doi.org/10.36721/PJPS.2022.35.3.REG.841-844.1

Using QSAR calculations on benzamide derivatives to inhibit reproduction in endothelial cells by CORAL SEA

Ghasem Ghasemi* and Nima Nasiri
Department of Chemistry, Rasht Branch, Islamic Azad University, Rasht, Iran

Abstract: The vascular endothelial growth factor is one of the most potent tumor angiogenic factors. The best descriptors with the imperialist competitive algorithm are MATS2m, DISPe, G(O…Cl), RDF090m, Mor30v, and R2u. Masses, electro negativities, van der Waals volumes, the d COMMA2 value, the sum of the geometrical distance between O…Cl and autocorrelations are important in our study. The SMILES files have been used with Coral Sea software. We had: n =5, R^2=0.5857, Q^2=0.5170, F=4 (Calibration set), n =11, R^2=0.9842, Q^2=0.9807, F=560 (Training set), n = 5, R^2 = 0.9999, Q^2 = 0.9995, F = 25054 (invTrain).

Keywords: Endothelial cells, QSAR, benzamide derivatives, ICA, CORAL SEA.

INTRODUCTION

Angiogenesis, plays an important role in tumor growth. The inhibition of angiogenesis can suppress the progression of tumor growth (Bird *et al*., 2013; Boeldt *et al*., 2015; Boeldt *et al*., 2011; Canti *et al*., 2012; Corcoran *et al*., 2014; Dekel *et al*., 2010; Drost *et al*., 2012; Dusse *et al*., 2013; Krupp *et al*., 2013; Liu *et al*., 2012; Penack *et al*., 2011; Leonhardt *et al*., 2013; Cutler *et al*., 2010).

SMILES can be used for the elucidation of the molecular structure in the analyses (Toropova and Toropov, 2013; Ibezim *et al*., 2012; Roy *et al*., 2012; Toropov *et al*., 2012; Hasand *et al*., 2014).

We have:

^{Graph} DCW= $\Sigma CW(Ak)+ \alpha\Sigma CW(^0ECk)+ \beta\Sigma CW(^1ECk)+ \gamma\Sigma CW(^2ECk)+ \delta\Sigma CW(^3ECk)$ (1) CW(x) is the correlation weight.

^{SMILIS} DCW is:

$^{SMILIS}DCW= \alpha\Sigma CW(Sk)+\beta\Sigma CW(SSk)+\gamma\Sigma CW(SSSk)+x.CW(NOSP)+y.CW(HALO)+z.CW(BOND)+\lambda.CW(ATOMPAIR)$ (2)

$RSS= \Sigma (Y_i-\bar{Y}_i)^2$ (3)

Total Sum of Squares (TSS) = $\Sigma (Y_i-\bar{Y})^2$ (4)

$\bar{Y}$ is the mean of the observed data $R^2= 1-\frac{RSS}{TSS}$ (6)

R^2 (LOO-Q2) is: $Q2 = 1 - \frac{\Sigma(Ypred-Y)2}{\Sigma(Y-\hat{Y})2}$ (7)

Y indicates observed activity values.

The *in vivo* antitumor activity of compounds was evaluated in the Calu-6 xenograft model. Antiangiogenic activity was confirmed by Microvessel density reduction (15%) in the xenograft tissue. Analogues were selected for a vascular endothelial growth factor receptor inhibition assay (Hada *et al*., 2012).

MATERIALS AND METHODS

Actual half-maximal inhibitory concentration (IC_{50}) values of all compounds were selected from literature (Hada *et al*., 2012). We had 21 compounds (fig. 1). Three random splits were: Subtraining, calibration, and test sets. Optimal smiles have been calculated by Coral Sea software (Version 2017). An ICA was used to create the QSAR models. The structures of all the compounds were optimized with the DFT method (B3LYP/6-31G(d)) using Gaussian software. The descriptors were calculated using DRAGON (Version 5.5). The quantitative structure activity relationship (QSAR) models were generated using a testing set of three molecules. Subsequently, 3,226 theoretical descriptors were calculated. Therefore, it was necessary to reduce the number of descriptors. ICA was used to create QSAR models.

STATISTICAL ANALYSIS

For 21 compounds, IC_{50} values have given from experimental work. The structures of all the compounds were optimized with using Gaussian (09) software. The descriptors were calculated using DRAGON (Version 5.5). We have to reduce the number of descriptors. ICA was used to create the QSAR models with Matlab(2016). SVM, LASSO and Jack-knife were generated using SPSS(23) and Unscrambler(10.4). The Monte-Carlo method has been calculated by Coral Sea software (Version 2017). The best descriptors were selected. We have evaluated RMSE, R^2, R, Q2, EPS, S, MAE and F in different statistical methods. The combined methods were better. ICA-SVM was better than SVM regression. The Monte-Carlo method and LASSO were the best models in different ways.

