based compounds of (training set)

Mol. No.	IUPAC Name	Molecular Formula	Mol. WT. (amu)	Biological Activity
1	6 –((2*R*)-tetrahydro-3,4,5- trihydroxy-6-(hydroxymethyl)- 2*H*-pyran-2-yloxy)-7- hydroxyl p-*H*-chromen-2-one	$C_{15}H_{16}O_9$	340.288	3.16
2	(2*S*)-tetrahydro-3,4,5- trihydroxy-6-(hydroxymethyl)- 2*H*-pyran-2-yloxy)-7- hydroxyl p-*H*-chromen-2-one	$C_{15}H_{16}O_9$	340.288	2.51
3	7-hydroxy-6-(tetrahydro-3,4,5-trihydroxy-6-(hydroxymethyl)- 2*H*-pyran-2-yloxy)-2*H*-chromen-2-one	$C_{15}H_{16}O_9$	340.288	12.58
4	7-hydroxy-6- (tetrahydro-3,4,5-trihydroxy-6- (hydroxymethyl) 2*H*-pyran-2-yloxy)-2*H*-chromen-2-one	$C_{15}H_{16}O_9$	340.288	7.94
5	7-hydroxy-6-methoxy-2*H* -chromen-2-one	$C_{10}H_8O_4$	192.172	3.16
6	6,7-dihydroxy-2*H*- chromen-2- one	$C_9H_6O_4$	178.145	12.58
7	7,8-dihydroxy-6-methoxy-2*H*- chromen-2- one	$C_{10}H_8O_5$	208.172	7.94
8	6,7-dihydroxy-4-methyl-2*H*- chromen-2- one	$C_{10}H_8O_4$	192.172	15.84
9	2-acetyl-2,3-dihydro-(1,4) dioxino(2,3-g)chromen-7- one	$C_{13}H_{10}O_5$	246.221	25.11
10	6,7,8-trimethoxy- 2*H*-chromen-2-one	$C_{12}H_{12}O_5$	236.221	25.11
11	7,8-dihydroxy-4-methyl-2*H*- chromen-2- one	$C_{10}H_8O_4$	192.172	0
12	7-(3-methylbut-2-enyloxy)-8-hydroxy-6-methoxy-2*H*-chromen-2- one	$C_{15}H_{16}O_5$	276.291	7.94
13	7,8-dihydroxy-2H-chromen-2-one	$C^9H_6O_4$	178.145	31.62
14	7,8-dihydroxy-4-methyl-2H-chromen-2-one	$C_{10}H_8O_4$	192.172	0
15	3-oxy-chromen-2-one-6-oxy-methyl-chromen-2-one	$C_{19}H_{12}O_7$	352.302	31.62
16	7-(3-methylbut-2-enyloxy)-8-hydroxy-6-methoxy-3H-chomen-2-one	$C_{15}H_{16}O_5$	276.291	39.81

This methodology is repeated until all the co-efficient are within the acceptable limits. The detail parameters passed to the model system along with the setting for running the multiple regression analysis. The multiple regression equation is listed in Table 6.

The computer based model system provides a number of additional information about the regression including the various plots.

The values of coefficients of various variables and multiple regression equations are listed in Table 6.

Table 3: Calculated parameters of Training Set molecules (Coumarin based molecules)

