CHAPTER 5

PRESENT WORK: QSAR STUDY OF SOME ANTICANCER DRUGS

CHAPTER 5

WORK DONE

1. **2D QSAR studies on New Stilbene Derivatives of Resveratrol as New Selective Aryl Hydrocarbon Receptor**
2. **QSAR studies on Hetaryl imidazoles Derivatives as novel dual inhibitors of VEGF receptors I and II**
3. **2D-QSAR studies on Carbazole Sulfonamide Derivatives as Antimitotic Agents Against solid Tumors**
4. **QSAR studies on 4,5,6,7- Tetrabromobenzimidazole Derivatives as inhibitors of Protein kinase CK**
5. **2D-QSAR studies on N-Phenyl-N^{I}{4-(4-quinolyloxy) phenyl} urea Derivatives as VEGFR-2 Tyrosine Kinase Inhibitors**
6. **QSAR studies of Pyrazolo(3, 4-d) pyrimidines Derivatives as potent inhibitors of the Insulin-like growth factor receptor (IGF-IR)**
7. **QSAR studies of Pyrrolidiones Derivatives as potent functional antagonists of human melanocortin-4 receptor**
8. **QSAR studies of 2-(alkyloxyaryl)-1*H*-benzimidazole Derivatives as antitumor agent**
9. **QSAR studies on 3, 17-Disubstituted 2-Alkylestra-1, 3, 5(10)-trien-3-ol Derivatives as Anticancer Agents**

CHAPTER 5

PRESENT WORK: QSAR STUDY OF SOME ANTICANCER DRUGS

QSAR study has been performed on some series of anticancer drugs, with a hope to find important physico chemical parameter which influence the anti cancer activity of drug molecules given in the series.

1. 2D QSAR studies on New Stilbene Derivatives of Resveratrol as New Selective Aryl Hydrocarbon Receptor

Introduction

Aryl hydrocarbon receptor[146](AhR) or Dioxin (TCDD,2,3,7,8-tetrachlorodibenzoparadioxin) receptor binds *cis* acting dioxin responsive elements and modulates the expression of various genes [147,148]. The AhR agonists such as TCDD, benzo(α)pyrene (BAP), and 7,12 dimethyl benzenthracene (DMBA) are environmental toxicants suspected to be responsible for numerous pathologies in human such as cancers, atherosclerosis, osteoporosis, skin disorders and reproductive failures. The main lines of evidence highlighting the importance of the AhR in the mediation of carcinogenicity of aryl hydrocarbons are:

1) BAP is not carcinogenic in mice lacking AhR.[149]

2) Mice expressing constitutively active AhR spontaneously develop cancers.[150, 151]

Keeping in mind these observations, it was thought that AhR antagonists are of interest for the development of prophylactic as well as curative drugs for major diseases involving AhR activation. Several ligands for AhR such as halogenated aromatic hydrocarbons(HAH) and

polycyclic aromatic hydrocarbons(PAH) have been studied. Besides environmental toxicants, the physiological oxysterol like 7- ketocholesterol, and phytochemicals like flavonoids and resveratrol have also been identified as AhR mixed agonist/antagoist[151-154]. In view of multiple potencies of resveratrol((E)-1-(4'-hydroxyphenyl)-2-(3,5-dihydroxyphenyl)ethene) (**fig.1.1**) in various pathologies from osteoporosis to chemoprevention against environmental contaminants, twenty four stilbene derivatives taken from literature[**146**] were subjected to QSAR analysis in order to understand the role of different physicochemical and structural parameters in binding of these molecule with AhR for which no 3-dimensional molecular structure is known. The QSAR[**155**] studies have been carried out using Linear Free Energy Relationship (LFER) approach taking the anticancer activity as dependent and different physicochemical parameters like hydrophobic (π), steric (MR) and electronic (f, R, σ) as independent parameters .

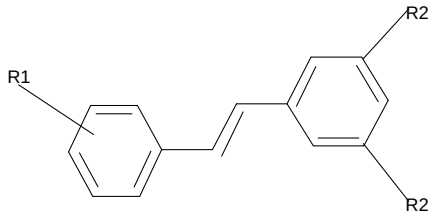


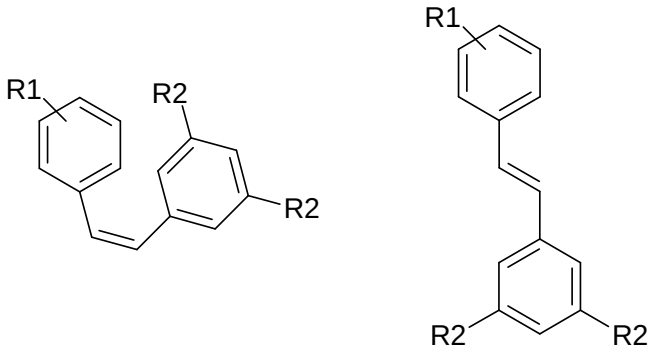
Fig1.1: Resveratrol [R_1=*4'OH*, R_2=OH**]**

Methodology

The AhR inhibitory activity data (k_i) of stilbene derivatives have been taken from the literature[**146**] and were converted to $-\log k_i$ for use in QSAR analysis as dependent parameter (table1.1). The physicochemical parameters viz. hydrophobic (π_1 and π_2), steric ($MR_1$and MR_2), electronic ($\sigma_1$andσ_2), resonance (RR_1 and RR_2) and field effect (f_1 and f_2) for the substituents at $R_1$and R_2 positions respectively were taken from literature[**156**]. In order to consider *cis* and *trans* series together, an indicator parameter (I) has been used with its value

I=1 for *trans* and I=0 for *cis*. These parameters are used as independent parameters in deriving QSAR equations. The parameters which showed some significant correlation with activity are described in table 1.1.

Table 1.1: Physicochemical parameters and AhR affinity of cis and trans stilbene derivatives.



C.No.	R_1	R_2	$®_2$	f_2	$-\log k_{i\ Cal.}$				$-\log k_{i\ obs}$
					eq. 3	**eq. 4**	**eq. 5**	**eq. 6**	
1	4'OMe	OMe	-0.04	0.52	-1.810	-1.799	-1.762	-1.751	-1.875
2	4'Cl	Cl	1.42	0.82	-1.376	-1.294	-1.399	-1.333	-1.114
3	4'OMe	F	0.28	0.86	-1.318	-1.688	-1.351	-1.659	-1.342
4	4'F	F	0.28	0.86	-1.318	-1.688	-1.351	-1.659	-1.799
5	4'CF_3	CF_3	1.76	0.76	-1.463	-1.777	-1.472	-1.235	-1.778
6	4'F	OMe	-0.04	0.52	-1.810	-1.799	-1.762	-1.751	-1.982
7	4'OEt	OMe	-0.04	0.52	-1.810	-1.799	-1.762	-1.751	-1.812
8	4'OBu	OMe	-0.04	0.52	-1.810	-1.799	-1.762	-1.751	-1.633
9	4'CF_3	Cl	1.42	0.82	-1.376	-1.294	-1.399	-1.333	-1.146
10	4'OMe	Cl	1.42	0.82	-1.376	-1.294	-1.399	-1.333	-1.079
11	3'CF_3	Cl	1.42	0.82	-1.376	-1.294	-1.399	-1.333	-1.278
12	3'OMe	Cl	1.42	0.82	-1.376	-1.294	-1.399	-1.333	-1.380
13	4'OMe	OMe	-0.04	0.52	-0.822	-0.810	-0.875	-0.857	0-.886
14	4'Cl	Cl	1.42	0.82	-0.388	-0.326	-0.512	-0.439	-0.079
15	4'OMe	F	0.28	0.86	-0.330	-0.700	-0.463	-0.765	-0.963
16	4'F	F	0.28	0.86	-0.330	-0.700	-0.463	-0.765	-0.579
17	4'CF_3	CF_3	1.76	0.76	-0.474	-0.188	-0.584	-0.342	-0.322
18	4'F	OMe	-0.04	0.52	-0.822	-0.810	-0.875	-0.857	-0.491
19	4'OEt	OMe	-0.04	0.52	-0.822	-0.810	-0.875	-0.857	-0.699
20	4'OBu	OMe	-0.04	0.52	-0.822	-0.810	-0.875	-0.857	-1.30
21	4'CF_3	Cl	1.42	0.82	-0.388	-0.306	-------	-------	0.698
22	4'OMe	Cl	1.42	0.82	-0.388	-0.306	-0.512	-0.439	-0.146
23	3'CF_3	Cl	1.42	0.82	-0.388	-0.306	-0.512	-0.439	-0.832
24	3'OMe	Cl	1.42	0.82	-0.388	-0.306	-0.512	-0.439	-0.756

Parameters were used to correlate the activity both independently and in combination of 2 or 3 parameters for each series (*cis*: compound number 1-12, *trans*: compound number.13-24) as well as for the combined series of both isomers by multiple linear regression analysis (MLR) using Systat 7.0 software. Among several parameters, the parameters which showed some correlations, both in individual and combined series are given in table 1.2.

Table –1.2: Pearson correlation matrix among different parameters

	f₂	**±₁**	**®₂**	**I**	**Jlogkᵢ**
f₂	1				
±₁	0.390	1			
®₂	0.702	0.628	1		
I	0	0	0	1	
Jlogkᵢ	0.323	0.313	0.384	0.764	1

Result and Discussion

The QSAR analysis using several physicochemical parameters, for the effect of substituents R_1 and R_2 at two different phenyl rings, either alone or in combination led to the identification of important parameters π_2, f_2, RR_2, σ_2 for *cis* series and MR_1, π_2, σ_2, f_2, σ_1 for *trans* series, with correlation coefficient r > 0.25. These parameters were used in QSAR analysis for their influence on activity alone or in combination. However, none of the combination of these parameters, keeping in view the orthogonality (r<0.5) among parameters used in the same equation, resulted in statistical significant QSAR equation in *trans* series. The parameters π_2 and f_2 showed moderate correlation of good statistical significance (>99%) %). **[$F_{1,10\ \alpha\ 0.01}$=7.95,**

$F_{1,10}$=8.244 for π_2 (eq.1) and $F_{1,10\ \alpha\ 0.01}$=7.95, $F_{1,10}$=9.358 for f_2 (eq.2)] for *cis* series.

$$-\log k_i = 0.295\ (\pm 0.103)\ \pi_2 - 1.746. \quad \text{---------- (Eq.1.1)}$$

n =12, *r* =0.672, F=8.244, s=0.256

$$-\log k_i = 1.515(\pm 0.495)\ f_2 - 2.612 \quad \text{---------- (Eq.1.2)}$$

n=12, *r* = 0.695, F=9.358, s=0.248

Where **'n'** is the no. of compounds used in the study, **r** is correlation coefficient, **F** is the variance ratio between the calculated and observed activities, and **s** is standard error of estimation.

In order to develop a combined QSAR model for *cis* and *trans* series, the indicator variable I was used in combination with above parameters. Among different combinations of these parameters, (>99.9%) the statistically significant eqs. with improved correlation (r>0.82) obtained were 1.3 and 1.4

$$-\log k_i = 1.447\ (\pm 0.545)\ f_2 + 0\ .988\ (\pm\ 0.158)\ I - 2.563. \text{----------(Eq.1.3)}$$

n=24, *r* =0.830, F =23.190, s =0.386

$$-\log ki = 0.346\ (\pm 0.102)\ \pi_2 + 0\ .988(\pm 0.146)\ I - 1.785 \text{----------(Eq.1.4)}$$

n=24, *r* =0.855, F =28.582, s =0.358

The careful analysis of both these equations led to the identification of compound no. 21 as an outlier, having the highest residual value (~3 times of the standard error) in the regression analysis. Hence its removal from the eq.1.3 and 1.4 led to the eq. 1.5 and 1.6 with significant

improved correlation values (r = 0.830, 0.855 in eq.1.3 and1.4 respectively and r= 0.859, 0.881 in eq.1.5 and 1.6 respectively).

$$-\log k_i = 1.210(\pm 0.430)\, f_2 + 0.888(\pm 0.126)\, I - 2.391 \text{---------- (Eq.1.5)}$$

n=23, *r* =0.859, F =28.272, s =0.301

==========J$\log k_i$ =0.286(ï0.081) ®$_2$ + 0.894(ï0.116) I J1.739---------- (Eq.1.6)

n =23, *r* =0.881, F =34.662, s =0.279

The outlier behaviour of the compound no.21 may be justified, as it is the only compound in the series with agonistic action on AhR while all other compounds are antagonist. It is pertinent to note that the affinity to the receptor does not normally differentiate between the agonist or antagonist which is normally based on functional assay but the QSAR analysis may differentiate because of subtle differences in the binding sites for the agonist and antagonist at the receptor Both the eqs. 1.5 and 1.6 are statistically significant but between these two equations, the eq.1.6 is superior to 1.5 both in terms of correlation coefficient value (r=0.85 and 0.88 for eq 1.5 and 1.6 respectively) and statistical significance. The graphs between calculated and observed activity of these two eqs. are given in Fig 1.2 and 1.3 respectively (r^2= 0.738 and 0.776 for eq.1.5 and 1.6 respectively) The cross validation for both the equations are also calculated (q^2=0.671 and 0.716 for eq.1.5 and 1.6 respectively, which proves the validity of the models .The r^2 and q^2 values of these eqs show that eq 1.6 is better than 1.5.The hydrophobicity of the substituent at R_2 position positively contributes to the

affinity of compound at the receptor along with *trans* spatial orientation of two phenyl rings. Further the similar regression coefficient values of indicator variable 0.88, (eq1.5) and 0.89 (eq.1.6) suggest the similar positive rate of contribution for activity with π_2 and f_2 parameters. The positive contribution of field effect of the substituent at R_2 position also cannot be ruled out because of high collinearity (r=0.7) between the hydrophobic and field effect in this series of compounds. Thus the affinity of antagonist in this series of compounds may be enhanced by the substituents with high hydrophobicity and high field effects keeping the *trans* spatial orientation of two phenyl rings.

