King Saud University

Arabian Journal of Chemistry

www.ksu.edu.sa
www.sciencedirect.com

ORIGINAL ARTICLE

QSAR study on 4-alkynyldihydrocinnamic acid analogs as free fatty acid receptor 1 agonists and antidiabetic agents: Rationales to improve activity

Rinnie [a], Vikas Gaba [a], Ketki Rani [a], Shilpa [a], Manish K. Gupta [a,b,*]

[a] Molecular Modeling and Pharmacoinformatics Lab, Department of Pharmaceutical Chemistry, ISF College of Pharmacy, Moga 142001, Punjab, India
[b] School of Pharmacy, Lloyd Institute of Management and Technology, Greater Noida 201306, UP, India

Abstract The free fatty acid receptor 1 (FFAR1) is a class A G-protein coupled receptor and a validated target to develop antidiabetic drugs. The present work describes the quantitative structure–activity relationship (QSAR) study of a series of 4-alkynyldihydrocinnamic acid analogs to rationalize their FFAR1 agonist activity. The various physicochemical and structural descriptors were derived from Molecular Operating Environment (MOE, 2011). The variable selection and model development were carried out using Combinatorial Protocol in Multiple Linear Regressions (CP-MLR) approach. The identified QSAR models highlighted the significance of solvent accessible surface area and associated hydrophobicity to the biological activity. The chemical features to improve hydrophobicity of 4-alkynyldihydrocinnamic acid analogs have been discussed.

Production and hosting by Elsevier

1. Introduction

The last decade witnessed a steep rise in the number of patients with type 2 diabetes mellitus (T2DM) especially in young and middle age individuals. The change in life style, irregular eating habits and lack of exercise are contributing factors for the development of T2DM. According to World Health Organization approximately, 347 million people worldwide are affected with diabetes. Here, T2DM alone accounts for more than 90% of total diabetes. It is estimated that the diabetes will be the 7th leading cause of death in 2030 (WHO, 2014). The T2DM is characterized by persistent hyperglycemia due to reduced insulin sensitivity, development of insulin resistance or impaired insulin secretion. If untreated, the hyperglycemia leads to the

serious pathological conditions such as coronary heart disease, atherosclerosis, nephropathy, neuropathy and retinopathy. T2DM is a major concern in patients with cardiovascular disease because cardiovascular deaths have been increased up to fourfold in diabetic patients when compared with their nondiabetic counterparts (Inzucchi and Sherwin, 2005). Several classes of drugs such as biguanide (metformin), sulphonylureas (gliclazide), thiazolidinediones (rosiglitazone), dipeptidyl peptidase-4 enzyme inhibitors (sitagliptin) are used as oral hypoglycemic agents, however, use of these drugs is often associated with serious risk factors (Stein et al., 2013). The biguanides and sulphonylureas force the β-cell to secrete insulin regardless of blood glucose levels and thus produce hypoglycemia. The hypoglycemia may lead to the loss of conscious (Tan et al., 2008; Graveling and Frier, 2009). These drugs also cause the loss of islet function (Maedler et al., 2005; Del Guerra et al., 2005). The side effects of mostly used thiazolidinediones also include fluid retention, weight gain, bone fracture and cardiovascular failure (Rizos et al., 2009). These drugs are also reported to be failed in providing desire hypoglycemic effect for the adequate duration of time (Kahn et al., 2006; Grant et al., 2007). Therefore, the development of new and effective antidiabetic drugs with improved patient compliance is of current interest (Waugh et al., 2010).

The free fatty acid receptor 1 (FFAR1 or GPR40) is a G-coupled, class A G-protein coupled receptor. FFAR1 is expressed in pancreatic β-cells and its activation by medium and longer chain free fatty acids (FFAs; endogenous ligands for FFAR1) results in enhancement of glucose stimulated insulin secretion (GSIS) (Itoh et al., 2003; Briscoe et al., 2003; Rayasam et al., 2007). The FFAR1 has been suggested as a potential drug target for T2DM. The studies have shown that FFAR1 agonists induce insulin secretion only when needed and thus prevent the patient from severe hypoglycemia (Tan et al., 2008; Feng et al., 2012). The various selective ligands including thiazolidinedione "glitazone" have been identified as FFAR1 agonists and antidiabetic agents (Kotarsky et al., 2003; Bharate et al., 2009).