Mol. No.	Partition Coefficient	Non-1,4 VDW Energy	Electrical Energy	HOMO	LUMO	Repulsion Energy	Balaben Index	Molecular Topological Index	Wiener Index	Shape Coeff.	Shape Attrib.	Total Energy
1	-0.7366	-12.0755	-33252.1	-9.4541	-0.6204	28230.4	414082	8690	1339	22.0417	1	3.8254
2	-0.7306	-10.3887	-33891.2	-9.47607	-0.6625	28869.3	414082	8690	1339	22.0417	1	5.0052
3	-0.7306	-12.9774	-33425.2	-9.40675	-0.5747	28403.6	414082	8690	1339	22.0417	1	3.6439
4	-0.7306	-18.3314	-34249.9	-9.4234	-0.56	29228	414082	8690	1339	22.0417	1	-2.9521
5	1.3518	-1.2659	-13104.1	-9.008	-0.4041	10435.8	39944	2036	288	12.0417	1	6.816
6	0.700681	-1.9839	-19261.5	-9.6068	-0.9255	15697.3	132086	3918	586	8	16.055	7.433
7	1.214	-2.4289	-11555.9	-9.397	-0.5791	9042.9	27687	1604	230	11.0759	0.75	0.1489
8	0.9147	-7.7323	-15145.6	-9.1204	-0.4929	12156.7	53266	2300	336	13.0667	1	-0.8005
9	1.713	-7.7961	-13153.5	-9.2441	-0.47634	10484.9	38334	1942	276	12.0714	0.75	-5.2843
10	0.9426	-3.1571	-18971.5	-9.24921	-0.4822	15571.8	105500	4156	592	16.0556	1	7.0976
11	1.0324	-5.1864	-1892.3	-9.5774	-0.6613	15625.3	94930	3302	470	15.0588	1	6.8062
12	3.1366	-3.3471	-19421.3	-9.1063	-0.3706	16158	105500	4338	592	16.0556	1	9.1542
13	1.214	-7.8966	-11758.6	-9.24249	-0.46547	9245.33	26979	1576	224	11.0769	1	-5.6594
14	1.713	-8.2341	-13296.5	-9.2251	-0.4123	10627.4	37517	1914	270	12.0714	1	-5.6412
15	2.7198	-7.3963	-31525.8	-9.32121	-0.6962	26693.3	520928	12148	1743	24.0385	0.8571	5.4919
16	2.6502	-7.0388	-23084.1	-9.4082	-0.6584	19345	226560	5902	819	18.05	1	3.5084

Table 4: Calculated parameters of Test Set II molecules (Coumarin based molecules)

Mole No.	MP	Pc	Tc	Vc	HF (Kj/mol)	Henry's Law	Log P	MP	Molar Refractivity	GE	CMR	Partition Coefficient
1	605.6	48.3595	852.713	425.5	-341.54	7.5312	1.373	410.14	43.2427	-226.79	4.3161	1.62348
2	640.976	60.1858	879.31	441.5	-518.85	11.514	1.0458	521.86	44.9368	-381.41	4.4692	1.2141
3	622.296	43.8577	840.475	428.5	-327.93	7.0858	1.1656	371.76	43.2193	-218.62	4.3161	1.9163
6	640.976	60.1858	879.31	441.5	-518.85	11.514	1.0458	521.86	44.9368	-381.41	4.4692	1.214
9	651.583	51.0204	879.9	497.5	-550.96	11.3189	1.2216	545.65	49.2188	-382.62	4.933	1.7131
10	664.629	49.8035	886.742	515.2	-683.18	12.742	0.9194	567.88	51.4	-487.62	5.0861	0.9147
11	577.466	34.6836	820.683	465.5	-196.34	3.3533	1.746	322.21	45.9057	-73.38	4.6268	1.911
14	629.794	40.7771	859.205	499.5	-505.87	8.7592	1.3089	456.16	49.7059	-333.1	4.933	1.3518
15	654.462	29.3134	896.938	637.5	-63.01	4.8993	3.3628	404.98	66.0007	81.13	6.6742	3.51

Table 5: Continued: Calculated parameters of Test Set molecules (Coumarin based molecules)

