EMPRO

Current Research in Pharmaceutical Sciences

Available online at www.crpsonline.com

crps online CRPS CURRENT RESEARCH IN PHARMACEUTICAL SCIENCES

ISSN 2250 – 2688

Received: 02/06/2012
Revised: 15/06/2012
Accepted: 22/06/2012

Khushboo Mishra, Neha Muzumdar, A.S. Parihar, R. S. Pawar, P. K. Singour
Computational & Synthetic Chemistry Division, Department of Pharmaceutical Chemistry, VNS Institute of Pharmacy, Neelbud, Bhopal, (M.P.) India

Correspondence
Khushboo Mishra
Computational & Synthetic Chemistry Division, Department of Pharmaceutical Chemistry, VNS Institute of Pharmacy, Neelbud, Bhopal, (M.P.) India
E-mail: khushboo09vns@gmail.com

QSAR analysis of 2-Acylamino-4, 6-Diphenylpyridine Derivatives as Novel antagonists of GPR-54

Khushboo Mishra, Neha Muzumdar, A.S. Parihar, R. S. Pawar, P. K. Singour

ABSTRACT

GPR54 is a G protein-coupled receptor (GPCR) which was formerly an orphan receptor. Recent functional study of GPR54 revealed that the receptor plays an essential role to modulate sex-hormones including GnRH. Antagonists of GPR54 (2-acylamino-4, 6-diphenylpyridine derivatives) are expected to be novel drugs for sex-hormone dependent diseases such as prostate cancer or endometriosis. The QSAR study on 2-acylamino-4, 6-diphenylpyridine derivatives as novel antagonists of GPR54 was carried out with 30 (21 training + 9 test) compounds. Molecular modeling was performed using ChemOffice 2006; supplied by Cambridge software, USA. The structures drawn were subjected to energy minimization by MM_2 and MOPAC and the lowest energy structure was used to calculate the physiochemical properties. The regression analysis was carried out using an automatic computer program called VALSTAT. The best models were selected from the various statistically significant equations. The analysis resulted in QSAR equation shows that biological activity is positively correlated to LogP (hydrophobicity constant), Partition Coefficient and Esb (stretch-bend energy). The statistical results of best model are, n=17, r=0.968213, r^2=0.937436, r^2adj=0.922999, q^2= 0.849, variance= 0.0678879, Std=0.260553 this study can help in rational drug design and synthesis of new novel GPR54 antagonists with predetermined affinity.

Keywords: QSAR, GPR54, 2-acylamino-4, 6-diphenylPyridine, Hypogonadotropic Hypogonadism

1. INTRODUCTION

Cancer is a group of diseases in which cells are aggressive (grow and divide without respect to normal limits), invasive (invade and destroy adjacent tissues) and sometimes metastatic (spread to other locations in the body). These three malignant properties of cancers differentiate them from benign tumors, which are self-limited in their growth and don't invade or metastasize (although some benign tumor types are capable of becoming malignant) [1, 2, 3].

Cancer may affect people at all ages, even fetuses, but risk for the more common varieties tends to increase with age. Cancer causes about 13% of all deaths. Nearly all cancers are caused by abnormalities in the genetic material of the transformed cells. These abnormalities may be due to the effects of carcinogens, such as tobacco smoke, radiation, chemicals, or infectious agents. Other cancer-promoting genetic abnormalities may be randomly acquired through errors in DNA replication or are inherited, and thus present in all cells from birth [4, 5].

GPR54 (OT7T175, AXOR12) is a G protein-coupled receptor (GPCR) that is highly expressed in brain, including hypothalamus and pituitary as well as peripheral regions. GPR54 was identified as an orphan receptor in rat in 1999 and Ohtaki et al. discovered that a 54-amino-acid product of a gene called Kiss-1 was its endogenous ligand. As the Kiss-1 gene was originally isolated as a tumor metastasis gene, the peptide product was named ‘metastin’. Later others also isolated the same peptide and named it ‘kisspeptin’. After much study to understand the function