Christiansen et al. (2011) synthesized a series of 4-alkynyldihydrocinnamic acid (hereafter referred as ADCA analogs) as FFAR1 agonists and antidiabetic agents (Table 1). These ADCA analogs have been prepared by modifying the ring B with substituted aryl and heteroaryls with an aim to optimize their lipophilicity for better activity (Table 1). These compounds have been evaluated for their efficacy to stimulate FFAR1 (Christiansen et al., 2011).

The quantitative structure–activity relationship (QSAR) study helps to optimize the desired physicochemical property(s) in a series of compound to improve their potency. The QSAR study assists medicinal chemists to select the optimum structural requirement to improve the activity of a series of compounds under investigation (Hansch et al., 2002). In QSAR study, the parameterization of chemical structure plays an essential role. In enumeration of chemical structures, it is important to note that in isolation, a data point is only a qualified number. A collection of such qualified numbers makes a variable or descriptor. In mathematical models each and every (independent; X) variable communicates with the target (dependent; Y) variable. A meaningful communication between X and Y variables results in the evolution of models with predictive value (Gupta and Prabhakar, 2008).

In this milieu, we have contemplated a QSAR study on FFAR1 stimulating activity of ADCA analogs. The aim of the study was to identify the specific physicochemical and structural requirement to improve the activity of ADCA analogs. For QSAR study, different physicochemical and structural descriptors have been calculated from the Molecular Operating Environment (MOE, 2011). The Combinatorial Protocol in the Multiple Linear Regression (CP-MLR) approach was used to develop the QSAR models (Prabhakar, 2003).

2. Material and methods

2.1. Dataset

A series of 4-alkynyldihydrocinnamic acid analogs (ADCA analogs) along with their FFAR1 stimulating activity in terms of pEC_{50} (negative logarithm of effective concentration which showed 50% activity) was selected for the study. The structure database of compounds was developed in MOE. The 3D structures were optimized by MMFF94x force-field as implemented in MOE. The optimized structure dataset was used for the parameterization of compounds. This has resulted in three hundred and seven non-zero descriptors depicting the physicochemical and topological properties of the ADCA analogs. The structure dataset was partitioned into training set (24 compounds) and test set (8 compounds) with the help of single linkage hierarchical clustering of all the descriptors (Table 1). A plot of molecular weights (MW) versus LogP of all ADCA analogs was used to compare the chemical space spanned by the training set and test set compounds (Fig. 1). The plot showed that the test set compounds cover almost entire chemical space corresponding to the training set compounds.

2.2. QSAR procedure

The QSAR models were developed using the CP-MLR procedure (Prabhakar, 2003). CP-MLR helps in identifying the information rich descriptors correspond to the biological activity under study from a large descriptor pool. It is a filter based variable selection procedure for QSAR and QSPR studies (Prabhakar, 2003; Gupta and Prabhakar, 2006; Saquib et al., 2007). Its implementation and detailed procedural aspects are given elsewhere (Gupta et al., 2005; Saquib et al., 2007; Prabhakar et al., 2006; Gupta and Prabhakar, 2008; Sharma et al., 2011; Kumar et al., in press). In CP-MLR, the filter-1 seeds the variables by the way of limiting inter-parameter correlations to a predefined level (default acceptable value ≤0.3); filter-2 controls the variables entry to a regression equation through t-values of coefficients (threshold value ≥2.0) which confer more than 95% regression coefficients significant level ; filter-3 provides comparability of equations with different numbers of variables in terms of square-root of adjusted multiple correlation coefficient of regression equation "r-bar" and filter-4 estimates the consistency of the equation in terms of cross-validated R^2 or Q^2. In the present study, the thresholds for filters-1, 2, 3 and 4 were kept 0.3, 2.0, 0.71 and $0.3 \leqslant Q^2 \leqslant 1.0$, respectively. Reassessment of the identified models was done for chance correlations by performing repeated randomization of the activity. Each identified model was subjected to 100 simulation runs with scrambled activity and the emerging