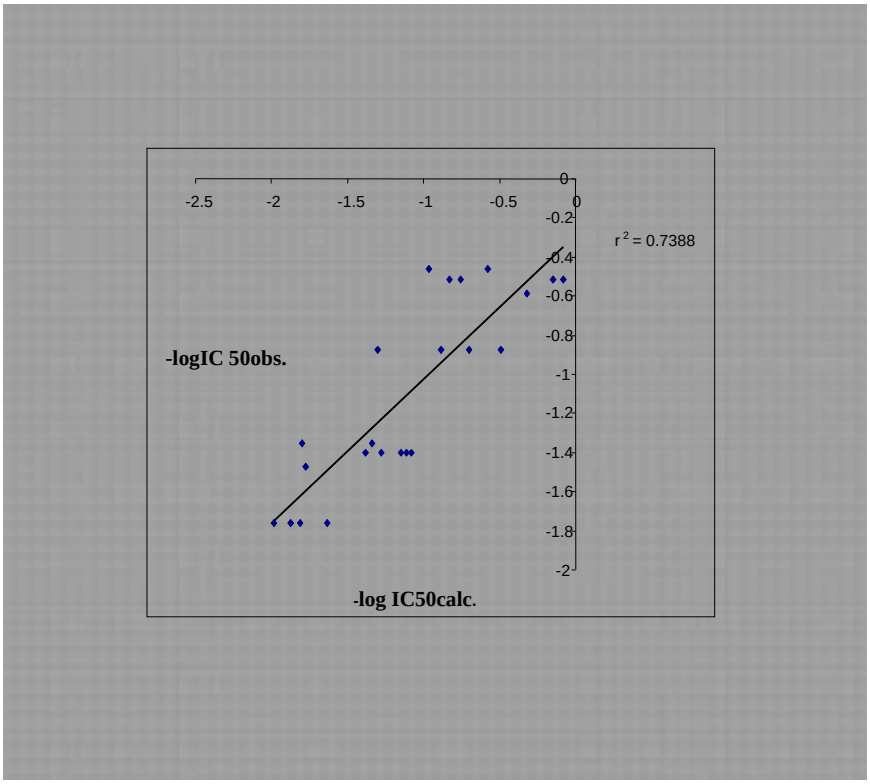


Fig: 1.2

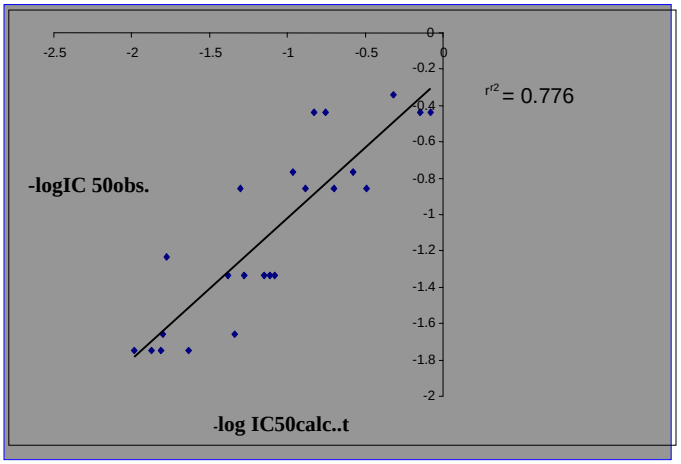


Fig: 1.3

Conclusion

The above studies highlight the importance of field effect and hydrophobicity of the substituents in one of the phenyl rings. In view of high intercorrelation between the two parameters (r=0.7), it is difficult to decide the role of each of these parameters, nevertheless it can be concluded that substituent with high field effects and hydrophobicity values will increase the affinity of the ligand to AhR.

2. QSAR studies on Hetaryl imidazoles derivatives as novel dual inhibitors of VEGF receptors I and II

Introduction

Vascular endothelial growth factor (VEGF), secreted by malignant tumor,is a key angiogenic factor.[157] Angiogenesis; formation of new vasculature from the existing vascular network[158], causes growth of solid tumor. Thus inhibitors of angiogenesis are considered to be novel therapeutic approach in oncology and ophthalmology[159,160] .A number of compounds, which inhibit the activity of VEGF, have been produced, as small molecule inhibitors inhibit VEGFR phosphorylation by directly competing with ATP binding site of respective intracellular kinase domain and finally lead to death of endothelial cells[157]. Here the QSAR[155] study on hetaryl imidazoles derivatives (Fig.2.1) has been reported to determine the physicochemical parameters that influence the anticancer activity of series.

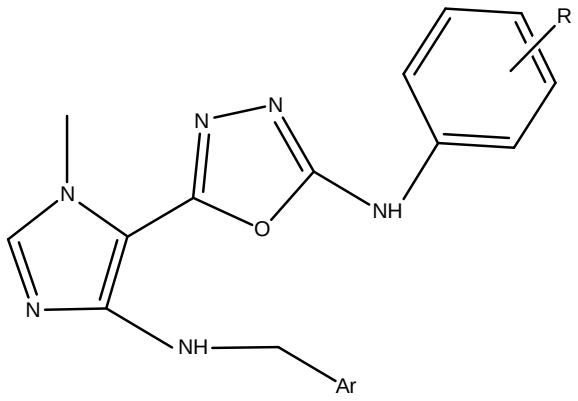


Fig.2.1: Hetaryl imidazoles derivatives

Methodology

The inhibitory activity data of hetaryl imidazole derivatives have been collected from literature [157] and calculated as –log BA.The values of physicochemical parameters for the substituents at R position are calculated as hydrophobic (π), steric (MR), electronic (σm,σp), resonance (RR) and field effect (f)[156].Different physico chemical parameters were considered as independent and inhibitory activity of hetaryl imidazole derivatives as dependent parameter in regression analysis for deriving QSAR equations, using Systat Software version 10.2. An indicator variable (I) having values "1"or' 0', is also introduced to indicate the presence or absence of 4-pyridine group at position Ar respectively. The parameters that showed some significant correlation during multi regression analysis are given in table. 2.1 along with the biological activities of all compounds of the series.

Result and Discussion

During regression analysis we considered 4-pyridine group as indicator parameter and assigned to it the value of '1' or '0' for its presence or absence at Ar position respectively. The parameters, which showed some significant correlation, are described in table 2.2

Table 2.1 Physicochemical parameters and inhibitory activity of Hetaryl imidazole Derivatives

C.No	R	^r :	±m	I	-logBA Cal		-logBA obs
					byeq.3	byeq.4	
1	4-Cl	4-Pyridine	0.00	1.00	0.532	0.632	0.602
2	4-Cl	Piperonyl	0.00	0.00	-0.294	-0.294	0.018
3	4-Cl	3,4-Di-F (C_6H_3)	0.00	0.00	-0.294	-0.294	-0.270
4	4-Cl	5-Indazole	0.00	0.00	-0.294	-0.294	-0.375
5	4-Cl	5-Quinoline	0.00	0.00	-0.294	-0.294	-0.549
6	4-t-Bu	4-Pyridine	0.00	1.00	0.532	0.632	0.420
7	4-i-Pr	4-Pyridine	0.00	1.00	0.532	0.632	0.357
8	4-ClF_2CO	4-Pyridine	0.00	1.00	0.532	0.632	0.886
9	4-F_3CO	4-Pyridine	0.00	1.00	0.532	0.632	1.032
10	4-F_3C	4-Pyridine	0.00	1.00	0.532	0.632	1.137
11	3-F_3CO	4-Pyridine	0.38	1.00	1.219	1.216	1.337
12	3-Me	4-Pyridine	-0.07	1.00	0.405	0.524	0.149
13	3-MeO	4-Pyridine	0.12	1.00	0.749	0.816	0.886
14	3,4 di Cl	4-Pyridine	0.37	1.00	1.200	1.201	0.745
15	3,4 di MeO	4-Pyridine	0.12	1.00	0.749	0.816	1.284
16	4-Cl, 3-CF_3	4-Pyridine	0.43	1.00	1.309	1.293	1.367
17	4-F, 4-Me	4-Pyridine	0.00	1.00	0.532	0.632	0.585
18	4-Br	4-Pyridine	0.00	1.00	0.532	0.632	0.137
19	4-Ph	4-Pyridine	0.00	1.00	0.532	-------	-0.508

Table –2.2: Pearson correlation matrix among different parameters

	f	**$\pm_m$**	**$\pm_p$**	**I**	**JlogBA**
f	1				
$\pm_m$	0.586	1			
$\pm_p$	0.235	0.001	1		
I	0.175	0.252	0.179	1	
JlogBA	0.200	0.569	0.209	0.658	1

Since no satisfactory results were obtained by linear regression analysis, multiregression analysis was performed. Thus several permutations and combinations of the above parameters were tried keeping in view the inter correlation($r<0.5$) among the parameters used in the same equation.

$$-\log BA = 0.702(\pm 0.374)f + 1.073(\pm 0.260)I - 0.582 \quad \text{---------- (eq.2.1)}$$

n =19 r =0.732 F=9.229 s=0.454

$$-\log BA = 1.028(\pm 0.523)\sigma_p + 1.079(\pm 0.258)I - 0.530 \quad \text{---------- (eq.2.2)}$$

n =19 r =0.737 F =9.526 s=0.451

$$-\log BA = 1.808(\pm 0.679)\sigma_m + 0.825(\pm 0.243)I - 0.294 \quad \text{---------- (eq.2.3)}$$

n =19 r =0.779 F =12.366 s=0.418

Where **'n'** is the no. of compounds used in the study, **r** is correlation coefficient, **F** is the variance ratio between the calculated and observed activities, and **s** is standard error of estimation.

On analyzing the above equations, equation 2.3 was found to be statistically significant (>99%) with moderate correlation coefficient (r> 0.77)

On careful analysis of equation 2.3, compound number 19 was identified as an outlier having highest residual value (~3 times of the standard error) in the regression analysis. Thus the removal of compound number 19 from equation 2.3 led to equation 2.4 with significant improved correlation coefficient (r> 0.85) The graph between calculated and observed activity of eq no 2.4 is given in Fig:2.2 (r^2= 0.6826)

-JlogBA = NKRPTE ï MKRPVF ±$_{m}$ H-MKVOSE ï MKNVPF**I**-J-MKOVQ======---------- **(eq.2.4)**

n =18 *r* =0.857 F =20.785 s=0.328

Since the coefficient of parameters in eq.4 possess positive sign which indicates their positive contribution to activity. Thus substitution (R) by group or atom, at meta position of benzene ring and presence of 4-pyridine group at Ar position play a crucial role in enhancing the activity of given series on drug molecule.

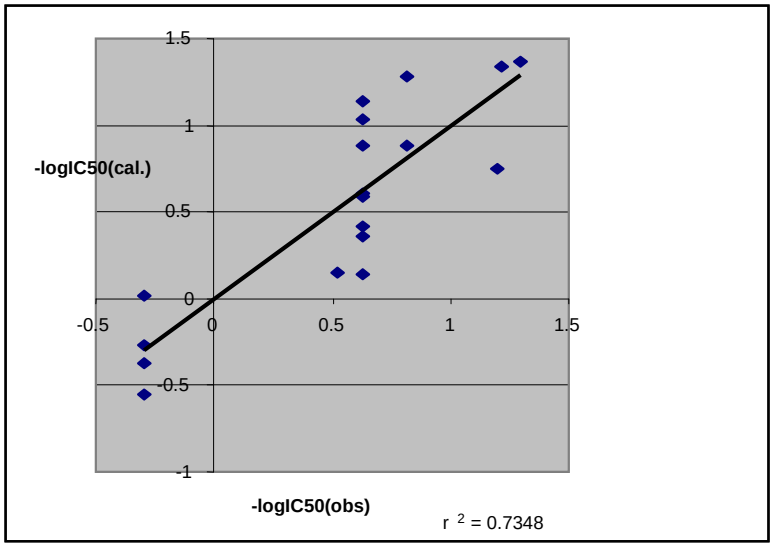


Fig: 2.2

Conclusion

The above QSAR studies on Hetaryl imidazole derivatives highlight the important parameters which affect its activity and thus, led us to know that substitution at meta position of benzene ring (R) by atom or group contribute positively to activity and presence of 4- pyridine group at Ar position will also be helpful in designing new drug molecule or enhancing the activity of given series of compounds.

3. 2D-QSAR studies on Carbazole Sulfonamide derivatives as Antimitotic Agents Against solid Tumors

Introduction

Antimitotic agents cause cell death and regress tumors by arresting cell cycle at G_2/M phase.[161,162] The classical antimitotic agents as toxins and Vinca alkaloids are tubulin-binding agent. However they have limited use due to their difficult synthesis, development of multidrug résistance, and neurotoxicity[163]. Combrestatin A4 **(Fig:3.1)**, a novel small molecule, binds to colchicines binding site, and inhibits tubulin polymerization.[164] Due to low water solubility of Combrestatin A4, its use is limited. Studies have shown that 3,4,5-trimethoxy substituent at ring A are important for its activity.[165,166] Studies are going on to improve activity of Combrestatin A4 by replacing linker group and ring B.[167,168] Here QSAR studies on carbazole sulphonamide derivatives **(Fig:3.2)** of Combrestatin A4, which is obtained by replacing the olefin linker with sulphonamide group, ring B with carbazole moiety and replacing 3,4,5-trimethoxy substituent at ring A with various subsituent [169]has been reported. The studies have been carried out in order to gain an insight to the essential structural and physicochemical parameters, that affect the inhibitory activity of Carbazole sulfonamide derivatives, considering antimitotic activity as dependent and different physicochemical parameters like hydrophobic (π), steric (MR),electronic (f, R, σ) and indicators as independent parameters in the analysis. .