of GPR54, several findings were reported that suggest GPR54 plays an important role in reproduction and pubertal development. The phenotype of GPR54-null mice was consistent with lack of steroid sex hormone production. In the mutant male mice, the serum testosterone level was similar to that found in normal females. In the case of the mutant females, they did not show the rise in estradiol normally found during estrus. Moreover, mutations of GPR54 were found in patients with idiopathic hypogonadoropic hypogonadism (IHH). Individuals with IHH fail to undergo puberty and are infertile because of failure to secrete the gonadotropic hormones, such as follicle stimulating hormone (FSH) and luteinizing hormone (LH) from the pituitary[6]. Hypogonadotropic hypogonadism is absent or decreased function of the male testes or the female ovaries. It is considered a form of secondary hypogonadism, which means the condition is due to a problem with the pituitary or hypothalamus gland. Hypogonadotropic hypogonadism is caused by a lack of the gonadal stimulating pituitary hormones: follicle stimulating hormone (FSH) and luteinizing hormone (LH). Normally, the hypothalamus in the brain releases gonadotropin-releasing hormone (GnRH), which stimulates the pituitary gland to release other hormones, including FSH and LH. These hormones tell the female ovaries and male testes to secrete hormones that are responsible for normal sexual development in puberty. A disruption in this chain of events causes a deficiency of the sex hormones and prevents normal sexual maturity.

Thus GPR54 antagonists may suppress the release of gonadotropic hormones and such compounds would be novel orally available drugs for sex hormone dependant diseases including prostate cancer and endometriosis. Current research has focused on developing safer GPR54 antagonists[7]. The development of drugs from this class of compounds through lead optimization or through sophisticated computer-aided drug design (CADD) techniques. The present QSAR study on various Pyridines attempts to address this need by arriving at the physico-chemical properties required for high specific GPR54 Antagonistic activity in the form of a mathematical equation, according to the QSAR analysis. This study should therefore help in designing newer molecules with better specific GPR54 Antagonistic activity.

2. MATERIALS AND METHODS

Data set for analysis

The first step in developing QSAR equations is to compile a list of compounds for which the experimentally determined inhibitory activity is known. The GPR54 antagonistic activity data of 2-acylamino-4, 6-diphenyl pyridine derivatives were taken from the reported work of *Kobayashi et al.* 2010 [7]. The list of reported compounds with their IC_{50} values is given in table 1. The biological activity data (IC_{50} in μM) was converted to negative logarithmic dose (pIC50in moles) for QSAR analysis. For the external validation of QSAR models, the molecules were rationally divided into training having 21 and test set having 6 compounds on the basis of structural diversity and cover the complete range of variations in inhibitory activity as the guidelines for dividing into training and test sets [8].

Chemical Structure Construction and optimization

The molecular modeling study was performed by using the ChemOffice 2006 software, supplied by Cambridge Software Company, USA. The two-dimensional (2D) structures were transformed into three dimensional (3D) structures by using the Chem3D Ultra 10.0 module. The structure of the molecules was drawn and saved as cdx, chm file. The resulting 3D structures were then subjected to an energy-minimization by using the molecular mechanics (MM2) method. The energy minimized molecules were re-optimizing using molecular orbital package (MOPAC). The numerical descriptors are responsible for encoding important features of the structure of the molecules and can be categorized as electronic, steric, and thermodynamic characters.

Descriptors calculation

The thermodynamic, spatial, electronic, and topological descriptors were calculated for QSAR analysis. The thermodynamic parameters describe free energy change during drug receptor complex formation. Spatial parameters were quantified for steric feature of drug molecules required for its complimentary fit with the receptor. Electronic parameters describe weak non-covalent bonding between drug molecules and the receptor. To avoid the local stable conformations of the compounds, geometry optimization was run many times with different starting points for each molecule, and conformation with the lowest energy was considered for calculation of the molecular descriptors. Various physicochemical parameters belonging to different classes' viz. hydrophobic, steric and electronic etc. were calculated using "ChemDraw 3D Ultra" program and values of the calculated parameters are given in table 2.

Division of Training and Test Set

It is proven that the only way to estimate the true predictive power of a model is to test it on a sufficiently large collection of compounds from an external test set. The test set must include not less than five compounds, whose activities and structure must cover the range of activities and structures of compounds from the training set. This application is necessary for obtaining trustful statistics for comparison between the observed and predictive activities for these compounds. In this series 9 compounds were selected as a test set and remaining 21 compounds as training set and shown in table 1. The test set was used for the validation of model.