Mole. No.	Non-1,4 VDW Energy	Electrical Energy	HOMO	LUMO	Repulsion energy	Balaben index	Molecular Topological index	Wiener index	Total Energy	Calculated B.A.	ΔG (Kcal/mol)
1	-1.5737	-9892.78	-9.4241	-0.3909	7700.06	19099	1358	185	2.0287	7.94	-11.5101
2	-2.4289	-11555.9	-9.3981	-0.5782	9042.89	27687	1604	230	0.1489	3.16	-7.7178
3	-1.5714	-9965.3	-9.6338	-0.3956	7772.65	18490	1326	179	1.722	24.11	-9.2386
6	-7.8966	-11758.6	-9.2424	-0.4654	9245.33	26979	1576	224	-5.6594	12.87	-8.2386
9	-2.7506	-13122	-9.3364	-0.5313	10453.2	38334	1942	276	0.113	7.94	-8.4427
10	-9.1089	-15160.5	-9.3241	-0.5725	12171.5	53266	2300	336	-4.7894	17.85	-7.9872
11	-1.4225	-9738.51	-9.4402	-0.2365	7710	18797	1426	182	3.5026	25.11	-8.04014
14	-2.7483	-13125.4	-9.4183	-0.5151	10457.1	39944	2036	288	2.8342	21.18	-6.82113
15	-0.6323	-15462.1	-8.6242	-1.023	12766.6	79721	3906	498	9.8545	31.64	-8.29126

Table 6: Estimated Model for calculation of biological activity on the basis of regression analysis for Coumarin based molecule series (Training Set).

Y=1326.13854810737-0.17547746126664*C1 +4.53265107717623 * C10 + 0.710061473430893 * C11 - 347.020842882972 * C12 + 0 * C13 -0.490300659985753 * C14 + 0*C15 + 0*C16 + 0*C17+ 0*C18+ 0*C19+ 0*C2+ 0*C20+ 0*C21 + 0 * C22 + 18.887467251569 * C23 + 0 * C24 + 0 * C25 + 0 * C3 + 0 *C4+ 0*C5+ 0*C6+ 0*C7+ 0*C8+ 0*C9
The filtered equation after removing all those variables having 0 (zero) coefficient values.
Y=1326.13854810737 - 0.17547746126664 * C1 + 4.53265107717623 * C10 + 0.710061473430893 * C11 - 347.020842882972 * C12 - 0.490300659985753 * C14 + 18.887467251569 * C23
Removed variables along with their coefficient are as follows:
0*C2, 0 * C3, 0 *C4, 0*C5, 0*C6, 0*C7, 0*C8, 0*C9, 0 * C13, 0*C15, 0*C16, 0*C17, 0*C18, 0*C19, 0*C20, 0*C21, 0 * C22, 0*C24, 0 * C25,

Where,

C1 - Partition Coefficient,	C2 - Non 1,4 VDW Energy,	C3 - Electrical Energy,
C4 - HOMO,	C5 - LUMO,	C6 - Repulsion Energy,
C7 - Balaben Index,	C8 - Mol.topological Index,	C9 - Wiener Index,
C10 - Shape Coefficient,	C11- Shape Attrib.,	C12 - Total Energy,
C13 - Boiling Point (BP),	C14 - Melting Point (MP),	C15-Critical Temp.,
C16 - Critical Pressure(Pc),	C17 - Critical Volume (Vc),	C18-Gibb's Free Eenergy,
C19 - Log P,	C20 - Molar Refractivity,	C21- Henry's law Constant,
C22 - GE	C23 - CMR,	C24 - ClogP
C25 – Biological Activity,	C26 – Binding Energy (ΔG)	

The report is generated by computer based statistical analysis software NCSS [6].

2.4. Validation of Model

It is necessary to test the resulted regression equations. In the present study, the regression equations were tested with the known molecules, which were not included in training set.

It is reported that the experimental biological activity and calculated biological activities of these molecules are nearly matching gives us confidence of our approach.