Table 1 Observed and predicted activities of alkynyldihydrocinnamic acid (ADCA) analogs
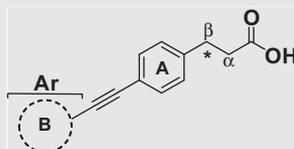
Comp.	Ar	logEC$_{50}$			
		Observed[b]	Predicted[c]		
		Obsd.	Eq. (1)	Eq. (2)	Eq. (3)
1	-phenyl	6.70	6.820	7.148	6.898
2[a]	-pyridin-2-yl	4.88	5.568	5.251	5.607
3	-pyridin-3-yl	5.22	5.637	5.656	5.675
4	-pyridin-4-yl	5.39	5.677	5.914	5.654
5	-6-methylpyridin-2-yl	5.21	5.857	5.441	5.777
6	-3-nitropyridin-2-yl	4.96	4.674	4.614	4.906
7	-2-chloropyridin-3-yl	6.22	6.355	6.005	6.460
8	-6-chloropyridin-3-yl	6.15	6.280	6.241	6.467
9	-2-fluoro-5-methylpyridin-3-yl	6.18	6.759	5.940	6.684
10	-2-chloropyrimidin-4-yl	5.04	5.160	5.487	5.256
11	-2-methoxypyrimidin-4-yl	5.02	4.956	5.085	4.431
12	-2-chloro-5-methylpyrimidin-4-yl	5.37	5.509	5.486	5.350
13	-2-chloro-6-methylpyrimidin-4-yl	5.56	5.437	5.462	5.314
14	-thiazol-5-yl	5.41	5.567	5.877	5.425
15[a]	-thiophen-2-yl	6.60	6.736	7.139	6.636
16	-3-methylthiophen-2-yl	7.12	7.084	7.263	6.669
17	-o-tolyl	7.34	7.046	7.039	6.901
18[a]	-2-methyl-pyridin-4-yl	5.95	5.828	5.882	5.763
19	-2-fluoropyridin-4-yl	6.29	6.259	6.816	6.423
20[a]	-2-chloropyridin-4-yl	6.53	6.353	6.389	6.458
21	-2-methoxypyridin-4-yl	6.28	6.090	5.490	5.770
22	-2-phenoxypyridin-4-yl	6.32	5.522	5.946	6.754
23	-2-phenyl-pyridin-4-yl	6.73	6.194	6.111	6.422
24[a]	-2-o-tolylpyridin-4-yl	5.97	6.285	6.347	6.672
25[a]	-2-fluoro-5-methylpyridin-4-yl	6.75	6.606	6.689	6.515
26	-2-fluoro-3-methylpyridin-4-yl	6.92	6.372	6.569	6.285
27	-2-chloro-3-methylpyridin-4-yl	6.87	6.601	6.531	6.552
28	-3-allyl-2-chloropyridin-4-yl	6.50	7.479	6.751	6.930
29[a]	-7-chloroquinolin-4-yl	7.10	6.955	6.699	6.935
30	-2,6-dichloropyridin-4-yl	7.39	6.874	7.170	7.246
31[a]	-2,6-dichloro-3-methylpyridin-4-yl	7.36	7.355	7.410	7.410
32[d]	-2,6-dichloropyridin-4-yl	7.37	7.341	7.501	7.125

[a] Test set compounds.
[b] Observed activity.
[c] Predicted activity by corresponding equation.
[d] Compound with a cyclopropane ring at α β carbons of parent structure.

correlations were counted to express the percent chance correlation of the model under examination (Prabhakar et al., 2004). The proposed models were validated through test set. The robustness of models was further evaluated by calculating their overall metric values (Ojha et al., 2011; Roy et al., 2012) and slopes (k and k') for the least squares regression line (Golbraikh and Tropsha) (Table 2). The metric is represented by the following equation:

$$r_m^2 = r^2 \left(1 - \sqrt{r^2 - r_0^2} \right)$$

Here r^2 and r_0^2 are correlations between the observed and predicted values with and without intercept, respectively, for the least squares regression lines. The metric r_m^2 values help to circumvent the overestimation of the quality of prediction due to

a wide response range (Y-range) (Ojha et al., 2011; Roy et al., 2012). The slopes k and k' were calculated by the following equations:

$$k = \frac{\sum Y_i \widetilde{Y}_i}{\sum \widetilde{Y}_i^2} \quad \text{and} \quad k' = \frac{\sum Y_i \widetilde{Y}_i}{\sum Y_i^2}$$