Statistical Analysis

The correlation matrix to show inter-correlation among the parameters is shown in table 3. Descriptors were selected for

the final equation having inter-correlation coefficient below 0.5 were considered. The QSAR models were developed by multiple linear regression (MLR) analysis using VALSTAT software.[9] The regression methods are used to build a model in the form of an equation that gives relationship between dependent variable (usually biological activity) and independent variable ("descriptors"). The model can then be used to predict activities for new molecules. The best QSAR model has characters of large F, low error s, low p-value, r^2 and q^2 value close to 1, as well as P <0.001. The large F means proposed regression model fits the data well. The low error means less standard deviation of the sampling distribution associated with the estimation method. The lower the p-value, more "significant" the result is, in the sense of statistical significance. The r^2 and q^2 value close to 1 means model explained well the activity variations in the compounds. Internal and external validation was performed to validate the QSAR model. In this approach, the activity of each compound in test set is computed. With the help of observed activity and calculated activity cross-validation coefficient q^2 was calculated. Cross-validation coefficient q^2 can be considered as an indicator of the predictive performance and stability of a model. For a reliable model the square of cross-validation coefficient q2 should be ≥0.5.

3. RESULTS AND DISCUSSION

When data set of 21 compounds was subjected to stepwise multiple linear regression analysis, in order to develop QSAR model, several models were obtained. The different QSAR models were developed using 21 training compounds and the best equation was obtained by using the optimal combination of descriptors. The stepwise development of model along with changes in statistical qualities on gradual addition of descriptors was done. The results suggest that the model-IV is best models using 17 compounds among other significant models and are good enough to rank molecular activities for further drug discovery.

Model: I

BA = [7.38232(± 0.358143)] +Vc [3.55547e-005(± 2.56433e-005)] +Pc [3.3048e-005(± 2.0108e-005)] +Y(PM) [-0.000748346(± 0.000512923)]

n=20, r=0.822557, r^2=0.6766, r^2adj=0.615962, variance=0.36814, std=0.606745, QF=1.35569, PE=0.0482097, F=11.1581, FIT=1.15429, LOF=13.9414, AIC=0.55221

Standard F max value at 95% confidence=11.1581. This model explains 67% of GPR54 antagonistic activity for the training set. The compound 1 is shown as outlier. So by removing this outlier, we developed model II.

Model: II

BA = [7.51287(± 0.284541)] +Vc [3.71539e-005(± 1.961e-005)] +Pc [3.2263e-005(± 1.53659e-005)] +Y(PM) [-0.000855788(± 0.000397039)]

n=19, r=0.899126, r^2=0.808428, r^2adj=0.770113, variance=0.211978, Std.= 0.460411, QF=1.95288, PE=0.0292998, F=21.0998, FIT=2.32826, LOF=7.97127, AIC=0.298095

Standard Fmax value at 95% confidence=21.0998. This model explains 80.8% of GPR54 antagonistic activity for the training set. The compound 19 is shown as outlier. So by removing this outlier, we developed model III.

Model: III

BA = [7.60346(± 0.221667)] +Vc [3.84795e-005(± 1.48226e-005)] +Pc [3.20661e-005(± 1.15982e-005)]+Y(PM)[-0.000907935(±0.000301332)

n=18, r=0.945675, r^2=0.8943,r^2adj=0.87165, variance=0.119085, std=0.345087, QF=2.7404, PE=0.0166091, F=39.4836, FIT=4.66801, LOF=4.46421, AIC=0.156299

standard Fmax value at 95% confidence=39.4836, compound 9l is shown as outlier so by removing this outlier model IV was developed.