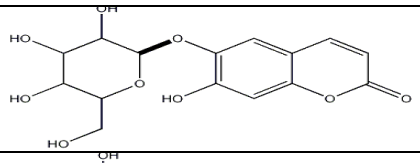
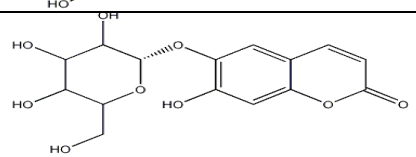
2.5. Design of Coumarin based "Unknown" molecules

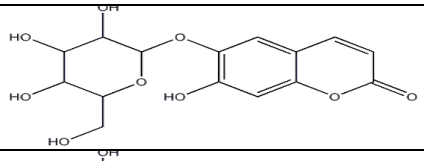
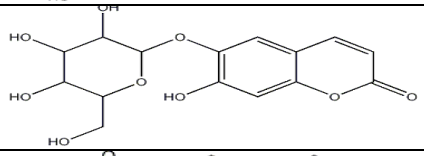
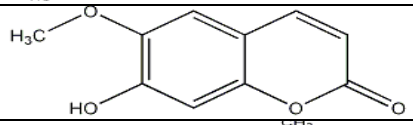
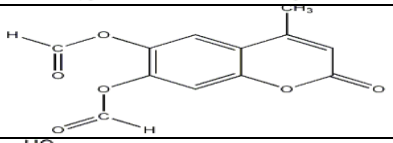
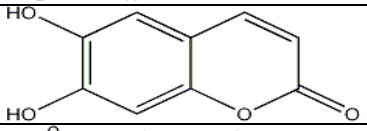
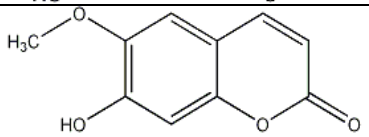
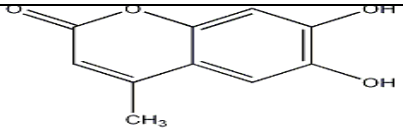
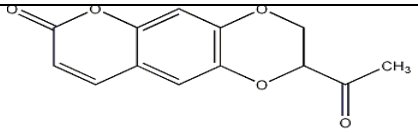
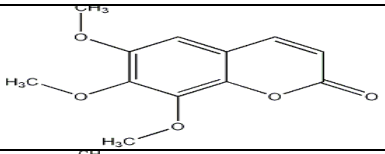
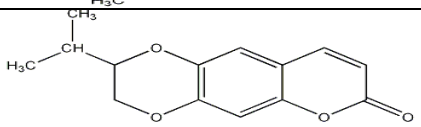
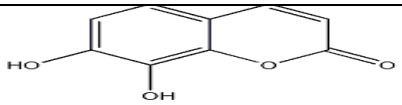
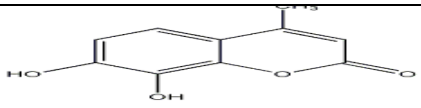
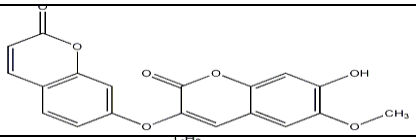
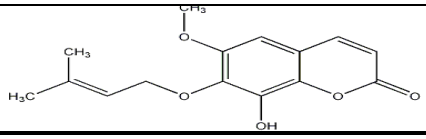
Once equation for calculating B.A. is known, knowing the needed contributions by any typical properties of molecules, new set of molecules are designed *in-silico* for both the types.

While preparing these molecules care is being taken to see that those properties which enhance the BA are keenly selected in new molecules while those properties which reduce BA are rejected from series.

In this table (Table 7) the molecule number provides in unique way. For example, Molecule Number 4a is actually the modification of molecule number 4 from same category of training set. Molecule No. 4 is from the Training Set, whereas Molecule No. 4a is from testing set. New molecule was designed by modifying the molecule no. 4. Hence, Molecule No.4a is a modification (virtually) of a molecule no. 4.

Table 7: Structure, IUPAC name and molecular formula of Coumarin based compounds of Test Set.