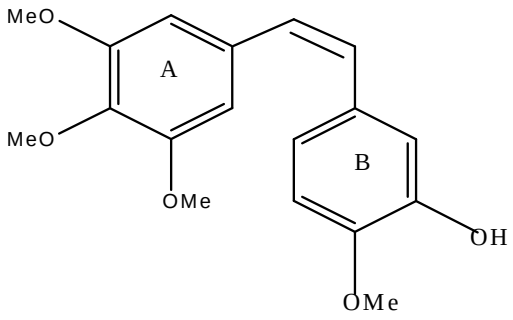


Fig: 3.1

Methodology

For QSAR analysis the antimitotic activity data (IC_{50}) of Carbazole sulfonamide derivatives, obtained from literature [169] and converted to $-\log IC_{50}$, is used as dependent parameter. The values of physicochemical parameters as hydrophobic (π), steric (MR), electronic ($\sigma m, \sigma p$), resonance effect(RR) and field effect (f) for the substituents at R position are calculated[156] and considered them as independent parameters in regression analysis[155],for obtaining QSAR equations, using Systat Software version 10.2. An indicator variable (IR_1) having values "1"or' 0', is also introduced to indicate the presence or absence of Et group respectively at position R_1 Another indicator parameter (IY) having value "1" for the presence of SO_2NH group at position "Y" and "0" for its absence, was also used in the analysis All the compounds of the series and their biological activities along with the parameters that showed significant correlation with activity are given in table 3.1.

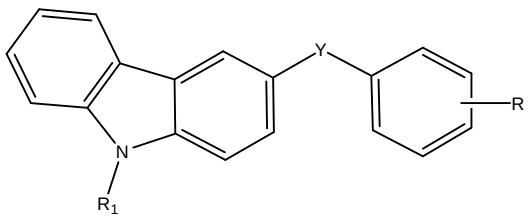


Fig: **3.2**

TableJ3.1: Physicochemical parameters and antimitotic activity of Carbazole sulfonamide derivatives

C.no	R	IR_1	IY	f	±p	$-\log k_{i\ cal.}$		$-\log k_{i\ obs}$
						Eq. 11	Eq12	
1	3,4,5- OMe_3	Et	SO_2NH	0.780	0.270	--------	0.437	-1.748
2	3,4,5- OMe_3	Me	SO_2NH	0.780	0.270	--------	0.523	-1.663
3	3,4,5- OMe_3	H	SO_2NH	0.780	0.270	-0.019	-0.597	-2.782
4	3,4 -OMe_2	Et	SO_2NH	0.520	0.270	-0.623	-0.505	-3.387
5	3,4- OCH_2O	Et	SO_2NH	1.150	0.100	-0.626	-------	-3.796
6	3,5- OMe_2	Et	SO_2NH	0.520	0.000	0.420	-0.107	-2.989
7	2,4 -OMe_2	Et	SO_2NH	0.520	0.540	0.362	------	-1.756
8	2,5- OMe_2	Et	SO_2NH	0.520	0.270	--------	-------	-1.785
9	2,4 6- OMe_3	Me	SO_2NH	0.780	0.810	-0.373	0.340	-1.845
10	4 -F	Et	SO_2NH	0.430	0.060	-0.071	-0.214	-3.337
11	3-Cl,4- OMe	Et	SO_2NH	0.670	0.270	0.080	-0.203	-2.683 -
12	5-Cl, 2,4- OMe_2	Et	SO_2NH	0.930	0.540	0.291	-0.042	-1.826
13	4-Cl, 2,5 -OMe_2	Et	SO_2NH	0.930	0.500	0.465	0.036	-1.748
14	4-Cl, 2,5- OMe_2	Me	SO_2NH	0.930	0.500	0.550	0.121	-1.663
15	3 -NH_2,4 -OMe	Et	SO_2NH	0.280	0.270	-0.579	0.182	-3.342
16	3,4-- OMe_2	Et	NH SO_2	0.520	0.270	-0.660	-0.542	-3.423
17	2,5 OMe_2	Et	NH SO_2	0.520	0.270	-0.017	0.101	-2.780
18	3- OMe	Et	NH SO_2	0.260	0.000	0.512	0.681	-2.897
19	3-CN	Et	NH SO_2	0.510	0.000	0.289	-0.212	-3.120

Results and Discussion

During the study, among several physicochemical parameters *viz*. π, MR, f, R, σ_m, σ_p IR_1 and IY, the parameters π (r=0.577), MR(r=0.592),f (r = 0.378), σ_p (r=0.665) and IR_1 (r=0.402) showed better correlation with activity than the remaining parameters and were thus considered for further study in the thesis, also described in table 3.2.

Table –3.2: Pearson correlation matrix among different parameters

	®=	**MR**	**f**	**± p**	**IR_1**	**JlogIC$_{50}$**
®	1					
MR	0.202	1				
f	0.203	0.690	1			
± p	0.388	0.636	0.435	1		
IR_1 =	0.159	0.546	0.392	0.431	1	
JlogIC$_{50}$	0.577	0.592	0.378	0.665	0.402	1

We performed linear regression analysis of the above parameters and obtained several models.

–log IC_{50} = 0.775(±0.266) π–0 .2521 ---------- (eq.3.1)

n =19 *r* = 0.577 F =8.502 s =0.630

–log IC_{50} = 2.323(±0.633) σ p– 3.226---------- (eq3.2)

n =19 *r* = 0.665 F =13.447 s =0.577

–log IC_{50} = 0.078(±0.026) MR– 4.078---------- (eq3.3)

n =19 *r* = 0.592 F =9.151 s =0.622

$-\log IC_{50} = 0.720(\pm 0.398)\ IR_1 - 1.988$---------- (eq.3.4)

n= 19 r = 0.402 F =3.276 s =0.701

$-\log IC_{50} = 1.225(\pm 0.728)\ f - 3.348$---------- (eq.3.5)

n =19 r = 0.378 F =3.785 s =0.714

Where **'n'** is the no. of compounds used in the study, **r** is correlation coefficient, **F** is the variance ratio between the calculated and observed activities, and **s** is standard error of estimation.

None of the above models were found to be satisfactory since they possess low correlation coefficient, low F value, high stantdard error value.We,then carefully analysed these models and found certain outliers in them ,having highest residual value (~3 times of the standard error) in the regression analysis. In model 3.1,which is based on hydrophobic parameter, compound no. 10 was found to be as an outlier.Similarly in model.no. 3.2 compound no.1 and in model.no 3.5, compound no.5 was identified as an outlier.Removal of these outliers from the particular equations resulted in equations 3.6,3.7,3.8, with improved correlation coefficient (r>0.60) and stasistical significance.

$-\log IC_{50} = 0.807(\pm 0.256)\ \pi - 2.468$ ---------- (eq 3.6)

n =18 r = 0.620 F =9.967 s =0.604

$-\log IC_{50} = 2.345(\pm 0.597)\ \sigma\ p - 3.285$---------- -- (eq3.7)

n=18 r = 0.701 F =15.418 s =0.543

$-\log IC_{50} = 2.377(\pm 0.599)\ f - 3.964$---------- (eq.3.8)

n =18 r = 0.704 F =15.765 s =0.517

The maximum improvement in correlation coefficient was observed in models based on electronic parameter σ p (eq.no.3.2) and field effect parameter (eq.no.3.5), which **indicated that f and ±ᵺp can be important parameters**, affecting the activity of given series of drug molecules Thus,further studies were targeted on these two parameters and performed the reanalysis of the models based on these parameters (eqs.3.7 and 3.8) This time we identified compound no. 1 (in eq .3.7) and 7 (in eq 3.8) as outliers. Their removal led t o eqs.3.9 and 3.10 respectively, with improved correlation coefficient (r>0.70).

$-\log IC_{50} = 2.367(\pm 0.558)\ \sigma p - 3.344$---------- -- (eq3.9)

n=17 r = 0.739 F =18.024 s =0.507

$-\log IC_{50} = 2.518(\pm 0.544)\ f - 4.110$---------- (eq3.10)

n =17 r = 0.767 F =21.432 s =0.467

On reanalysis of these models compound no. 8 was identified as an outlier in both the models. Its removal from eqs3.9.and 3.10 resulted in eqs.3.11 and 3.12 with significant improved correlation values (r = 0.739, 0.767 in eq.3.9 and 3.10 respectively and r= 0. 784, 0. 841 in eq.3.11 and 3.12 respectively).

$-\log IC_{50} = 2.392(\pm 0.506)\ \sigma p - 3.409$---------- (eq3.11)

n=16 r = 0.784 F =22.375 s =0.460

Jlog IC_{50} = 2.678(ï 0.461) f– 4.274---------- (eq3.12)

n =16 *r* = 0.841 F =33.739 s =0.392

Both of these eqs. are statistically significant. Since the parameters present in eq.3.11and 3.12 possess positive sign before their coefficients, which indicates that these parameters contribute positively to the activity. Among these two equations, eq. 3.12 is found to be superior, both in terms of correlation coefficient ((r=0.78 and 0.84 for eq 3.11 and 3.12 respectively) and statistical significance. The graphs between calculated and observed activity of these two equations are given in Fig 3.3 and 3.4 respectively (r^2=0.615 and 0.706 for eq.3.11 and 3.12 respectively) which proves the validity of the models and shows that equation 3.12 is better than 3.11.The equation **3.12** thus obtained above is statistically significant as it has good correlation coefficient (r = 0.84) of high (>99.9 %) statistical significance **($F_{1,14\sim\ 0.001}$ = 17.14; $F_{1,14\ 0.001}$ = 33.739)**.Therefore the activity of given series of drug molecules can be increased by substituting group at position R, having high field effect values.

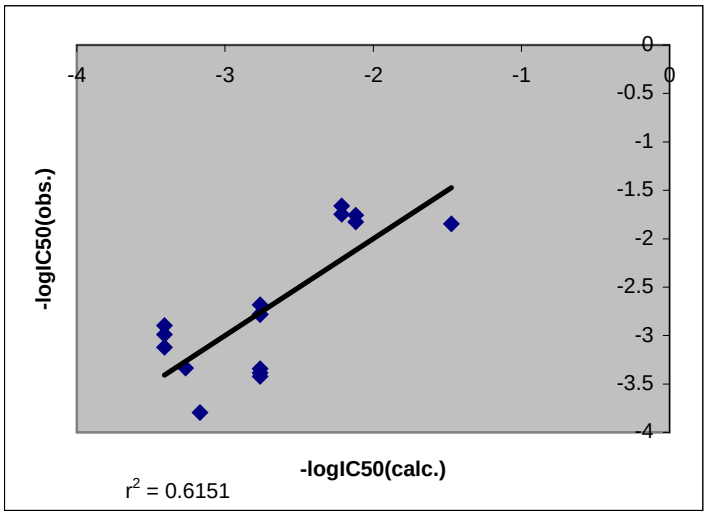


Fig: 3.3

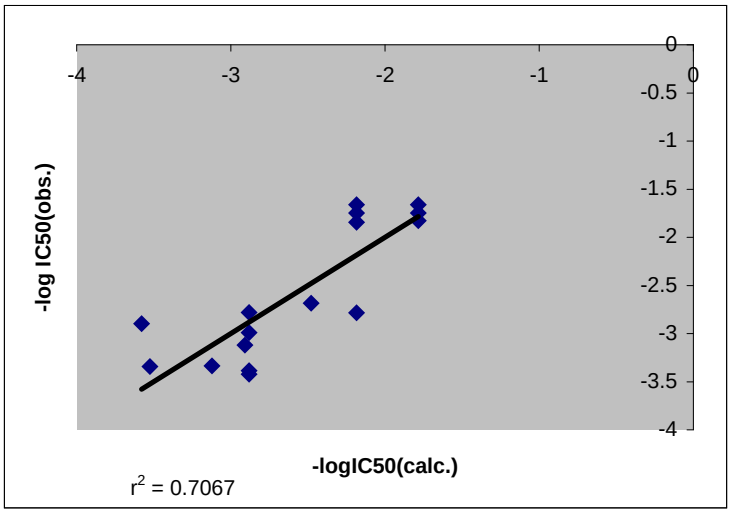


Fig: 3.4

Conclusion

The above QSAR studies on Carbazole sulphonamide derivatives highlight the **importance of field effect parameter** On the basis of these studies it is inferred that, for increasing the activity of given series of compounds or designing the new chemical entities for cancer, the substituents with high field effect values will be preferred at position R.Thus it can be concluded that **substituents with high field effect values** will increase the antimitotic activity of Carbazole sulfonamide derivatives against solid tumors.