**Corresponding author:* e-mail: ghasemi@iaurasht.ac.ir

RESULTS

Table 1 presents the experimental results, and the calculations were done with the pIC_{50} values. table 2 contains correlation weights for the DCW calculation.

Table 3 contains representations of analyzed compounds and three of their distributions into the training, test, and validation sets.

Symbols ‘)’ and ‘]’ are branching and ions, respectively. The SMILES and correlation weights are shown in table 1.

Intercept (c0) and slope (c1) were calculated for each set individually:
Training set: c_0= -3.10462 c_1= 0.02209
InvTraining set: c_0= -6.68229 c_1= 0.06804
Calibration set: c_0= -3.68639 c_1= 0.02818
Slope and intersect were calculated with subtraining set to give the model:

pIC_{50} = -3.1046166 (± 0.0059835) + 0.0220918 (±0.0000716) * DCW (1, 25)

The most significant descriptors that were selected by ICA) are: MATS2m, DISPe, G (O…Cl), RDF090m, Mor30v, and R2u (table 4).

The RMSE, R, P-value, Q2, R^2 pred and R^2 value calculated using ICA, are shown in table 5.

Table 2: Correlation weights over three probes of the Monte Carlo optimization

Structural attribute(SA)	CW(SA)
EC0-C...4...	0.7491
EC0-C...3...	0.0026
EC0-O...2...	2.2469
EC0-O...1...	3.9331
EC0-H...1...	-1.2514
C.....	1.2535
O.....	2.8742
1.....	5.2481
=.....	1.6865
(.....	-1.6240
2.....	4.4423
O...C.....	3.5007
C...1.....	2.5669
=...1.....	6.4990
C...=.....	0.5021
C...C.....	2.3751
C...(.....	-0.6836
O...C.....	3.5007
C...2.....	5.5629
=...2.....	4.1235
2...(.....	0.3735
=...(.....	2.2455
O...=.....	0.3757
$10001000000	6.1838

Table 1: Experimental and calculated pIC_{50} and DCW for benzamide derivatives to inhibit reproduction in endothelial cells

SMILES	DCW	Expr	Calc
+:COC1=CC=C(C=C1COC2=CC(=C(Cl)C=C2)C)C(C)=O	75.58903	-1.4100	-1.4347
+:COC1=CC=C(C=C1COC2=CC(=CC=C2)C)C(C)=O	90.35589	-1.1100	-1.1085
+:COC1=CC=C(C=C1COC2=CC(=C(Cl)C=C2)C)C(C)=O	75.58903	-1.4700	-1.4347
+:COC(=O)C1=CC=C(O)C(=C1)COC2=CC(=C(Cl)C=C2)C	77.32605	-1.3900	-1.3963
+:COC1=CC=C(C=C1COC2=CC(=C(Cl)C=C2)C)C#N	83.04404	-1.2700	-1.2700
+:COC1=CC=C(C=C1C(=O)NC2=CC=C(Cl)C=C2)C(N)=O	77.58256	-1.3900	-1.3907
+:COC1=CC=C(C=C1\C=C\C2=CC=C(Cl)C=C2)C(N)=O	77.27434	-1.4300	-1.3975
+:COC1=CC=C(C=C1\C=C/C2=CC=C(Cl)C=C2)C(N)=O	78.26049	-1.3400	-1.3757
+:CNC(=O)C1=CC=C(OC)C(=C1)/C=C/C2=CC=C(Cl)C=C2	65.64009	-1.6500	-1.6545
+:COC1=CC=C(C=C1/C=C/C2=CC=C(Cl)C=C2)C(=O)N(C)C	63.65444	-1.7000	-1.6984
+:COC1=CC=C(C=C1\C=C/C2=CC=C(Cl)C=C2)C(=O)NCC(O)CO	80.82807	-1.3200	-1.3190
-:CC1=C(Cl)C=CC(=C1)OCC2=CC(=CC=C2)C#N	76.17347	-1.5000	-1.4218
-:COC1=CC=C(C=C)C=C1COC2=CC(=C(Cl)C=C2)C	75.74324	-1.5300	-1.4313
-:NC(=O)C1=CC=C(OC=C)C(=C1)/C=C/C2=CC=C(Cl)C=C2	70.32180	-1.9000	-1.5511
-:NC(=O)C1=CC=C(OCC=C)C(=C1)/C=C/C2=CC=C(Cl)C=C2	73.31649	-1.6900	-1.4849
-:COC1=CC=C(C=C1/C=C/C2=CC=C(Cl)C=C2)C(=O)NC=C	78.82307	-1.3200	-1.3633
#:COC1=CC=C(C=C1COC2=CC(=C(Cl)C=C2)C)C(N)=O	81.93502	-1.5000	-1.2945
#:COC1=CC=C(C=C1COC2=CC=C(Cl)C=C2)C(N)=O	79.03482	-1.5100	-1.3586
#:CN(C)CCOC1=CC=C(C=C1/C=C/C2=CC=C(Cl)C=C2)C(N)=O	73.34056	-1.4400	-1.4844
#:CN1CCN(CCOC2=CC=C(C=C2/C=C/C3=CC=C(Cl)C=C3)C(N)=O)CC1	67.53614	-1.9000	-1.6126
#:COC1=CC=C(C=C1/C=C/C2=CC=C(Cl)C=C2)C(=O)NCCO	79.38889	-1.3400	-1.3508