Here, Y and $\widetilde{Y}$ are observed and predicted activities, respectively. For a good predictive QSAR model the value of k and k' should be near to 1 (Golbraikh and Tropsha, 2002).

3. Results and discussion

In the first step of the study, CP-MLR procedure with default parameter settings was used to develop simple one and two
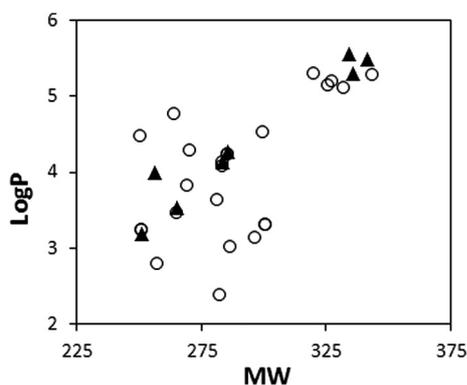
Figure 1 Plot of molecular weights (MW) versus logP of training (circle) and test (triangle) set of ADCA analogs.

descriptor models (base line models). This resulted in identification of three one-parameter and six hundred and sixty-four two-parameter models embedding total one hundred and sixty-five descriptors in them. These one hundred and sixty-five descriptors were selected to discover higher models i.e. four-parameter models to explain the activity of ADCA analogs. The following are selected four-parameter models (Eqs. (1) and (2)) from the pool of equations:

$$-\log EC_{50} = 0.661(0.117)a\_nCl + 0.601(0.223)a\_nF$$
$$- 1.158(0.128)a\_nN + 1.678(0.501)std\_dim2$$
$$+ 4.844 \quad (1)$$

$$n = 24, \quad r = 0.920 \quad Q^2_{(LOO)} = 0.743, \quad Q^2_{(L3O)} = 0.733,$$
$$s = 0.347, \quad F = 26.403$$

$$FIT = 2.640, \quad LOF = 0.215, \quad AIC = 0.184,$$
$$R^2_T = 0.857, \quad r_{yrand}(SD) = 0.424(0.137) \quad (2)$$

$$-\log EC_{50} = 0.127(0.026)SlogP\_VSA2$$
$$+ 0.124(0.029)SlogP\_VSA6$$
$$+ 14.392(1.596)vsurf\_CP$$
$$+ 1.059(0.377)dipoleX + 0.287$$

$$n = 24, \quad r = 0.920, \quad Q^2_{(LOO)} = 0.781, \quad Q^2_{(L3O)} = 0.752,$$
$$s = 0.348, \quad F = 26.321$$

$$FIT = 2.632, \quad LOF = 0.216, \quad AIC = 0.185, \quad R^2_T = 0.840,$$
$$r_{yrand}(SD) = 0.395(0.116)$$

In all regression equations, n is number of compounds, r is correlation coefficient, Q^2 is cross-validated R^2 from

leave-one-out (LOO) procedure, Q^2_{L3O} is cross-validated R^2 from leave-three-out procedure (where a group of three compounds are randomly kept outside the analysis each time in such a way that all compounds are in the predictive groups for once), s is standard error of the estimate and F is F-ratio between the variances of calculated and observed activities. The $r_{yrand}(SD)$ is the mean correlation coefficient of the regressions in the activity (Y) randomization study with its standard deviation from 100 simulations. In the randomization study, none of the identified models has shown any chance correlation. Additional statistical parameters such as, the "Akaike's information criterion" (AIC) (Akaike, 1973, 1974), the Kubinyi function 'FIT' (Kubinyi, 1994a, 1994b) and the Friedman's "lack of fit" (LOF) (Friedman, 1990) have also been calculated to further validate the derived QSAR models. The AIC takes into account the statistical goodness of fit and the number of parameters that have to be estimated to achieve that 'degree of fit'. The Kubinyi function 'FIT' closely related to the F-value, proved to be a useful parameter for assessing the quality of the models. The model that produces the lowest AIC value and highest FIT value is considered potentially the most useful. The LOF factor takes into account the number of terms used in the equation and is not biased, as are other indicators, toward large number of parameters (Sharma et al., 2009). The Eqs. (1) and (2) have been validated with a test set of 8 compounds ($R^2_T = R^2$ of test set compounds). The test set predictions are in agreement with their experimental values (Table 1). The metric r^2_m values and slopes (k and k') of developed models (Eqs. (1)–(3)) have been found within their acceptable limits and thus, ensure the good predictability of the models (Table 2).