4. QSAR studies on 4, 5, 6, and 7- Tetrabromobenzimidazole Derivatives as inhibitors of Protein kinase CK

Introduction

Protein Kinase plays a vital role in controlling mostly all the cellular functions, specially signalling transduction path[170]. They act as catalyst during the transfer of γ phosphate of ATP to serine, threonine and tyrosine residue of protein substrates. They are activated by the presence of specific stimuli. The unscheduled activation of protein kinase caused by genetic rearrangements and mutations, results in several pathologies as neoplasia[171]**.** Casien Kinase2 (CK2) is a pleiotropic protein kinase having 300-protein substrate[172] due to which it has lack of control over catalytic activity. CK2 is essential in several important cell functions[173].Several evidences show that its catalytic subunit acts as oncogenes[174-177],displays antiapoptotic effect in prostrate cell lines[178] .Thus CK2 has been identified as an important target for antineoplastic drugs. A number of CK2 inhibitors have been developed as Emodin, TBB (4,5,6,7- tetrabromobenzotriazole) and IQA (5-oxo-5, 6-dihydroxyindolo- [1,2-α] quinazolin-7-ylacetic acid). Among these inhibitors TBB is found to be the most effective .Its potency is higher than other CK2 inhibitors, but not outstanding. This chapter contains the QSAR study on 4,5,6,7- tetrabromobenzimidazole derivative in order to gain an insight to the essential structural and physicochemical parameters, that affect the inhibitory activity of 4,5,6,7 tetrabromo benzimidazole derivatives **(Fig:4.1).** 4,5,6,7- tetrabromobenzimidazole is obtained by replacing triazole ring of TBB with imidazole moiety. The QSAR[155]studies have been carried out considering anticancer activity as dependent and different physicochemical

parameters like hydrophobic (π), steric (MR) and electronic (f, R, σ) as independent parameters .

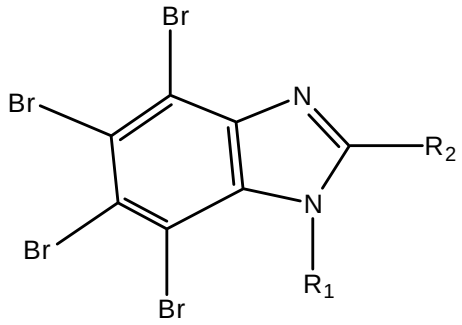


Fig: 4.1

Methodology

A series of 22 molecules reported in literature[170] having the prototype **1** were selected for the QSAR studies in order to find the influence of physicochemical parameters *viz.* hydrophobicity (π); electronic effect (σ-Hammet); resonance effect (RR); field effect (f) and steric effect (MR) for the substituent $R_1$and R_2 of prototype **1** on protein kinase CK2 inhibitory activity. During QSAR analysis the inhibitory activity data (IC_{50}) of 4,5,6,7 tetrabromobenzimidazole derivatives has been taken from the literature[170] and were converted to -log IC_{50} for use as dependent parameter **(table4.1).** The values for the physicochemical parameters were taken from the literature [156] and are used as independent parameter. The multiparameter regression analysis was carried out using SYSTAT software. The parameters which showed significant correlation during analysis are given in table. 4.1 along with the biological activities of all compounds of the series.

Result and Discussion

Among different physicochemical parameters *viz.* π_1, MR_1 f_1 RR_1, σ_1 π_2, MR_2 f_2 RR_2, σ_2 the parameters $\boldsymbol{\pi_1}$(r=0.004), MR_1 (r = 0.173), $\mathbf{f_1}$(r = 0.132), $\mathbf{RR_1}$ (r = 0.069),

±₁(r=0.058) and **MR₂**(r=0.159)showed the **least correlation with activity** and were thus not considered for further study. The parameters **®₂** (r =0.177), **f₂**(r = 0.327) **RR₂** (r = 0.221) and **± ₂**(r = 0.218) showed **better correlations with activity** than the above parameters, and are described in table 4.2.These parameters were considered for further study.

Table4.1: Physicochemical parameters and inhibitory activity of 4,5,6,7-tetrabromo benzimidazole derivatives

C.No.	R_1	R_2	±₂	f_2	-logIC$_{50cal}$		-logIC$_{50obs}$
					eq. 8	eq. 9	
1	H	Br	0.23	0.44	0.177	0.116	0.131
2	H	NH_2	-0.66	0.02	0.797	0.817	0.523
3	H	$NHCH_3$	0.84	-0.11	0.607	0.587	0.678
4	H	N $(CH3)_2$	-0.83	0.10	0.752	0.769	0.854
5	H	$NHCH_2CHOH$	0.84.	-0.11	.---------	.------	-1.431
6	H	NHCH $(CH_3)_2$	0.54	-0.19	0.753	0.754	0.796
7	H	$NHCH_2CH(OH)C$	0.67	-0.15	0.684	0.675	0.553
8	H	$NHCH_2CH_2N(CH_3)$	0.52	-0.05	0.615	0.600	0.602
9	CH_3	Br	0.23	0.44	0.177	0.116	0.367
10	$CH_2CH=CH_2$	Br	0.23	0.44	0.177	0.116	0.056
11	$CH_2C(O)NH2$	Br	0.23	0.44	0.177	0.116	-0.196
12	CH_3	$OCH(CH_3)_2$	-0.45	0.30	--------.	.-------	-1.309
13	CH_3	=O	-0.37	0	0.459	0.437	0.409
14	CH_3	$N(CH_3)_2$	-0.83	0.10	0.752	0.769	0.886
15	CH_3	$NHCH(CH_3)_2$	0.54	-0.19	0.753	0.754	0.745
16	H	SCH_3	0.00	0.20	0.471	0.446	0.602
17	H	SCH_3	0.00	0.20	0.471	0.446	0.222
18	H	SCH_2 $(C_6H_4)NO_2$	0.77	0.95	-0.460	-0.599	-0.542
19	H	SCH_2COOH	-0.45	0.53	0.037	.-------	0.602
20	H	$SCH_2CH(OH)CH_2$	0.00	0.20	0.470	0.446	0.824
21	H	$SCH_2COOC_2H_5$	0.45	0.53	0.037	-0.042	-0.190
22	H	=O	-0.37	0.29	0.459	0.437	0.444

Table –4.2: Pearson correlation matrix among different parameters

	®$_2$	**MR$_2$**	**f$_2$**	**RR$_2$**	± $_2$	**-logIC$_{50}$**
®$_2$	1					
MR$_2$	0.391	1				
f$_2$	0.598	0.231	1			
RR$_2$	0 .306	-0.027	0.735	1	1	
± $_2$	0.171	0. 454	-0. 023	0.068	0.523	
-logIC$_{50}$	-0.177	-0.159	-0.327	-0.221	-0.218	1

First of all linear regression analysis of the above parameters was carried out but satisfactory results were not obtained. Hence different combinations of the above independent parameters were used for correlating the activity in QSAR analysis (LFER model), keeping in view the inter correlation($r<0.5$) amongst parameters used in same equation.

$-\log IC_{50} = -0.006\ (\pm 0.014)\ MR_2 - 0.638\ (\pm 0.494)\ f_2 + 0.492$ ---------- (eq.4.1)

n = 22 r = 0.338 F =1.226 s =0.639

$-\log IC_{50} = -0.081\ (\pm 0.158)\ \pi_2 - 0.267\ (\pm 0.338)\ RR_2 + 0.125$ ---------- (eq.4.2)

n = 22 r = 0.249 F =0.627 s =0.658

$-\log IC_{50} = -0.300\ (\pm 0.317)\ RR_2 - 0.251(\pm 0.270)\ \sigma_2 + 0.126$ ---------- (eq.4.3)

n = 22 r = 0.301 F =0.945 s =0.648

$-\log IC_{50} = -0.740\ (\pm 0.469)\ f_2 - 0.278\ (\pm 0.259)\ \sigma_2 + 0.451$ ---------- (eq.4.4)

n = 22 r = 0.348 F =1.7 88 s =0.623

Where **'n'** is the no. of compounds used in the study, **r** is correlation coefficient, **F** is the variance ratio between the calculated and observed activities, and **s** is standard error of estimation.

On analyzing the above equations, compound no.5 and 12 were identified as outliers in all the models having residual value three times more than the standard error. Thus the removal of these compounds from equations resulted in equations 4.5, 4.6, 4.7 and 4.8 respectively with significant improved correlation coefficient (r> 0.6)

$-\log IC_{50} = -0.006\ (\pm 0.006)\ MR_2 - 1.025\ (\pm 0.205)\ f_2 + 0.690$ ---------- (eq.4.5)

n = 20 r = 0.788 F =13.901 s =0.256

$-\log IC_{50} = -0.216\ (\pm 0.074)\ \pi_2 - 0.354\ (\pm 0.145)\ RR_2 + 0.283$ ---------- (eq.4.6)

n = 20 r = 0.743 F =10.488 s =0.278

$-\log IC_{50} = -0.465\ (\pm 0.158)\ RR_2 - 0.222\ (\pm 0.144)\ \sigma_2 + 0.202$ ---------- (eq.4.7)

n = 20 r = 0.641 F =5.929 s =0.319

$-\log IC_{50} = -1.021\ (\pm 0.180)\ f_2 - 0.215\ (\pm 0.104)\ \sigma_2 + 0.675$ ---------- (eq.4.8)

n = 20 r = 0.833 F = 19 .215 s =0. 237

On the analysis of equations 4.5, 4.6,4.7 and 4.8, equation 4.8 was found to be statistically significant (>99%) with moderate correlation coefficient (r> 0.83) On analyzing equation 4.8 carefully; compound number 19 was identified as an outlier. Its removal from eq.4.8 led to equation 4.9 with improved correlation coefficient (r> 0.90)

$\log IC_{50} = -1.136\ (\pm 0.147)\ f_2 - 0.252\ (\pm 0.083) \pm_2 + 0.674$ ---------- (eq.4.9)

n = 19 r = 0.902 F = 35.114 s =0. 184

The equation **4.9** thus obtained above is statistically significant as it has good correlation coefficient (r = 0.902) of high (>99.9 %) statistical significance **($F_{2,16}$ ~ $_{0.001}$ = 10.97; $F_{2,16\ 0.001}$ = 35.114)** Since in eq.4.9 the coefficient of parameters f_2 and σ_2 possess negative sign which indicates their negative contribution to activity. The negative contribution by f_2 and σ_2 suggests that the substitutions at R_2 position by the molecules with less field effect value and less electronic effect (σ) value should be preferred for enhancing the activity of given series on drug molecule The graph between calculated and observed activity of eq no4.9 is given in Fig.4.2 **(r^2= 0.814)** The cross validation for the equation is also calculated **(q^2=0.732)** ,which also proves the validity of this model

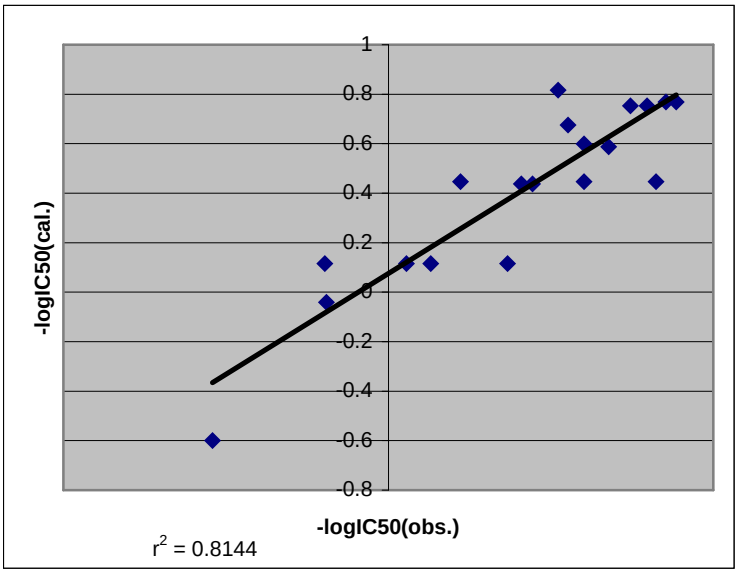


Fig: 4.2

Conclusion

The above QSAR studies on 4,5,6,7 tetrabromobenzimidazole derivatives highlight the important parameters which affect its activity The results led us to know that **substitution at R_2 position by atom or group having less field effect (f) and less electronic value (±)** will be helpful in designing new drug molecule or enhancing the activity of given series of compounds. Hence, these results will be useful in optimizing the activity in this class of molecules and also in designing new molecules.

5. 2D-QSAR studies on N-Phenyl-N^{I}{4-(4-quinolyloxy) phenyl} urea derivatives as VEGFR-2 Tyrosine Kinase Inhibitors

Introduction

Angiogenesis is the process of new blood vessels growth and involves the proliferation of endothelial cells in response to specific growth stimuli such as Vascular Endothelial Growth Factor (VEGF)[179]. VEGF, which is secreted by malignant tumors, is a key angiogenic factor. It causes migration of vascular endothelial cells[180,181]. Therefore inhibition of angiogenesis is a novel therapeutic approach against tumor. In view of it, development of inhibitors for VEGFR- 2 represents a challenge for the development of prophylactic as well as curative drug for major diseases involving VEGF activity Several compounds as antibodies to VRGF[182], its receptor, small molecule receptor tyrosine kinase inhibitor such as 3-substituted indolinones, 4-anilinoquinazolines and anilinophthalazines which inhibit the biological activity of VEGF have been produced. Here the QSAR study on N-Phenyl-N^{I} {4-(4-quinolyloxy) phenyl} urea derivatives (fig. 5.1) has been reported using linear free energy relationship (LFER)[155]approach to understand the role of different physico chemical and structural parameters on its biological activity.