Mol. No.	IUPAC Name	Molecular Structure	Molecular Formula	Mol. Wt. (amu)
1a	6 –((2*R*)-tetrahydro-3,4,5- trihydroxy-6-(hydroxymethyl)-2*H*-pyran-2-yloxy) -7- hydroxyl p-*H*-chromen-2-one		$C_{15}H_{16}O_9$	340.288
2a	(2*S*)-tetrahydro-3,4,5- trihydroxy-6-(hydroxymethyl)-2*H*-pyran-2-yloxy)-7- hydroxyl p-*H*-chromen-2-one		$C_{15}H_{16}O_9$	340.288

3a	7-hydroxy-6-(tetrahydro-3,4,5-trihydroxy-6-(hydroxymethyl)-2*H*-pyran-2-yloxy)-2*H*-chromen-2-one		$C_{15}H_{16}O_9$	340.288
4a	7-hydroxy-6- (tetrahydro-3,4,5- trihydroxy-6-(hydroxymethyl) 2*H*-pyran-2-yloxy)- 2*H*-chromen-2-one		$C_{15}H_{16}O_9$	340.288
5a	7-hydroxy-6- methoxy-2*H* -chromen-2-one		$C_{10}H_8O_4$	192.172
6a	4-Methyl, 6-7-dioxyaldehyde-chromen-2-one		$C_{12}H_8O_6$	248.193
7a	6,7-dihydroxy-2*H*- chromen-2- one		$C_9H_6O_4$	178.145
8a.	7,8-dihydroxy-6- methoxy-2*H*- chromen-2-one		$C_{10}H_8O_5$	208.172
9a	6,7-dihydroxy-4- methyl-2*H*- chromen-2- one		$C_{10}H_8O_4$	192.172
10a	2-acetyl-2,3-dihydro- -(1,4) dioxino(2,3-g) chromen-7- one		$C_{13}H_{10}O_5$	246.221
11a	6,7,8-trimethoxy- 2*H*-chromen-2-one		$C_{12}H_{12}O_5$	236.221
12.a	2,3-dihydro-2-isopropyl-(1,4) dioxino(2,3-g) chromen-2- one		$C_{14}H_{14}O_4$	246.264
13a	7,8-dihydroxy-2*H*- chromen-2- one		$C_9H_6O_4$	178.145
14a	7,8-dihydroxy-4- methyl-2*H*- chromen-2- one		$C_{10}H_8O_4$	192.172
15a	3-oxy-chromen-2-one-6-oxy-methyl-chromen-2-one		$C_{19}H_{12}O_7$	352.302
16a.	7-(3-methylbut-2- enyloxy)-8-hydroxy- 6-methoxy -2*H*-chromen-2- one		$C_{15}H_{16}O_5$	276.291

The modification was carried out by removing or by adding new functional groups to this molecule. These changes help in either enhancing or reducing the required physio-chemical properties of newly designed molecules. The molecules which show the positive change in the properties are selected. Positive changes means those changes which enhances the anti-cancer properties of the molecule by carrying out structural changes such as the functional group change and that can be judged by knowing the values of parameters.

2.6. Confirmation of QSAR model and derivation of equations.

The properties of designed molecules (Test Set) are utilized to find out the "Calculated Biological Activities" using the equation which is already acquired.

Table 8: Calculated Properties values of designed (Test Set II) Coumarin based molecules.