MODEL IV:

BA = [7.52292(± 0.176554)] +Vc [3.7487e-005(± 1.1308e-005)] +Pc [3.26806e-005(± 8.84284e-006)] +Y(PM) [-0.000844437(± 0.000233098)]

n=17, r=0.968213, r^2=0.937436, r^2adj=0.922999, q^2=0.849, variance= 0.0678879, std.=0.260553, QF=3.71599, PE=0.0101159, F=64.9296, FIT=8.2669, LOF=2.55055, AIC=0.0827384, standard Fmax value at 95% confidence=64.9296

The QSAR model IV was selected as best model among the four models on the basis of statistical parameters discussed earlier. This model explains 93.7 % of GPR54 antagonistic mactivity for the training set. The internal predictive power of the model was confirmed by LOO cross validation (q^2) and q^2 value 0.849 indicates robust model. F-ratio is 64.9296 shows that the model is 95% significant. Variance inflation factor (VIF) is a measure of multicollinearity. Data in the model of the series shows that there is absence of any serious multicollinearity problem among the descriptors. (VIF<10)

The model IV was internally and externally validated with the training and test set respectively and judged by LOO method. The observed and calculated biological activity of training and test set by using model III is given in table 4 and 5 respectively.

It is evident from the QSAR studies that model-IV shows that Pc (critical pressure), Vc (critical volume) and Y (PM) (the principal moment of inertia), were affecting the biological activity. Pc and Vc were positively correlated while Y (PM) was negatively correlated to the biological activity and all these three parameters became key parameter in studies of the environmental fate of

chemicals. Partition coefficient in octanol-water system plays an important role in transportation of molecules through lipid membranes. In this model Pc and Vc are positively affecting which means compounds should have high pressure and high volume groups for higher biological activity. Positive contribution of Pc (critical pressure) & Vc (critical volume) to the biological activity indicates that maximizing the increase of critical pressure na volume both of the molecules does favor for activity. And the Y (PM) indicates A descriptor that calculates the moment of inertia and radius of gyration. Moment of inertia (MI) values characterize the mass distribution of a molecules related to the MI values, ratios of the MI values along the three principal axes are also well know modeling variables. This descriptor calculates the MI values along the X, Y and Z axes as well as the ratio's X/Y, X/Z and Y/Z. Finally it also calculates the radius of gyration of the moments. The shadow area descriptors align the rest two moments of inertia of the molecule along the X and Y axes and then calculate the area of the projection of the molecule on the XY, XZ and YZ planes. In general these types of descriptors capture features related to molecular size and shape and thus are generally physically interpretable. The drawback to these descriptors is that they require accurate molecular geometries and thus for large sets of molecules the optimization step can become time consuming, energy of the molecule increases the activity.

4. CONCLUSION

It was observed from the selected QSAR models that biological activities of derivatives are governed by thermodynamic, electronic and steric properties of the molecules. The models also suggest about the groups that responsible to increase the activity. This information can be explored for the designing of new molecules having better anticancer activity.

5. ACKNOWLEDGEMENT

Authors are thankful to Head, department of Pharmaceutical Science, V.N.S. Institute of Pharmacy, Bhopal (M.P.) for providing necessary facilities.

REFERENCES

1. Lemon SM, Thomas DL. Vaccines to prevent viral hepatitis. N Engl J Med. 1997; 336: 196–204.

2. Pae HO, Chung HT. Induction of apoptosis by 4-acetyl-12, 13-epoxyl-9-trichothecene-3,15-diol from Isaria japonica Yasuda through intracellular reactive oxygen species formation and caspase-3 activation in human leukemia HL-60 cells, Toxicology in Vitro. 2003;17: 49-57.

3. Bhalla KN. Microtubule-targeted anticancer agents and apoptosis. Oncogene. 2003; 22: 9075 -9086.

4. Sui XC, Bao N, Shaojuan J, John H, John G, Sanjeeva R, Ben T, John D. A role for transferrin receptor in triggering apoptosis when targeted with gambogic acid. J Med Chem. 2003; 46: 2474-2481.

5. Maristella CA, Patricia SM, Nelson D, Marcela H. Dehydrocrotonin and its derivative, dimethylamide-crotonin induce apoptosis with lipid peroxidation and activation of caspases-2, -6 and -9 in human leukemic cells HL60. Toxicology. 2004; 203: 123-137.

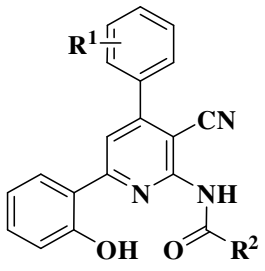
6. Kobayashi T, Sasaki S, Tomita N, Fukui S, Kuroda N, Nakayama M, Kiba A, Takatsu Y, Ohtaki T, Itoh F, Baba A. 2-acylamino-4,6-diphenylpyridine derivatives as novel GPR54 antagonists with good brain exposure and in vivo efficacy for plasma LH level in male rats, Bioorg Med Chem, 2010; 15 & 18(14): 5157-5171.