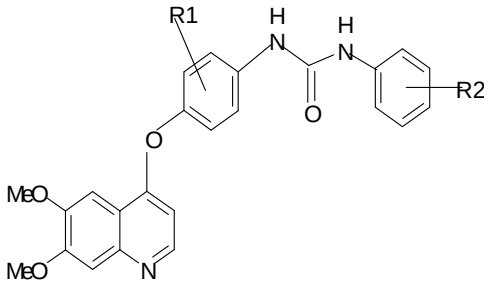


Fig: 5.1

Methodology

All the inhibitory activity data(IC_{50}) of N-Phenyl-N'{4-(4-quinolyloxy) phenyl} urea derivative have been collected from literature[183]which is expressed as $-\log IC_{50}$. The values of physicochemical parameters for the substituents (R_1 and R_2) were taken from literature[156] and are calculated as hydrophobic (π_1 and π_2), steric ($MR_1$and MR_2), electronic (σm_1,σp_1 and σm_2 σp_2), resonance (RR_1 and RR_2) and field effect ($f_1$andf_2).For deriving QSAR equations, we used inhibitory activity of N-Phenyl-N'-{4-(4-quinolyloxy) phenyl} as dependent and different physico chemical parameters as independent parameter Multiparameter regression analysis was performed using Systat Software version 10.2 All the compounds of the series and their biological activities along with the parameters that showed some significant correlation with activity are given in table 5.1

Table JR1 Physicochemical parameters and inhibitory activity of N-Phenyl-N' {4-(4-quinolyloxy) phenyl} urea Derivatives

C.No.	R_1	R_2	j R_2	f_2	JlogIC$_{50Cal}$			JlogIC$_{50obs}$
					byeq.4	byeq.5	byeq.6	
1	H	H	5.150	0.00	0.773	0.701	0.641	0.700
2	H	2-OMe	11.990	0.260	-0.122	-------.	---------.	0.700
3	H	4-OMe	11.990	0.260	-0.122	-0.203	-0.264	-0.040
4	H	2-Me	9.770	-0.040	0.374	0.275	0.200	-0.360
5	H	3-Me	9.770	-0.040	0.374	0.275	0.200	0.700
6	H	4-Me	9.770	-0.040	0.374	0.275	0.200	0.300
7	H	2-NO_2	11.480	0.670	-0.466	-0.502	-0.536	-1.280
8	H	3-NO_2	11.480	0.670	-0.466	-0.502	-0.536	-0.040
9	H	4-NO_2	11.480	0.670	-0.466	-0.502	-0.536	-0.750
10	H	2-F	5.040	0.430	0.372	0.346	0.312	0.300
11	H	3-F	5.040	0.430	0.372	0.346	0.312	-0.080
12	H	4-F	5.040	0.430	0.372	0.346	0.312	0.400
13	H	3-Cl	10.150	0.410	-0.092	-0.148	-0.194	-0.200
14	H	4-Cl	10.150	0.410	-0.092	-0.148	-------.	0.700
15	H	2,3-F_2	4.930	0.860	-0.029	-0.010	-0.017	0.150
16	H	2,4-F_2	4.930	0.860	-0.029	-0.010	-0.017	0.150
17	H	2,5-F_2	4.930	0.860	-0.029	-0.010	-0.017	0.400
18	H	2,6-F_2	4.930	0.860	-0.029	-0.010	-0.017	-0.260
19	H	3,4-F_2	4.930	0.860	-0.029	-0.010	-0.017	-0.040
20	H	3,5-F_2	4.930	0.860	-0.029	-0.010	-0.017	-0.260
21	H	2,3-Cl_2	15.150	0.820	-0.957	-0.997	-1.030	-1.430
22	H	2,4-Cl_2	15.150	0.820	-0.957	-0.997	-1.030	-0.750
23	H	2,5-Cl_2	15.150	0.820	-0.957	-0.997	-1.030	-0.780
24	H	2,6-Cl_2	15.150	0.820	-0.957	-0.997	-1.030	-1.570
25	H	3,4-Cl_2	15.150	0.820	-0.957	-0.997	-1.030	-0.570
26	H	3,5-Cl_2	15.150	0.820	-0.957	-0.997	-1.030	-0.950
27	2-F	2,4-F_2	4.930	0.860	-0.029	-0.010	-0.017	0.050
28	2-Cl	2,4-F_2	4.930	0.860	-0.029	-0.010	-0.017	0.400
29	3-Cl	2,4-F_2	4.930	0.860	-0.029	-0.010	-0.017	-0.410

Result and Discussion

The parameters, which showed some correlation, are described in table 5.2.

Table –5.2: Pearson correlation matrix among different parameters

	MR_2	**f_2**	**R_2**	**$\pm p_2$**	**$JlogIC_{50}$**
MR_2	1				
f_2	0.013	1			
R_2	0.624	0.235	1		
$\pm p_2$	0.431	0.326	0.022	1	
$JlogIC_{50}$	0.604	0.473	0.494	0.523	1

Linear regression analysis of the above parameters has been performed but could not obtain satisfactory results. Then multi regression analysis was carried out.

Several permutations combinations of the above parameters were tried, keeping in view the inter correlation($r<0.5$) amongst parameters used in the same equation.

$-\log IC_{50} = -0.072(\pm 0.025)MR_2 - 0.096(\pm 0.454)\sigma p_2 + 0.629$ ---------- (eq5.1)

n =29 r =0.671 F=10.628 s=0.492

$-\log IC_{50} = -1.480(\pm 0.325)f_2 - 1.295(\pm 0.374)R_2 + 0.257$ ---------- (eq5.2)

n =29 r =0.684 F=11.452 s=0.484

$-\log IC_{50} = -0.583(\pm 0.156)\pi_2 - 1.502(\pm 0.379)\ \sigma p_2 + 0.342$ ---------- (eq5.3)

n =29 r = 0.727 F=14.567 s=0.456

$-\log IC_{50} = -0.095(\pm 0.019)MR_2 - 0.957(\pm 0.248)\ f_2 + 1.260$ ---------- (eq5.4)

n =29 r =0.772 F=19.199 s=0.422

Where **'n'** is the no. of compounds used in the study, **r** is correlation coefficient, **F** is the variance ratio between the calculated and observed activities, and **s** is standard error of estimation.

On analyzing the above equations, eq.5.4 was found to be the best fit model with statistical significance (>99.9%) and moderate correlation coefficient(r>0.77) On carefully analyzing eq.no.5.4 compound no 2 was found as an outlier in regression analysis, having high residual value (~3 times more than the standard error). Its removal from eq.5.4 resulted in eq.5.5 with improved correlation value. (r>0.79).

$-\log IC_{50} = -0.100(0.018)MR_2 - 0.854(0.237)f_2 + 1.216$ ---------- (eq5.5)

n =28 *r* =0.798 F=21.882 s=0.394

Again, compound no.14 was identified as an outlier and its removal from eq 5.5 led to the eq.5.6

JlogIC$_{50}$= JMKNMOEïMKMNTFMR_2JMKTVNEïMKONUFf_2+1.167---------- (eq5.6)

n =27 *r* =0.826 F=25.808 s=0.361

Thus eq.5.6 was found to be the best both in terms of correlation value(r>0.82) and statistical significance (>99.9%) The graph between calculated and observed activity of eq no.5.6 is given in fig 5.2 (**r^2= 0.826)**. Since the parameters present in eq.5.6 possess negative sign before their coefficients, which indicates that these parameters contribute negatively to the activity. Thus substitutions at R_2 position by the group possessing less, field effect and molar refractivity (less bulky group) will enhance the activity of given series of compound

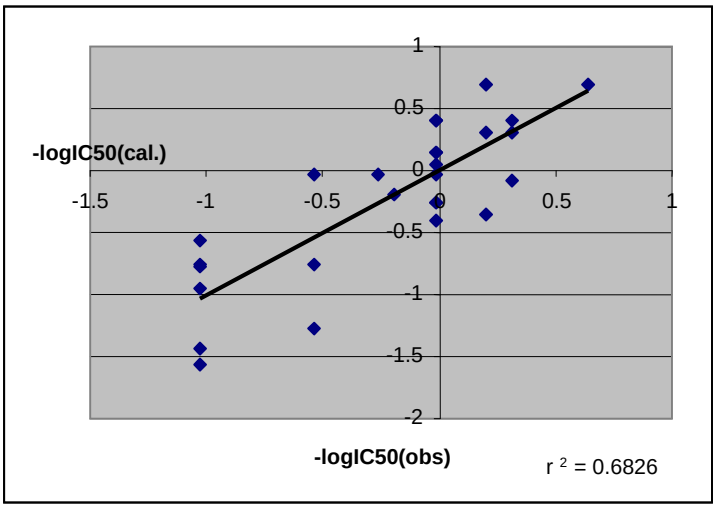


Fig:5.2

Conclusion

The above studies reveal the importance of field effect and molar refractivity parameters for the substituents at R_2 position. Hence the substituents having **lower value of field effect and molar refractivity (small and less bulky substituent groups/atoms)** will be preferred in order to increase the activity of given series of compounds or designing the new chemical entities for cancer

6. QSAR studies of Pyrazolo(3, 4-d) pyrimidines derivatives as potent inhibitors of the Insulin-like growth factor receptor(IGF-IR)

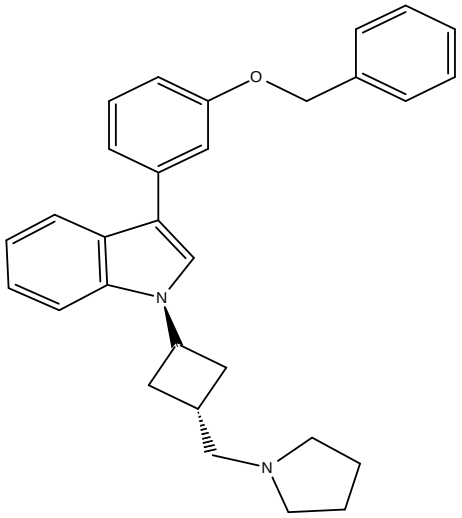
Introduction

Receptor tyrosine kinases (RTKs) are recent targets for cancer chemotherapy[184].Certain agents that inhibit RTKs are Bcr-Abl,EGFR and EGFR/Erb B-2.Search for additional RTKs led to the point that small molecule inhibitors of insulin like growth factor receptor(IGF-IR)may yield effective anticancer agent [185].Large number of approaches are used to target IGF-IR eg antibodies, antisense RNA, small molecule inhibitors[186] · Small molecule inhibitors are divided into two types:

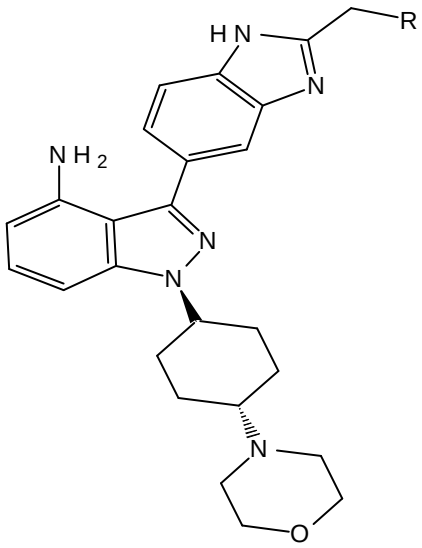
1) substrate inhibitors of IGF-IR[187] .

2) ATP competitive inhibitors of IGF-IR.[188]

One of the example of ATP competitive inhibitors of IGF-IR in both enzymatic and cellular essay is NVP-ADW742 **1**[189] .During its study, it was noted that Pyrazolopyrimidines class of kinase inhibitors possess potency for IGF IR. [190] QSAR studies have been done on Pyrazolo (3, 4-d) pyrimidines derivatives **2** in order to find important physico chemical parameters which affects its activity.



1

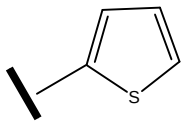


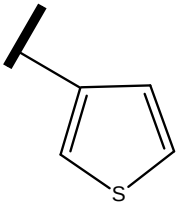
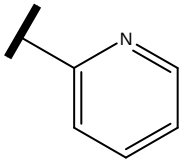
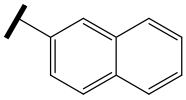
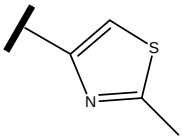
2

Methodology

In order to find the influence of different physicochemical parameters on the activity of Pyrazolo (3, 4-d) pyrimidines derivatives, QSAR studies were carried out using – log of inhibitory constant (IC_{50}) as dependent parameter[184]. The values of different parameters *viz.* hydrophobic (π), steric (MR), electronic (σ), field effect (f), resonance (RR) were taken from literature.[155, 156] and considered them as independent parameter. The multiparameter regression analysis was carried out using SYSTAT software. The parameters that showed significant correlation during analysis are given in table.6.1 along with the biological activities of all compounds of the series.