Mol No	BP	MP	Tc	Pc	Vc	GE (Kj/mol)	Log P	MR cm^3\mol	Henry's Law	HF Kcal/mol	Clog P	CMR	Cal B.A.	ΔG kcal/mol
1a	1143.54	772.78	1057.81	38.39	805.5	-930.69	-0.79	79.67	19.3	-1377.03	-0.7306	7.8401	3.16	-8.8825
2a	1143.54	772.78	1057.81	38.39	805.5	-930.69	-0.79	79.67	19.3	-1377.03	-0.7306	7.8401	2.51	-12.1373
3a	1143.54	772.78	1057.81	38.39	805.5	-930.69	-0.79	79.67	19.3	-1377.03	-0.7306	7.8401	12.58	-8.5781
4a	1143.54	772.78	1057.81	38.39	805.5	-930.69	-0.79	79.67	19.3	-1377.03	-0.7306	7.8401	7.94	-9.5101
5a	632.6	456.16	859.2	40.78	499.5	-333	1.31	51.16	8.76	-505.87	1.3518	4.933	3.16	-11.7178
6a	727.44	498.25	892.8	30.66	647.5	-484.84	0.94	59.72	6.92	-696.16	0.70068	5.932	25.11	-12.2386
7a	662.94	521.86	879.31	60.19	441.5	-381.41	1.05	45.72	11.51	-518.85	1.214	4.4692	12.58	-8.2386
8a	713.22	567.88	886.74	49.8	515.5	-487.62	0.92	52.97	12.74	-683.18	0.9148	5.0861	7.94	-12.4427
9a	690.8	545.65	879.9	51.02	497.5	-382.62	1.22	49.96	11.32	-550.96	1.7131	4.933	15.84	-8.9872
10a	721.96	486.03	892.36	30.83	630.5	-310.26	0.6	64.65	9.53	-579.67	0.9426	6.1827	25.11	-8.04003
11a	652.54	436.48	849.01	25.74	631.5	-390.8	1.45	63.84	7.23	-657.22	1.0324	6.0137	25.11	-6.82113
12a	690.53	432.37	872.83	26.24	674.5	-175.36	2.53	68.78	5.99	-493.01	3.1366	6.6108	31.64	-9.29191
13a	690.53	432.37	872.83	26.24	674.5	-175.36	2.53	68.78	5.99	-493.01	3.1366	6.6108	31.62	-8.64611
14a	690.8	545.65	879.9	51.02	497.5	-382.62	1.22	49.96	11.32	-550.96	1.7131	4.933	0	-8.68494
15a	967.08	712.59	1047.52	25.46	891.5	-478.55	2.12	93.65	13.11	-781.93	2.7198	9.0717	31.62	-9.90318
16a	778.44	528.22	899.14	24.48	778.5	-333.86	2.42	78.01	9.48	-645.33	2.6502	7.3797	39.81	-8.14903

3. RESULTS

3.1. Regression Analysis

Present study envisages the task to find out a best regression analysis for two types of molecular series. They are Coumarin based molecules. The known series of molecules are already tested for their Biological Activities. No model is valid until it is checked for the accuracy and reproducibility. In this section, the observed values *i.e.* the equations for test series is used to predict the biological activities of known and unknown set of molecules.

The derived equation for test set is depicted in Table 6.

The values of a coefficient for Training Set ' indicated that Balaben index, Molecular Topological Index, Wiener Index, Repulsion energy, Electrical Energy had the negligible contribution for making Coumarin based molecules as an anti-cancer agent. Log P, CMR, HOMO, Henry's Law value contributed negatively in making Coumarin molecule as an anti-cancer agent. These values reduced the inhibition ability against cancer proteins.

Using this equation the biological activities for the test series are calculated and it is reported that the values are within 98% confidence level which is the great achievement of present work. Table 10 depicts the comparison between the experimental BA with calculated BA. The calculated BA is in excellent agreement with the desired one and hence the work envisage in present study successfully accomplished.

3.2. Docking Studies

After obtaining the regression equation, the equation was tested with the Training set and Test set of molecules. To extend the verification of our model, docking studies were performed for designed set (test set) of molecules. The docking gives the "binding

energy" between the ligand (small molecule) and protein. For good inhibitor the binding energy is always negative.

The model needs further validation on the basis of enhancing anti-cancer properties. To accomplish this, both known (Training Set) and designed (Test Set) are docked with 1GII proteins, which is known to be responsible for the cancer activities.