7. Kobayashi T, Sasaki S, Tomita N, Fukui S, Kuroda N, Nakayama M, Kiba A, Takatsu Y, Ohtaki T, Itoh F, Baba A. Synthesis and structure–activity relationships of 2-acylamino-4,6-diphenylpyridine derivatives as novel antagonists of GPR54. Bioorg Med Chem. 2010; 18 : 3841–3859.

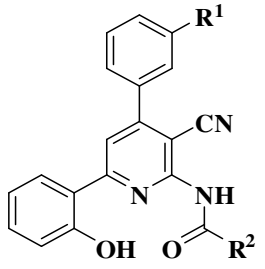
8. Opera TJ, Waller GL., Marshall GR. 2D-QSAR Studies on Anthranilic Acid Derivatives: A Novel Class of Allosteric Inhibitors of Hepatitis C NS5B Polymerase. J Med Chem. 1994; 37: 2206-2215.

9. Valstat Software [A regression analysis program specially designed for generation of QSAR models & their Validation program developed by Mr. Arun Kumar Gupta, Department of pharmacy SGSITS, Indore (M.P.)]

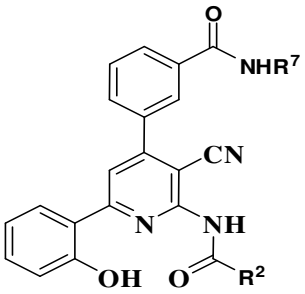
Table 1: Human GPR54 binding affinities and antagonistic activities

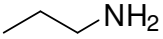
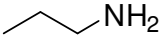
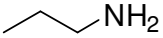
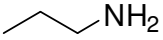
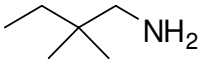
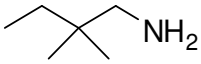
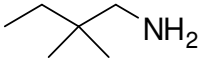


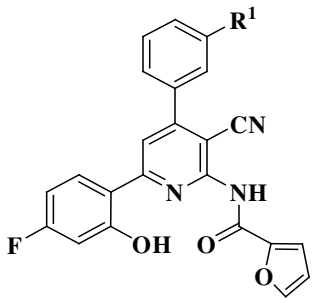
Compound	R^1	R^2	IC_{50} (µm)	BA(-log IC_{50})
1	3'-CH_2OH	2-Thienyl	1.5	5.8239
11*	3'-CO_2H	2-Thienyl	9.9	5.0044
7a	NH_2	2-Thienyl	2.1	5.6777
19	3'-CH_2NH_2	2-Thienyl	0.57	6.244
9a	3'-$NHCOCH_2NH_2$	2-Thienyl	0.051	7.2924
9b	3'-$NHCO(CH2)_2NH_2$	2-Thienyl	0.021	7.6778
9c	3'-$NHCO(CH_2)$2NH2	Phenyl	0.051	7.2924
9d*	3'-$NHCO(CH_2)_3NH_2$	2-Thienyl	0.034	7.9685
13a	3'-$CONH(CH_2)_2NH_2$	2-Thienyl	33	4.4814
13b*	3'-$CONH(CH_2)_2NH_2$	2-Thienyl	0.045	7.3467
13c	3'-$CONH(CH_2)_2NH_2$	2-Thienyl	0.12	6.9208



Compound	R^1	R^2	R^3	IC_{50} (µm)	BA(-log IC_{50})
25*	3'-$NHCO(CH_2)_2NH_2$	Phenyl	2'-NH_2	0.89	6.05
9e	3'-$NHCO(CH_2)_2NH_2$	2-Thienyl	2'-OH-5'-Cl	0.49	6.3098
9f	3'-$NHCO(CH_2)_2NH_2$	2-Thienyl	2'-OH-5'OMe	0.48	6.31875
9g*	3'-$NHCO(CH_2)_2NH_2$	2-Thienyl	2'-OH-4'-F	0.0074	8.13076
9h	3'-$NHCO(CH_2)_2NH_2$	2-Thienyl	2'-OH-4'-Cl	0.0098	8.00877
9i	3'-$NHCO(CH_2)_2NH_2$	2-Thienyl	2'-OH-4'-Br	0.019	7.72125
9j*	3'-$NHCO(CH_2)_2NH_2$	2-Thienyl	2'-OH-4'-Me	0.012	7.79208
9k	3'-$NHCO(CH_2)_2NH_2$	Phenyl	2'-OH-4'OMe	0.12	6.9208