Table6.1: Physicochemical parameters and inhibitory activity of Pyrazolo(3, 4-d) pydimidines derivatives

C.No	R_1	j R	RR	-logIC$_{50\ Cal.}$				-logIC$_{50obs}$
				eq. 3	eq. 4	eq. 5	eq. 6	
1	-Ph	25.36	-0.08	-1.900	-1.547	-1.518	-1.499	-1.806
2	-2-Cl-Ph	30.36	-0.23	-2.249	-2.098	-2.129	-2.179	-1.568
3	-2-Me-Ph	29.98	-0.21	-2.217	-2.044	-2.070	-2.115	-1.580
4	-2-F-Ph	25.25	-0.42	-2.111	-1.992	-2.001	-1.992	-1.708
5	-2-OMe-Ph	32.20	-0.59	-2.571	-2.707	-------	---------	-1.908
6	-2-Br-Ph	23.21	-0.25	-1.899	-1.622	-1.594	-1.553	-1.851
7	-3-Cl-Ph	30.36	-0.23	-2.249	-2.098	-2.129	-2.179	-2.090
8	-3-Me-Ph	29.98	-0.21	-2.217	-2.044	-2.070	-2.115	-2.270
9	-3-F-Ph	25.25	-0.42	-2.111	-1.992	-2.001	-1.992	-1.763
10	-4-Cl-Ph	30.36	-0.23	-2.249	-2.098	-2.129	- 2.179	-2.808
11	-4-Me-Ph	29.98	-0.21	-2.217	-2.044	-2.070	-2.115	-1.845
12	-4-F-Ph	25.25	0.42	-1.576	--------	--------	---------	-2.966
13	-2,6-di-F-Ph	25.14	-0.76	-2.322	-2.438	-2.484	-2.485	-1.898
14	-2-F,6-Cl-Ph	30.25	-0.57	-2.460	-2.543	-2.612	-2.672	-1.954
15	-3,5-di-F-Ph	25.14	-0.76	-2.322	-2.438	-2.484	-2.485	-2.182
16	-2,3-di-F-Ph	25.14	-0.76	-2.322	-2.438	-2.484	-2.485	-2.238
17	-3,4-di-F-Ph	25.14	-0.76	-2.322	-2.438	-2.484	-2.485	-2.829
18	-2,6-di-Cl-Ph	35.36	-0.38	-2.597	-2.649	-2.740	---------	-2.097
19	-3,4-di-Cl-Ph	35.36	-0.38	-2.597	-2.649	-2.740	-2.859	-2.688
20	-2,3-di-Cl-Ph	35.36	-0.38	-2.597	-2.649	-2.740	-2.859	-3.366
21		24.04	0.04	-1.757	-1.294	-1.240	- 2.859	-1.362

22		24.04	-0.06	-1.820	-1.427	-1.385	-1.348	-1.505
23		26.59	-0.76	-2.395	-2.539	-2.598	-2.618	-3.369
24		38.79	-0.19	-2.650	-2.637	-2.735	-2.893	-3.117
25		26.16	-1.05	-2.558	-2.896	-2.983	-3.007	-3.520

Results and Discussion

Among different physicochemical parameters *viz*. π, MR f RR σ_m, σ_p , the parameters **π, f and σ_p**, were not considered for further study since they **showed poor correlation with activity** [π_1(r=0.123), f(r=0.064), σ_p(r=0.019)].The parameters **MR**(r=0.315),**RR**(r=0.292) and **±m**(r = 0.162) showed **better correlations with activity** than the above parameters, and are described in table 6.2.

Table –6.2: Pearson correlation matrix among different parameters

	MR	**RR**	**± m**	**JlogIC50**
MR	1			
RR	0 .071	1		
± m	0.009	0.327	1	
JlogIC50	0.315	0.292	0.162	1

First of all linear regression analysis of the above parameters were tried but could not obtain satisfactory results .Hence different combinations of the above independent parameters were tried keeping in view the inter correlation($r<0.5$) amongst parameters used in same equation.

$-\log IC_{50} = -0.047\ (\pm 0.030)\ MR - 0.541\ (\pm 0.652)\ \sigma_m - 0.844$ ---------- (eq.6.1)

n = 25 r = 0.355 F =1.590 s =0.627

$-\log IC_{50} = -0.245\ (\pm 0.704)\ \sigma_m + 0.535\ (\pm 0.431)\ RR - 2.023$ ---------- (eq. 6.2)

n = 25 r = 0.300 F =1.091 s =0.640

$-\log IC_{50} = -0.051\ (\pm 0.029 MR) + 0.637\ (\pm 0.383)\ RR - 0.565$ ---------- (eq. 6.3)

n = 25 r = 0.447 F =2.743 s =0.600

Where **'n'** is the no. of compounds used in the study, **r** is correlation coefficient, **F** is the variance ratio between the calculated and observed activities, and **s** is standard error of estimation.

None of the above three models were found to be satisfactory but **Eq. 6.3 was found to be better in terms of correlation coefficient value(r,) standard error(s), Fvalues**. It possess large correlation coefficient value (r = 0.355 , r = 0.300, r = 0.447 in eq . 6.1, 6.2, 6.3 respectively) ,low stantard error value (s =0.627, s =0.640, s =0.600 in eq. 6.1, 6.2, 6.3 respectively) and high F value (F=1.590 , F =1.091, F =2.743 in eq. 6.1, 6.2, 6.3 respectively) with respect to remaining two models So

Further studies were concentrated on model 6.3. On reanalysis of the model no. 6.3 compound no.12 was identified as an outlier having high residual value ((~3 times of the standard error) Its removal led to eq 6.4 having improved correlation coefficient and statistical significance(r= 0.447 , 0.682 in eq. 6.3, 6.4 respectively)

$-\log IC_{50} = -0.070\ (\pm 0.024)\ MR + 1.333(\pm 0.370)\ RR + 0.342$ ---------- (eq. 6. 4)

n = 24 r = 0.682 F =9.114 s =0.489

On analyzing equation 6.4 carefully; compound number 5 was identified as an outlier. Its removal from eq. 6.4 resulted in equation 6. 5 with improved correlation coefficient (r> 0.70),decreased standard error (s =0.489 ,0.464 in eq. 6.4, 6.5 respectively) and increased F value (F =9.114, F =11.508 in eq. 6. 4, 6.5 respectively)

$-\log IC_{50} = -0.079\ (\pm 0.023)\ MR + 1.446\ (\pm 0.357)\ RR + 0.597$---------- (eq6.5)

n = 23 r = 0.731 F =11.508 s =0.464

While reanalyzing model 5, compound no 18 was identified as an outlier. Its removal from eq 6.5 gave eq. 6.6 having high statistical significance (99.9%) and improved correlation coefficient(r>0.75)

JlogIC$_{50}$= J0.092 (ï 0.024) j R H=1.480 (ï 0.346)= RR H=0.944---------- (eq. 6.6)

n = 22 r = 0.766 F =13.504 s =0.448

Model 6.6 was found to be the best, statistically significant model **($F_{2,19\sim\ 0.001}$ = 10.16; $F_{2,19\ 0.001}$ = 13.504)**, having moderate correlation coefficient (r=0.766) The graph between calculated and observed activity of eq no. 6.6 is given in Fig: 6.1 (**r^2=0.587)**.As the model contains steric (MR) and resonance parameter(RR) which

suggests that the activity of Pyrazolo (3, 4-d) pydimidines analogs are influenced mainly by these two parameters. Since the coefficient of MR parameter contains negative sign and there is positive sign before the resonance parameter which suggests their negative and positive contributions to activity, respectively Thus the activity of given series of drug molecule can be enhanced by substitutiting groups having less MR value(small and less bulky groups) and large resonance parameter value ,at position R

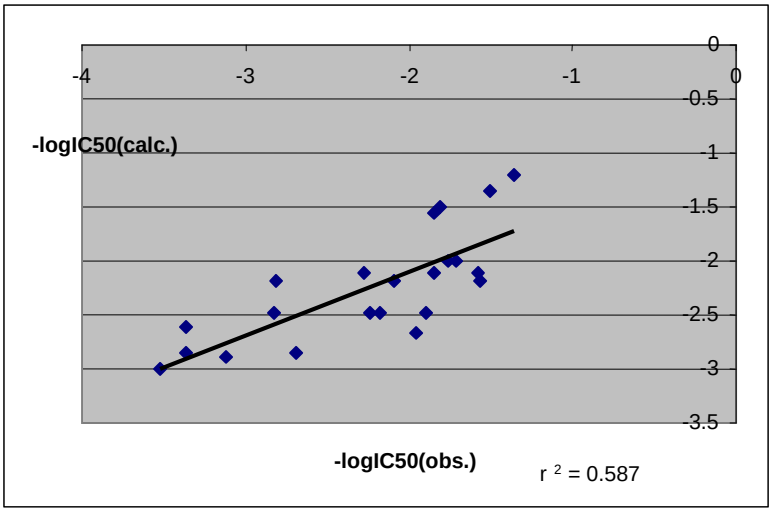


Fig: 6.6

Conclusion

The above QSAR studies on Pyrazolo (3, 4-d) pydimidines analogues highlight the important parameter influencing its activity. The studies led us to know that MR and RR are important parameters. It can be concluded that substituents with low MR value and high resonance parameter values at position R will increase the activity of given series of compounds and will be helpful in designing the new chemical entities for cancer,

7. QSAR studies of Pyrrolidiones derivatives as potent functional antagonists of human melanocortin-4 receptor

Introduction

Melanocortin-4 receptor (MC4R) plays a vital role in regulating feeding behaviour and other biological function[191].It is a member of G –Protien-Coupled receptor(GPCR)[192]A number of its antagonists are found to reverse mass loss and enhance intake of food in animal model of cachexia which suggests that it can be used for treatment of cancer cachexia.[193,194] In order to find important essential structural and physicochemical sparameters, that effect the inhibitory activity of Pyrrolidiones derivatives(Fig:7.1),antagonists of MC4R,QSAR study was performed using linear free relationship (LFER) approach.The QSAR[155]studies have been carried out considering anticancer activity as dependent and different physicochemical parameters like hydrophobic (π), steric (MR) and electronic (f, R, σ) as independent parameters

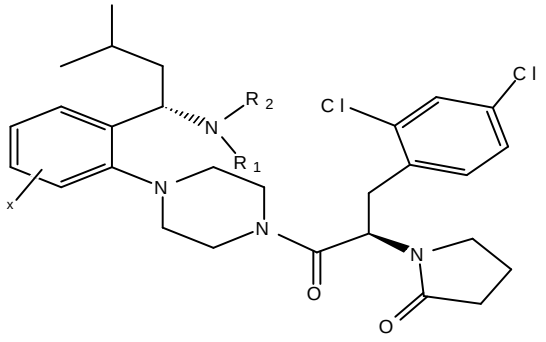


Fig:7.1

Methodology

The MC4R inhibitory activity data (k_i) of Pyrrolidiones derivatives have been taken from literature[192] and were converted to –log k_i for use in QSAR analysis as dependent parameter (table7.1). The physicochemical parameters viz. hydrophobic (π_1 and π_x), steric (MR_1 and MR_x), electronic (σ_1, andσ_x), resonance (RR_1 and RR_x) and field effect ($f_1$and f_x) for the substituents at R_1 and R_x positions respectively were taken from literature[156] An indicator variable (IR_1) having values "1"or' 0', is also introduced to indicate the presence or absence of **H** atom, respectively at position R_1 These parameters are used as independent parameters in deriving QSAR equations. The parameters which showed some significant correlation with activity are described in table 7.1.

Table 7.1: Physicochemical parameters and inhibitory activity of Pyrrolidiones derivatives

C.No	X	R_1	R_2	f_2	f_x	-logKi cal		-logKi obs
						eq. 5	eq. 6	
1	H	H	H	0.00	0.00	-0.490	-0.507	-0.806
2	6-F	H	H	0.00	0.43	-1.263	-1.217	-0.987
3	4-F	H	H	0.00	0.43	-1.263	-1.217	-1.415
4	4-Me	H	H	0.00	-0.04	-0.418	-0.441	-0.505
5	4-CF_3	H	H	0.00	0.38	-1.173	-1.131	-0.653
6	6-F	H	Et	-0.05	0.43	-1.115	-1.041	-0.792
7	6-F	Et	Et	-0.05	0.43	-1.115	-1.041	-0.898
8	6-F	H	CH_2CH_2OH	0.24	0.43	-0.749	-0.617	-0.806
9	4-CF_3	H	CH_2CH_2OH	0.24	0.38	-0.659	-0.535	-0.939
10	6-F	H	$CH_2CH_2NH_2$	-0.03	0.43	-0.463	-0.284	0.027
11	4-F	H	$CH_2CH_2NH_2$	-0.03	0.43	-0.463	-------.	-1.778
12	4-CF_3	H	$CH_2CH_2NH_2$	-0.03	0.38	-0.373	-0.202	0.000
13	6-F	Me	$CH_2CH_2NH_2$	-0.03	0.43	-0.463	-0.284	-0.176
14	H	H	CH_2CH_2NHMe	-0.16	0.00	-1.172	-1.298	-0.579
15	4-F	H	CH_2CH_2NHMe	-0.16	0.43	-1.945	-2.004	-1.690
16	4-F	H	$CH_2CH_2NMe_2$	0.05	0.43	--------.	--------.	-2.079
17	4-F	H	$CH_2CH_2NEt_2$	-0.04	0.43	--------.	--------.	-1.748
18	H	H	$CH_2CH_2CH_2NH_2$	-0.04	0.00	0.290	0.399	0.022
19	6-F	H	$CH_2CH_2CH_2NH_2$	-0.04	0.43	-0.482	-0.307	0.046
20	4-CF_3	H	$CH_2CH_2CH_2NH_2$	-0.04	0.38	-0.393	-0.225	-0.146
21	4-F	H	CH_2CONH_2	0.20	0.43	-1.451	-1.431	-2.146
22	4-CF_3	H	CH_2COOH	0.29	0.38	-1.450	-1.452	-1.431
23	4-F	H	$CH_2CH_2CH_2COOH$	0.28	0.43	-1.559	-1.557	-2.380
24	4-CF_3	H	$CH_2CH_2CH_2COOH$	0.28	0.38	-1.469	-1.475	-1.602

During the study, among several physicochemical parameters *viz*. π_1,π_x, MR_1,MR_x,σ_1 ,σ_x, RR_1, RR_x, f_1, f_x and IR_1, the parameters $\mathbf{f_x}$ (r=0.337), ®$_2$ (r=0.331),$\mathbf{f_2}$ (r = 0.440), and ±$_2$ (r=0.335) **showed better correlation with activity** than the remaining parameters and were thus considered for further study ,also described in table 7.2.