Table 3: Statistical Quality of the CORALSEA program (2018)

	n	R^2	CCC	IIC	Q2	s	MAE	F	RMSE
Training	10	0.9198	0.9582	0.6393	0.8910	0.049	0.037	92	0.0490
InvTrain	5	0.9897	0.4529	0.8536	0.9589	0.180	0.123	289	-
Calib	5	0.5371	0.6440	0.7326	0.7607	0.181	0.153	3	0.1699

Table 4: Selected parameters with different hidden units using ICA

Layer	0 or 1
2	00001011111001101000011111001110111110001100
3	01011001110011100101001111001101101111101000
4	01001010011111101110101010001000100111000000
5	01001000011011101110001000101100000100011110
6	11001100001010000010001000111001111111100001
7	01001010100011101010001001110100101101100010
8	10100111000011000110000011110000000101111110
9	01001000111111101110000000101100000101011110
10	01111010011111101110001000101100000100000100
11	11111110001111001010010101100100100101111101
12	00101100011010101010010110000101100110101000

Table 5: The statistical parameters of different constructed QSAR models (Matlab (2016),SPSS (23), Unscrambler (10.4))

Method	RMSE test	RMSE train	R^2	R	Q2	R^2 pred	Expected prediction std. error
ICA	0.1561	0.1351		0.7063			
ICA-SVM	0.1901	0.1208				0.7537	
SVM	0.1990	0.1275				0.7197	
ICA-LASSO			0.9960				0.076
Jack- Knife					0.8839		

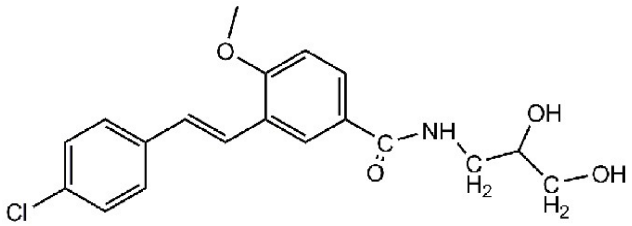

Fig. 1: The Structure of Benz amide Derivatives

DISCUSSION

The aforementioned models include the following: N is the number of compounds in the set; R is correlation coefficient; Q is cross-validated correlation coefficient; CCC is concordance correlation coefficient; IIC is index of ideality of correlation; s is standard error of estimation; MAE is mean absolute error; F is Fischer F-ratio. According to the statistical indices, the model has a good quality. The optimal descriptors were evaluated with Monte Carlo method.

If CW(x) < 0, it implies the decrease of target and vice versa; however, if CW(x) > 0-, it implies the increase of the target. The most significant descriptors are: MATS2m, DISPe, G(O…Cl), RDF090m, Mor30v, and R2u. Also important for our study are: Masses, electro negativities, van der Waals volumes, the d COMMA2 value, the sum of the geometrical distance between O…Cl and autocorrelations.

The presence of branching in the stable compound increases pIC_{50}. The presence of oxygen in the stable compound decreases pIC_{50}.

CONCLUSION

The CORAL SEA software gives models of reproduction in endothelial cells of good statistical quality. The statistical quality is dependent on the distribution of three

probes. These data can use for reproduction in endothelial cells.

ACKNOWLEDGMENTS

The author would like to thank the Research vice President of the Islamic Azad University, Rasht Branch for their support.

REFERENCES

Bird IM, Boeldt DS, Krupp J, Grummer MA, Yi FX and Magness RR (2013). Pregnancy, programming and preeclampsia: Gap junctions at the nexus of pregnancy-induced adaptation of endothelial function and endothelial adaptive failure in PE. *Curr. Vasc. Pharmacol*., **11**(5): 712-729.

Boeldt DS, Grummer MA, Yi F, Magness RR and Bird IM (2015). Phosphorylation of Ser-279/282 and Tyr-265 positions on Cx43 as possible mediators of VEGF-165 inhibition of pregnancy-adapted Ca^{2}+ burst function in ovine uterine artery endothelial cells. *Mol. Cell. Endocrinol*., **412**: 73-84.