Table 9: Comparison between the experimental B. A. with calculated B. A. for training set

Molecule	B. A.	Calculated B. A.
1	1.92	1.9212
2	3.76	3.7622
3	4.29	4.2932
4	4.99	4.9923
5	4.71	4.7123
6	4.53	4.535
7	5.02	5.012
8	5.12	5.1356
9	4.31	4.367
10	4.56	4.534
11	4.87	4.8567
12	4.89	4.8834
13	3.62	3.645
14	3.3	3.43
15	3.79	3.734
16	3.52	3.525
17	4.31	4.356
18	3.98	3.943
19	4.13	4.158
20	4.73	4.726
21	5	5.012
22	4.3	4.33

4. CONCLUSION

The main aim of the Present study was to find out the best regression analysis equation for Coumarin based molecules. It is already known that Coumarin molecules are the anti-cancer agents but there is a need for the enhancement in their effectively.

The Table 10 show the docking results of Coumarin based (Test Set) molecules in the form of "Binding Energy i.e. ΔG Kcal/mol. These molecules are docked with the oral cancer prone protein 1GII.

The reported binding energy values are comparable with the calculated BA values and hence the models for Coumarin based molecules are valid and can be used to predict the molecule for its anti-cancer activities.

From present study, it is concluded that;

- It is possible to find out the predicted biological activities for Coumarin based molecules by using the above equation.
- The multiple regression analysis reports the various coefficient values for various physio-chemical properties of a molecule. These coefficients help us in rejecting or accepting a particular property for designing a molecule as an anti-cancer agent.
- It is reported that on increasing the chain length, the biological activity increases.
- Hydrophobic nature of a molecule enhances the anticancer activities of Coumarin based molecules.
- It is observed that when less number of carbon atoms are present, decreases the biological activity whereas reverse is reported when the carbon atoms increases.

Table 10: Docking results in the form of "Binding Energy *i.e.* ΔG Kcal/mol. for Coumarin based (Test Set) molecule

S. N.	Molecule No.	Molecular Formula	Calculated B.A.	Binding Energy ΔG Kcal/mol.)
1.	Mol 2	$C_9H_6O_3$	2.51	-9.4241
2.	Mol 3	$C_9H_6O_3$	12.58	-9.3981
3.	Mol 4	$C_9H_6O_3$	7.94	-9.6338
4.	Mol 10	$C_9H_6O_4$	3.16	-9.2424
5.	Mol 17	$C_{10}H_8O_4$	25.23	-9.3364
6.	Mol 18	$C_{10}H_8O_5$	12.34	-9.3241
7.	Mol 22	$C_{10}H_8O_2$	7.89	-9.4402
8.	Mol 28	$C_{10}H_8O_4$	15.84	-9.4183
9.	Mol 32	$C_{15}H_{10}O_2$	25.80	-8.6242

QSAR model has its own limitation. It is a statistical model and the designed molecules are from computer-based software. Furthermore, properties are evaluated using computer-based tools. Though this model does have its certain limitations, but still it helps in finding the better anti-cancer molecules well before synthesizing in the laboratory. Using derived equation, it is now possible to check any Coumarin based molecule for their anti - cancer activity.

In present study 1GII protein is used as receptor and Coumarin based molecules as a ligand. Nearly all molecules dock successfully and giving negative ΔG values in Kcal/mol. Negative values supports the stability of complex and hence it is also proved form the docking study that Coumarin based molecules are good anti-cancer agents.

The derived equation will help synthetic chemists to design the right molecule before synthesizing it in the laboratory and hence saving number of years and also huge cost. It also helps the environment, as unnecessary chemicals are not wasted and a supportive method for Green Chemistry.

5. REFERENCES

1. http://www.who.int/en
2. King, R.J.B. (2000), Cancer biology, 2nd Ed, School of biological science university of surrey 158-168.
3. Xia Z, Kinaak C, Ma J, Beharry Z, McInnes C, Wang W, Kraft A, Smith C. *J. Med. Chem.,* 2009; **52(1):**74-86.
4. CambridgeSoft\ChemOffice2004\ChemDraw\READMECP.HT M
5. http://sdk.cambridgesoft.com/
6. www.ncss.com