Table –7.2: Pearson correlation matrix among different parameters

	$\mathbf{f_x}$	®$_2$	$\mathbf{f_2}$	±$_2$	**JlogKi**
$\mathbf{f_x}$	1				
®$_2$	0.124	1			
$\mathbf{f_2}$	0.242	0.029	1		
±$_2$	0.219	0. 051	0.151	1	
JlogKi	0.337	0.331	0.440	0.335	1

Different combinations of the independent parameters were used for correlating the activity in QSAR analysis keeping in view the inter correlation($r<0.5$) amongst parameters used in same equation.

$-\log Ki = -2.167\ (\pm 1.082) f_2 - 0.414(\pm 0.263)\sigma_2 - 1.011$ ---------- (.eq7.1)

n =24 r =0.528 F=4.059 s=0.668

$-\log Ki = -0.360\ (\pm 0.207)\ \pi_2 - 2.361(\pm 1.005) f_2 - 0.756$ ---------- (.eq7.2)

n =24 r =0.543 F=4.395 s=0.661

$-\log Ki = -1.127\ (\pm 0.897)\ f_x - 2.088(\pm 1.068)\sigma_2 - 0.513$ ---------- (.eq7.3)

n =24 r =0.500 F=3.498 s=0.682

$-\log Ki = -2.004\ (\pm 0.857)\ f_x - 0.632\ (\pm 0.261)\ f_2 - 0.513$ ---------- (.eq7.4)

n =24 r =0.554 F=4.644 s=0.655

Where **'n'** is the no. of compounds used in the study, **r** is correlation coefficient, **F** is the variance ratio between the calculated and observed activities, and **s** is standard error of estimation.

Among the different permutations and combinations tried,eq. 7.4 was found to be better, since it possess better correlation coefficient (r =0.528 ,r =0.543 r =0.500 , r =0.554 in eq 7.1, 7.2, 7.3,7.4 respectively) ,less standard error value(s=0.668, s=0.661 s=0.682 ,s=0.655 in eq. 7.1, 7.2, 7.3, 7.4 respectively) and high F value(F=4.059, F=4.395, F=3.498, F=4.644 in eq. 7.1, 7.2, 7.3, 7.4 respectively) than the remaining one. So further studies were targeted on these two parameters and while reanalysis of the model no. 7.4 , compound no.16 and 17 were identified as outliers having high residual value ((~3 times of the standard error). Their removal led to eq. 7.5 having improved correlation coefficient and statistical significance(r= 0.554, 0.746 in eq. 7. 4, 7.5 respectively)

$$-\log Ki = -1.797\ (\pm 0.667)\ f_x - 0.988\ (\pm 0.222)\ f_2 - 0.490 \quad \text{---------- (.eq7.5)}$$

n =22 *r* =0.746 F=11.894 s=0.508

On reanalysis of model 7. 5 , compound no. 11 was identified as an outlier Its removal from eq 7.5 gave eq. 7.6 having high statistical significance (99.9%) and good correlation coefficient(r>0.80)

JlogKi = JNKSQP=EïMKRPOF $\mathbf{f_x}$=J=NKNQT=EïMKNUOF=$\mathbf{f_2}$-J=MKRMT====---------- **(.eq7.6)**

n =21 *r* =0.844 F=22.307 s=0.404

Thus **eq. 7.6 was found to be the best** both in terms of correlation value(r>0.82) and statistical significance **($F_{2,18\sim\ 0.001}$ = 10.39; $F_{2,18\ 0.001}$ = 22.307)**.The graph between calculated and observed activity of eq no. 7.6 is given in Fig: 7.2 (**r^2=0.712**)which proves the validity of the model As the eq7.6 consist of field effect parameter(f_x and f_2) which suggests that the activity of Pyrrolidiones derivatives are influenced mainly by these two parameters. Since coefficient of both parameters contains negative sign which indicates their negative contribution to activity Thus the activity of given series of drug molecule can be enhanced by substituting groups at position R_2 and x, having less field effect value.

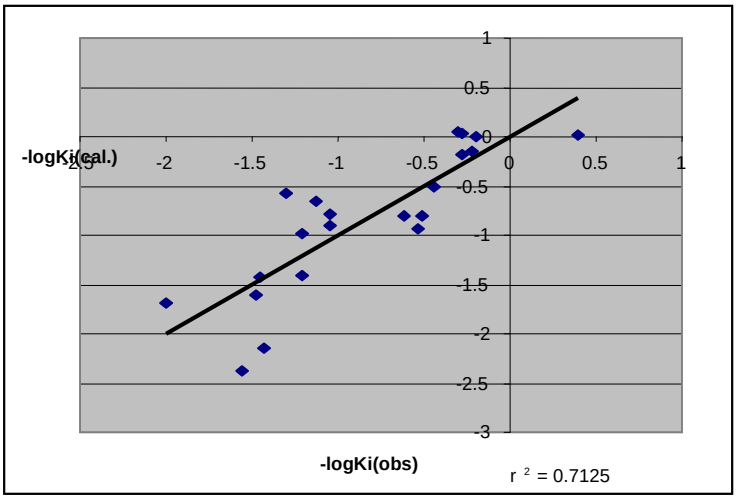


Fig 7.2

Conclusion

On the basis of above QSAR studies on Pyrrolidiones derivatives it can be concluded that **field effect is an important parameter influencing its activity** The studies led us to know that the activity of given series of compounds can be increased by substituting **substituents with low field effect value at position R_2 and X** The result is also helpful in designing new chemical entities for cancer

8. Qsar studies of 2-(alkyloxyaryl)-1*H*-benzimidazole derivatives as antitumor agents

Introduction

Benzimidazole moiety shows several pharmacological activities as antitumorvasorelaxant, antifungal, antiparasitic etc[195-197]. It is an important structural moiety The condensation of 1,2 phenylenediamines and aromatic substituted aldehyde in presence of sodium metabisulfite generate bezamidazole derivatives[198]Here QSAR studies have been performed on 18 derivatives of 2-(alkyloxyaryl)-1*H*-benzimidazole derivatives(**Fig:8.1**), to find important physico chemical parameter which influences their activity.

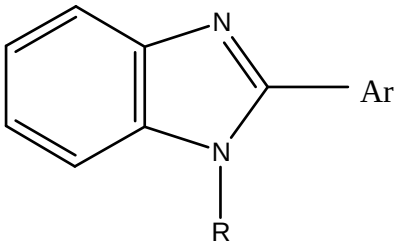


Fig: 8.1

Methodology

All the inhibitory activity data(IC_{50}) of 2-(alkyloxyaryl)-1*H*-benzimidazole derivatives have been collected from literature[199],converted to $-\log IC_{50}$ and used as dependent parameter during the study .The physicochemical parameters viz. hydrophobic (π_{Ar}), steric (MR_{Ar}), electronic ($\sigma_{M\ Ar}$, $\sigma_{P\ Ar}$), resonance (RR_{Ar}) and field effect (f_{Ar}) for the substituents at Ar positions were taken from literature[156]. An indicator variable (IR) having values “1”or’ 0’, is also introduced to indicate the presence or absence of **H** atom, respectively at position R These parameters are used

as independent parameters in deriving QSAR equations[155]. The parameters which showed some significant correlation with activity are described in table 8.1.

Table 8.1: Physicochemical parameters and inhibitory activity of 2-(alkyloxyaryl)-1*H*-benzimidazole derivatives

C.No.	R	Ar	RR	IR	$\log IC_{50cal}$		$\log IC_{50obs}$
					eq.4	eq. 5	
1	H	2-HydroxyPhenyl	-0.720	1	-1.051	-0.990	-0.712
2	H	2-MethoxyPhenyl	-0.590	1	-1.176	-1.119	-0.886
3	H	2-EthoxyPhenyl	-0.520	1	.------	------.	-0.288
4	H	2-PropoxyPhenyl	-0.530	1	.------	------.	-0.279
5	H	2-(Benzyloxy)Phenyl	-0.730	1	-1.041	-0.980	-0.450
6	H	2-(4-Cl Benzyloxy)Phenyl	-0.880	1	-0.897	------.	-1.771
7	H	2-(4-Me-Benzyloxy)Phenyl	-0.860	1	-0.916	-0.852	-1.176
8	H	2-NitroPhenyl	0.080	1	-1.822	-1.782	-1.895
9	H	4-HydroxyPhenyl	-0.720	1	-1.051	-0.990	-0.720
10	H	4-MethoxyPhenyl	-0.590	1	-1.176	-1.119	-1.025
11	H	2-(*N*,*N* Dimethyl amino)Phenyl	-0.640	1	-1.128	-1.069	-1.477
12	H	4-Hydroxy,3-Methoxy Phenyl	-1.230	1	-0.559	-0.486	-0.076
13	H	3,4-(DiMethoxy) Phenyl	-1.100	1	-0.685	-0.614	-0.748
14	H	2,3,4-(TriMethoxy) Phenyl	-1.610	1	-0.193	-0.110	-0.255
15	H	2,4,5-(TriMethoxy) Phenyl	-1.610	1	-0.193	-0.110	-0.602
16	H	3,4-(Methylenedioxy)Phenyl	-0.640	1	-1.128	-1.069	-0.531
17	H	4-Pyridyl	-0.160	1	-1.590	-1.544	-2.279
18	CH_2CH_3	2-EthoxyPhenyl	-0.520	0	-0.236	-0.236	- 0.236

Results and Discussion

Among different physicochemical parameters *viz*. π_{Ar}, MR_{Ar} f_{Ar} RR_{Ar}, $\sigma_{M\ Ar,}$ $\sigma_{P\ Ar}$, and IR, the parameters π_{Ar} (r=0.160), MR_{Ar} (r = 0.009), $\sigma_{P\ Ar}$ (r = 0.187), showed the least scorrelation with activity and were thus not considered for further study. The parameters $\mathbf{RR_{Ar}}$, (r =0.506), $\pm_{j\ =Ar}$ (r = 0.377) $\mathbf{f_{Ar}}$ (r = 0.223) and **IR**(r = 0.244) **showed better correlations** with activity than the above parameters, and are described in table8. 2.

Table –8.2: Pearson correlation matrix among differenr parameters

	$\mathbf{RR_{Ar}}$,	$\pm_{j\ =Ar}$	$\mathbf{f_{Ar}}$	**IR**	**J**$\mathbf{logIC_{50}}$
$\mathbf{RR_{Ar}}$,	1				
$\pm_{j\ =Ar}$	0.812	1			
$\mathbf{f_{Ar}}$	0.580	0.672	1		
IR	0.136	0.130	0.111	1	
J$\mathbf{logIC_{50}}$	0.506	0.377	0.223	0.244	1

These Parameters were used to correlate the activity both independently and in combination of 2 or 3 parameters keeping in view the inter correlation(r<0.5) amongst parameters used in same equation. No satisfactory results were obtained by linear regression analysis. The different equations obtained by multi regression analysis are:

$-\log IC_{50}$= 0.612(±0.589) f_{Ar} – 0 .733 (± 0.656) IR – 0.419---------- (eq.8.1)

n = 18 *r* = 0.350 F =1.048 s =0.633

$-\log IC_{50}$= 5.145(±2.827) $\sigma_{M\ Ar}$ – 0 .802 (± 0.616)IR – 0.236---------- (eq. 8.2)

n = 18 *r* = 0.479 F =2.234 s =0.593

$-\log IC_{50} = -0.813\ (\pm 0.310)\ RR_{Ar} - 0.858\ (\pm 0.564)\ IR - 0.658$ ---------- (eq. 8.3)

n = 18 r = 0.596 F =4.136 s =0.543

Where **'n'** is the no. of compounds used in the study, **r** is correlation coefficient, **F** is the variance ratio between the calculated and observed activities, and **s** is standard error of estimation.

Among these equations,**eq. 8.3 was found to be better**, since it possess better correlation coefficient (r =0.350,r =0.479,r =0.596 in eq. 8.1, 8.2, 8.3 respectively) , less standard error value(s=0.633, s=0.593 s=0.543 in eq. 8.1, 8.2, 8.3 respectively) and high F value (F=1.048 F=2.234, F=4.136 in eq. 8.1, 8.2, 8.3 respectively) than the remaining two equations. On reanalyzing equation 8.3, compound no.3 and 4 were identified as outliers having high residual value (~3 times of the standard error).Their removal led to eq. 8.4 having improved correlation coefficient and statistical significance(r= 0.596, 0.728 in eq. 8.3, 8.4 respectively)

$-\log IC_{50} = -0.964\ (\pm 0.274)\ RR_{Ar} - 1.008\ (\pm 0.492)\ IR - 0.737$ ---------- (eq. 8.4)

n = 16 r = 0.728 F =7.317 s =0.471

On analyzing equation 8.4 carefully; compound number 6 was identified as an outlier. Its removal from eq. 8.4, led to equation 8.5 with improved correlation coefficient (r>0.78).