Boeldt DS, Yi FX and Bird IM (2011). eNOS activation and NO function: Pregnancy adaptive programming of capacitative entry responses alters nitric oxide (NO) output in vascular endothelium - new insights into eNOS regulation through adaptive cell signaling. *J. Endocrinol*., **210**: 243-258.

Canti V, Maggio L, Ramirez GA, Locatelli A, Cozzolino S, Ramoni V, Ruffatti A, Tonello M, Valsecchi L and Rosa S (2012). Hypertension negatively affects the pregnancy outcome in patients with antiphospholipid syndrome. *Lupus*., **21**: 810-812.

Corcoran JJ, Nicholson C, Sweeney M, Charnock JC, Robson SC, Westwood M and Taggart MJ (2014). Human uterine and placental arteries exhibit tissue-specific acute responses to 17 beta-estradiol and estrogen-receptor-specific agonists. *Mol. Hum. Reprod*., **20**(5): 433-441.

Cutler C, Kim HT and Ayanian S (2010). Prediction of veno-occlusive disease using biomarkers of endothelial injury. *Biol. Blood Marrow Transplant*., **16**(8): 1180-1185.

Dekel N, Gnainsky Y, Granot I and Mor G (2010). Inflammation and implantation. *Am. J. Reprod. Immunol*., **63**(1): 17-21.

Drost JT, Arpaci G, Ottervanger JP, de Boer MJ, van Eyck J, van der Schouw YT and Maas AH (2012). Cardiovascular risk factors in women 10 years post early preeclampsia: The preeclampsia risk evaluation in females study (PREVFEM). *Eur. J. Prev. Cardiol*., **19**(5): 1138-1144.

Dusse LM, Alpoim PN, Lwaleed BA, de Sousa LP, Carvalho M and Gomes KB (2013). Is there a link between endothelial dysfunction, coagulation activation and nitric oxide synthesis in preeclampsia?. *Clinica. Chimica. Acta*., **415**: 226-229.

Hada K, Suda A and Asoh K (2012). Angiogenesis inhibitors identified by cell-based high-throughput screening: Synthesis, structure-activity relationships and biological evaluation of 3-[(E) -styryl]benzamides that specifically inhibit endothelial cell proliferation. *Bioorg. Med. Chem*., **20**(4): 1442-1460.

Hasand VH, Tropov AA, Tropova AP and Mahajan DT (2014). QSAR models for Anti malaria Activity of 4 - Amonoquinolines, *Curr. Comput. Aided. Drug. Des*., **10**(1): 75-82.

Ibezim E, Duchowicz PR, Ortiz EV and Castro EA (2012). QSAR on arylpiperazine derivatives with activity on malaria, *Chemometr. Intell. Lab. Syst*., **110**: 81-88.

Jodkowska A, Martynowicz H, Kaczmarek-Wdowiak B and Mazur G (2015). Thrombocytopenia in pregnancy - pathogenesis and diagnostic approach. *Postępy Higieny i Medycyny Doswiadczalnej*., **69**: 1215-1221.

Krupp J, Boeldt DS, Yi FX, Grummer MA, Bankowski Anaya HA, Shah DM and Bird IM (2013). The loss of sustained $Ca(^{2+})$ signaling underlies suppressed endothelial nitric oxide production in preeclamptic pregnancies: Implications for new therapy. *Am. J. Physiol. Heart. Circ. Physiol*., **305**: H969-H979.

Leonhardt F, Grandmami S and Behe M (2013). Inflammatory neovascu- larization during graft-versus-host disease is regulated by a v integrin and miR-100. *Blood*, **121**(17): 3307-3318.

Liu Z, Afink GB and Dijke PT (2012). Soluble fms-like tyrosine kinase 1 and soluble endoglin are elevated circulating anti-angiogenic factors in pre-eclampsia. *Pregnancy Hypertension*, **2**(4): 358-367.

Penack O, Gerard Socie'G and van den Brink MRM (2011). The importance of neovascularization and its inhibition for allogeneic hematopoietic stem cell transplantation. *Blood*, **117**(16): 4181-4189.

Roy K, Mitra I, Kar S, Ojha PK, Das RN and Kabir H (2012). Comparative studies on some metrics for external validation of QSPR models, *J. Chem. Inf. Model.*, **52**(2): 396-408.

Toropov AA, Toropova AP, Benfenati E, Gini G, Puzyn T, Leszczynska D and Leszczynski J (2012). Novel application of the CORAL software to model cytotoxicity of metal oxide nanoparticles to bacteria Escherichia coli. *Chemosphere*, **89**(9):1098-1102.

Toropova AP and Toropov AA (2013). Optimal descriptor as a translator of eclectic information into the prediction of membrane damage by means of various TiO2 nanoparticles, *Chemosphere*, **93**(10): 2650-2655.