Compound	R^7	R^2	IC_{50} (µm)	BA(-log IC_{50})
27a*		2-Thienyl	0.030	7.5228
27b		2-Furyl	0.014	7.8538
27c*		2'-MeOPh	0.045	7.34678
27d		4'-MeOPh	0.031	7.5086
27e		2-Furyl	0.015	7.8239
27f		2'-MeOPh	0.032	7.4948
27g		4'-MeOPh	0.032	7.4948



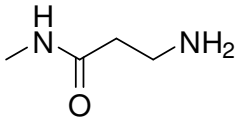
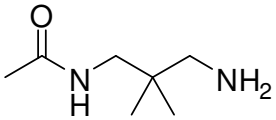
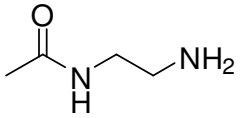
Compound	R^1	IC_{50} (µm)	BA(-log IC_{50})
91		0.0037	8.4317
13d*		0.0080	8.0969
13e		0.011	7.9586

Table 2: The descriptors values (independent variables) of training and test set calculated by ChemDraw 3D Ultra software

Compound	Vc	Pc	Y(PM)
1	1153.5	25.0501	4969.03
11*	1158.5	24.975	5562.73
7a	1107.5	26.653	4890.18
19	1163.5	24.195	4910.16
9a	1247.5	24.532	7458.96
9b	1303.5	22.355	7783.64
9c	1347.5	19.93	6269.32
9d*	1359.5	20.456	6363.79
13a	1303.5	22.355	6720.2
13b*	1303.5	22.355	8230.03
13c	1303.5	22.355	9007.63

25*	1360.5	18.154	6217.6
9e	1352.5	21.295	8964.6
9f	1377.5	19.877	8918.32
9g*	1321.5	21.139	8497.81
9h	1352.5	21.295	9419.5
9i	1365.5	23.542	12111.4
9j*	1359.5	20.181	8351.65
9k	1377.5	19.877	9732.22
27a*	1303.5	22.355	8198.7
27b	1278.59	22.292	8073.68
27c*	1580.5	16.116	8839.74
27d	1580.5	16.116	11991.1
27e	1351.5	17.347	7575.98
27f	1653.5	12.995	9227.81
27g	1653.5	12.995	10959.6
9l	1296.5	21.081	9024.93
13d*	1296.5	21.081	8716.63
13e	1397.5	18.092	10677

Table 3: Correlation matrix of descriptors used in model IV

	Vc	**Pc**	**Y(PM)**
Vc	1.000000		
Pc	0.014123	1.000000	
Y(PM)	0.132552	0.193415	1.000000

Table 4**:** Observed and calculated biological activity values of training set compounds

compound	Calculated BA	Observed BA
1	3.41	5.8239
7a	3.43	5.6778
19	3.39	6.244
9a	1.271	7.2924
9b	0.999	7.6778
9c	2.280	7.2924
13a	1.897	4.4814
13c	- 0.033	6.3098
9e	0.004	6.31875
9f	0.0442	6.3098
9h	- 0.379	8.00877
9i	- 2.652	7.72124
9k	- 0.6430	6.9208
27b	- 6.769	7.8538
27d	- 2.506	7.5086
27e	1.176	7.8239
27f	- 0.206	7.4948
27g	- 1.669	7.4948
9l	- 0.048	8.4317
13e	- 1.440	7.9586

Table 5: Observed and calculated biological activity values of test set compounds

compound	Calculated BA	Observed BA
11	2.87	5.8239
9d	2.2	5.00436
13b	0.6227	7.2924
25	2.3241	8.00877
9g	0.3972	7.721246
9j	0.6512	6.9208

27a	0.6493	7.8538
27c	0.1180	7.4948
13d	0.2115	8.0969