$J\log IC_{50} = J0.989\ (ï 0.242) RR_{Ar} - 0.953\ (ï\ 0.434)\ IR - 0.750$ ---------- (eq. 8.5)

n = 15 r = 0.785 F =9.631 s =0.414

The equation 8.5 thus obtained above is statistically significant as it has moderate correlation coefficient(r=0.785)of high(>99.95%) statistical significance **($F_{2,18\sim=0.001}$ = 3.89 ; $F_{2,16\ 0.001}$ = 9.631)** The graph between calculated and observed activity of eq no. 8.5. is given in Fig 8.2 **(r^2= 0.616)** Since in eq. 8.5 the coefficient of parameters **$RR_{=Ar}$** and **IR** possess negative sign which indicates their negative contribution to activity. The negative contribution by RR_{Ar} and IR suggests that the substitutions at Ar position by the molecules with less resonance parameter value and atom other than H at position R should be preferred for enhancing the activity of given series on drug molecule

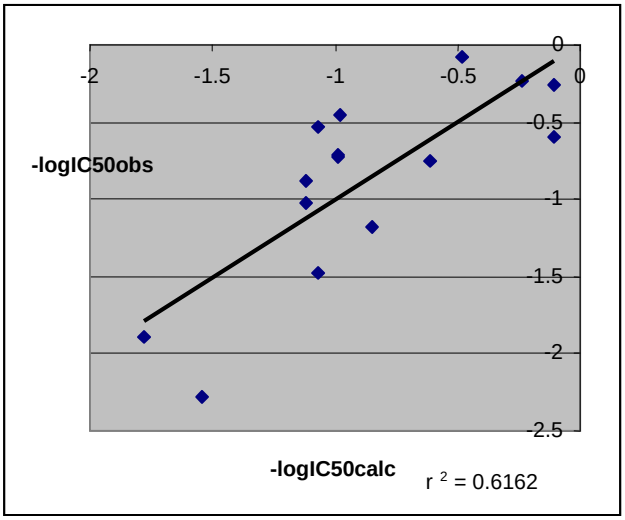


Fig :8.2

Conclusion

These studies led us to know that RR_{Ar} and IR are important parameters which influence the activity of given series of 2-(alkyloxyaryl)-1*H*-benzimidazole derivatives Thus for increasing the activity of given series of compounds or designing the new chemical entities for cancer, the **substituent with low resonance effect parameter values will be preferred at position Ar and H atom at position R is not preferred** since it contributes negatively to activity as shown by equation. **8.5**

9. QSAR studies on 3, 17-Disubstituted 2-Alkylestra-1, 3, 5(10) -trien-3-ol Derivatives as Anticancer Agents

Introduction

Steroid Sulfatase (STS) is a clinical target for treatment of breast cancer.[200,201] Sulfamoylated analogues of 2-methyl estradiol such as 2-methyl estradiol-3, 17- *O, O*-bis sulfamate are found to be inhibitors of STS and cancer cell growth[202-206].Phenolic Sulfamate inhibits STS either by transferring the sulfamate group to residue in enzyme's active site or nucleophilic attack on formylglysine residue of active site by sulfamate group.2-Methyl estradiol inhibits both angiogenesis and cancer cell proliferation. So its various analogues have been synthesized and tested[207]. In order to find important structural and physicochemical parameters which influence its activity,QSAR studies have been performed on 23 derivatives of 3, 17-Disubstituted 2-Alkylestra-1, 3, 5(10)-trien-3-ol**(Fig:9.1),** considering antimitotic activity as dependent and different physicochemical parameters like hydrophobic (π), steric (MR) ,electronic (f, R, σ) and indicators as independent parameters in the analysis[155].

.

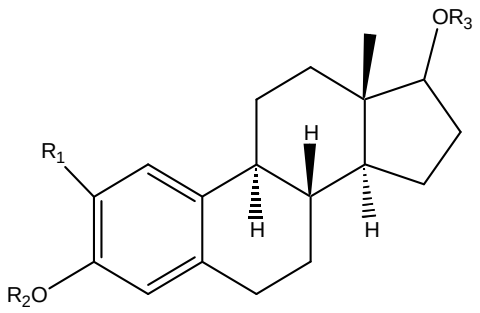


Fig: 9.1

Methodology

The QSAR studies were carried out using LFER approach, in order to find the influence of physicochemical parameters *viz.* hydrophobicity (π); electronic effect (σ); resonance effect (R); field effect (f) and steric effect (MR) on inhibitory activity of prototype 1. The values of physicochemical parameters viz. hydrophobic (π_1 and π_3), steric ($MR_1$and MR_3), electronic ($\sigma_1$andσ_3), resonance (RR_1 and RR_3) and field effect (f_1 and f_3) for the substituents at $R_1$and R_3 positions respectively were taken from literature[156] and were used as independent parameter. An indicator parameter was introduced with its value**"1"or' 0',** to indicate the presence or absence of **SO_2NH_2** group respectively, at position R_2. For QSAR analysis inhibitory activity data (GI_{50}) of 3,17-Disubstituted 2-Alkylestra-1, 3,5(10)-trien-3-ol have been taken from the literature[207,] converted to -log GI_{50} and were used as dependent parameter. The multiparameter regression analysis was executed using SYSTAT software. The parameters that showed significant correlation during analysis are given in table.9. 1 along with the biological activities of all compounds of the series.

Results and discussion

The QSAR analysis using several physicochemical parameters, for the effect of substituents R_1, R_2 and R_3 on activity, led to the identification of important parameters viz. π_1 , MR_1,π_3, MR_3 and f_3 with correlation coefficient $r > 0.40$. Pearson correlation matrix (Table 9.2) was constructed to determine the intercorrelation between these physiochemical parameters.

Table 9.1: Physicochemical parameters and antimitotic activity of 3,17-Disubstituted 2-Alkylestra-1,3,5(10)-trien-3-ol Derivative

C.no	R_1	R_2	R_3	$®_1$	$®_3$	-log$GI_{50\ Cal.}$				-log$GI_{50i\ obs}$
						eq. 3	eq. 4	eq. 5	eq. 6	
1	MeO	H	H	-0.02	0.00	-0.520	-0.363	-0.448	-0.237	-0.086
2	Et	H	H	1.02	0.00	-0.975	-0.979	-1.040	-0.906	-1.013
3	MeO	SO_2NH_2	SO_2NH_2	-0.02	-3.42	0.329	0.712	0.628	0.687	0.469
4	Et	SO_2NH_2	SO_2NH_2	1.02	-3.42	-0.126	0.096	0.036	0.0179	0.678
5	Me	H	H	0.56	0.00	-0.774	-0.706	-0.778	.------	-1.708
6	t-Bu	H	H	1.98	0.00	-1.395	-1.548	-1.586	-1.524	-1.380
7	i-Pr	H	H	1.53	0.00	-1.198	-1.281	-1.330	-1.234	-1.362
8	n-Pr	H	H	1.55	0.00	-1.207	-1.293	-1.341	-1.247	-1.176
9	H	SO_2NH_2	SO_2NH_2	0.00	-3.42	0.321	.------	.--------	--------.	-1.061
10	Me	SO_2NH_2	SO_2NH_2	0.56	-3.42	0.076	0.368	0.298	0.314	0.420
11	n-Pr	SO_2NH_2	SO_2NH_2	1.55	-3.42	-0.357	-0.218	-0.266	-0.323	-0.531
12	i-Pr	SO_2NH_2	SO_2NH_2	1.53	-3.42	-0.349	-0.206	-0.254	-0.310	-0.255
13	n-Bu	SO_2NH_2	SO_2NH_2	2.13	-3.42	-0.611	-0.562	-0.596	-0.697	-0.954
14	Et	SO_2NH_2	Ac	1.02	-0.55	-0.838	-0.806	-0.867	-0.758	-0.204
15	n-Pr	SO_2NH_2	Ac	1.55	-0.55	-1.070	-1.120	-1.168	-1.099	-1.787
16	n-Bu	SO_2NH_2	Ac	2.13	-0.55	-1.324	-1.464	-1.498	-1.472	-1.265
17	MeO	SO_2NH_2	$CONH_2$	-0.02	-1.49	-0.150	0.106	0.021	0.166	-0.127
18	Me	SO_2NH_2	$CONH_2$	0.56	-1.49	-0.404	-0.238	-0.309	-0.207	-0.279
19	Et	SO_2NH_2	$CONH_2$	1.02	-1.49	-0.605	-0.511	.-------	-------.	0.658
20	n-Pr	SO_2NH_2	$CONH_2$	1.55	-1.49	-0.837	-0.825	-0.873	-0.845	-1.322
21	i-Pr	SO_2NH_2	$CONH_2$	1.53	-1.49	-0.828	-0.813	-0.861	-0.832	-0.380
22	n-Bu	SO_2NH_2	$CONH_2$	2.13	-1.49	-1.090	-1.168	-1.203	-1.218	-1.176
23	t-Bu	SO_2NH_2	$CONH_2$	1.98	-1.49	-1.025	-1.079	-1.117	-1.122	-1.114

Table –9.2: Pearson correlation matrix among different parameters

	®N =	MR 1	®P	j RP	f3	JlogGI50
®N =	1					
MR 1=	0.937	1				
®P	0.145	0.180	1			
j RP=	0.001	0.039	0.829	1		
f3=	0.083	0.130	0.944	0.920	1	
JlogGI50	0.493	0.427	0.517	0.425	0.457	1

Since no satisfactory results were obtained by linear regressions analysis, multiregression analysis was performed. Thus several permutations and combinations of the above parameters were tried keeping in view the orthogonality ($r<0.5$) among the parameters used in same equation in order to develop meaningful QSAR equations.

$-\log GI_{50} = -0.507\ (\pm 0.174)\pi_1 + 0.066(\pm 0.026)MR_3 - 0.634$ ---------- (eq9.1)

n =23 r = 0.652 F =7.382 s =0.599

$-\log GI_{50} = -0.470\ (\pm 0.176)\pi_1 + 1.353(\pm 0.553)f_3 - 0.505$ ---------- (eq9.2)

n =23 r = 0.646 F =7.160 s =0.603

$-\log GI_{50} = -0.437\ (\pm 0.173)\pi_1 - 0.248(\pm 0.092)\pi_3 - 0.529$ ---------- (eq9.3)

n =23 r = 0.668 F =8.041 s =0.588

Where **'n'** is the no. of compounds used in the study, **r** is correlation coefficient, **F** is the variance ratio between the calculated and observed activities, and **s** is standard error of estimation.

On analyzing the above equations, equation 9.3 was found to be statistically significant (>99%) with moderate correlation coefficient (r> 0.66)Rest of the above models were not found to be satisfactory since they possess low correlation coefficient, low F value, high standard error value.

On reanalysis of model no. 9.3 , compound no.9 was identified as an outlier having highest residual value (~3 times of the standard error).Its removal led to eq. 9.4 with improved correlation coeficient(r>0.70) and statistical significance.

$-\log GI_{50} = -0.592\ (\pm 0.151)\pi_1 - 0.314\ (\pm 0.079)\ \pi_3 - 0.374$ ---------- (eq. 9.4)

n =22 r = 0.797 F =16.556 s =0.486

On careful analysis of equation 9.4, compound number 19 was identified as an outlier. Thus the removal of compound number 19 from equation 9.4, produced equation 9.5, with improved correlation coefficient (r> 0.80).

$-\log GI_{50} = -0.569\ (\pm 0.128)\pi_1 - 0.315(\pm 0.067)\pi_3 - 0.460$ ---------- (eq. 9.5)

n =21 r = 0.843 F =22.160 s =0.412

While reanalyzing this model, compound no. 5 was identified as an outlier. Its removal from equation 9.5 resulted in equation 9.6 with significant improved correlation coefficient (r> 0.85) and high statistical significance.

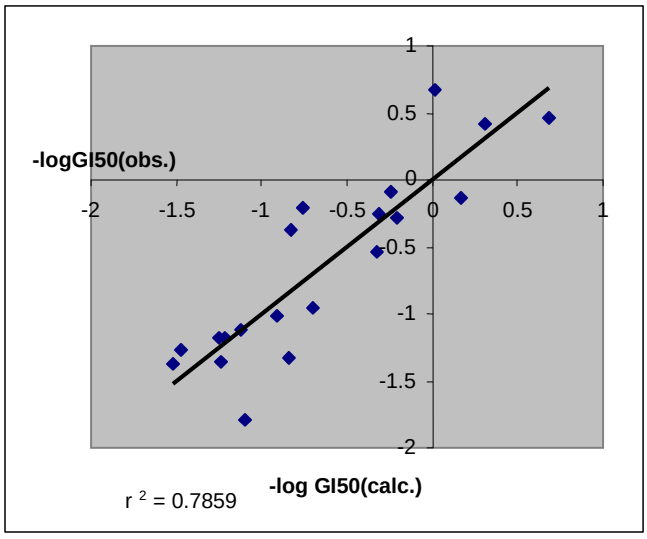
$\log GI_{50} = -0.644(\pm 0.110)\pi_N + 0.270$ (ïMKMRUF®3JMKOQV:---------- (eq. 9.6)

n =20 r = 0.886 F =31.183 s=0.346

The equation **9.6** thus obtained above was considered to be the best model explaining 78.6% variance in activity, having correlation coefficient (r = 0.88) The low standard error of estimate (s) and a high F value suggests that the model is statistically highly significant. The data showed overall statistical significance >99.99% with F=31.183 against tabulated value for Fischer's test at 99.9% significance **[$F_{2,17;\sim=0.001}$(obs) =10.66].** The eq. 9.6 led to conclusion that hydrophobic parameter influences the activity of given series of compounds, as the anticancer activity was best correlated with the hydrophobic parameter. The correlation between observed and predicted activities of all compounds using eq. 9.6 (best eq.) has been represented graphically as shown in Fig 9.2 which proves the validity of the model

Since the parameters present in eq. 9.6 (π_1 and π_3) possess negative sign before their coefficients, which indicates that these parameters contribute negatively to the activity. The eq. suggest that the activity of given series of molecule can be improved by substituting less hydrophobic group at R_1 and R_3 position of prototype 1.

Fig 9.2: Observed Vs calculated activity for 20 compounds



Conclusion

The above 2D QSAR analysis on 23 derivatives of 3, 17-Disubstituted 2-Alkylestra-1, 3, 5(10)-trien-3-ol, led to the identification of important physiochemical parameters in explaining the variation in activity in given series of molecules Thus for increasing the activity of given series of compounds or designing the new chemical entities for cancer, **the substituent with low hydrophobicity value will be preferred at position R_1 and R_3.**Hence the model can be useful in the optimization of activity in the given series of molecules and will also be helpful in designing new drug molecules.