Now in terms of physical meaning descriptors, a_nN and vsurf_CP (Pearson correlation $r = -0.691$ and 0.687, respectively) are among the highest correlating descriptors with the activity (Table 3). The a_nN refers to the number of nitrogen atoms in ADCA analogs. Negative correlation of a_nN with activity indicated that presence of nitrogen atom(s) in aryl rings is detrimental for the activity. Closure look into the structures of ADCA analogs showed that phenyl/thiophenyl group containing analogs (compounds 1 and 15, Table 1) exhibited better activity than nitrogen containing pyridyl (2, 3 and 4), pyrimidyl (11) and thiazolyl (14) analogs. The descriptors a_nCl (number of chlorine atoms) and a_nF (number of fluorine atoms) show positive correlation with the activity. Here, chloro (20, 30) or fluoro (19) substituted pyridyl derivatives appeared superior to the unsubstituted one (2, 3, 4). This may be suggested that the increase in activity of chloro or fluoro substituted pyridyl or pyrimidyl derivatives may be due their positive effect on the hydrophobicity of these analogs. The descriptor std_dim2 (Eq. (1)) and vsurf_CP (Eq. (2)) belongs to the descriptor class representing surface area,

Table 2 Metric r^2_m values and slopes of QSAR models.

Equations	Metric value[a]				Slope[b]	
	Metric r^2_m	Reverse r^2_m	Average r^2_m	Δr^2_m	k	k'
(1)	0.693	0.602	0.647	0.090	0.997	0.999
(2)	0.742	0.644	0.693	0.098	0.997	0.999
(3)	0.748	0.649	0.698	0.099	1.000	0.996

[a] For an acceptable QSAR model, the value of "Average r^2_m" should be >0.5 and Δr^2_m should be <0.2.
[b] Slope should be close to 1.

Table 3 Pearson correlation matrix for the descriptors correspond to the Eqs. (1)–(3).

Des. no.	Descriptors[a]	1	2	3	4	5	6	7	8	9
1	a_nCl	1.000								
2	a_nF	−0.260	1.000							
3	a_nN	0.236	−0.055	1.000						
4	dipoleX	0.291	0.084	0.235	1.000					
5	SlogP_VSA2	−0.269	1.000	−0.055	0.084	1.000				
6	SlogP_VSA6	0.946	−0.232	0.141	0.161	−0.232	1.000			
7	std_dim2	−0.069	−0.007	0.105	0.123	−0.007	0.105	1.000		
8	vsurf_CP	−0.038	−0.277	−0.896	−0.292	−0.277	0.041	−0.003	1.000	
9	logP$_{(o/w)}$	0.349	0.005	−0.554	−0.053	0.005	0.391	0.257	0.633	1.000
	−logEC$_{50}$[b]	0.294	0.150	−0.691	0.114	0.150	0.367	0.249	0.687	0.857

[a] a_nCl: number of chlorine atoms; a_nF: number of fluorine atoms; a_nN: number of nitrogen atoms; dipoleX: dipole moment; SlogP_VSA2 and SlogP_VSA6: accessible van der Waals surface area associated with SlogP; std_dim2 and vsurf_CP: surface area, volume and shape descriptor; logP$_{(o/w)}$: hydrophobicity.

[b] −logIC$_{50}$: Free fatty acid receptor 1 agonist activity.
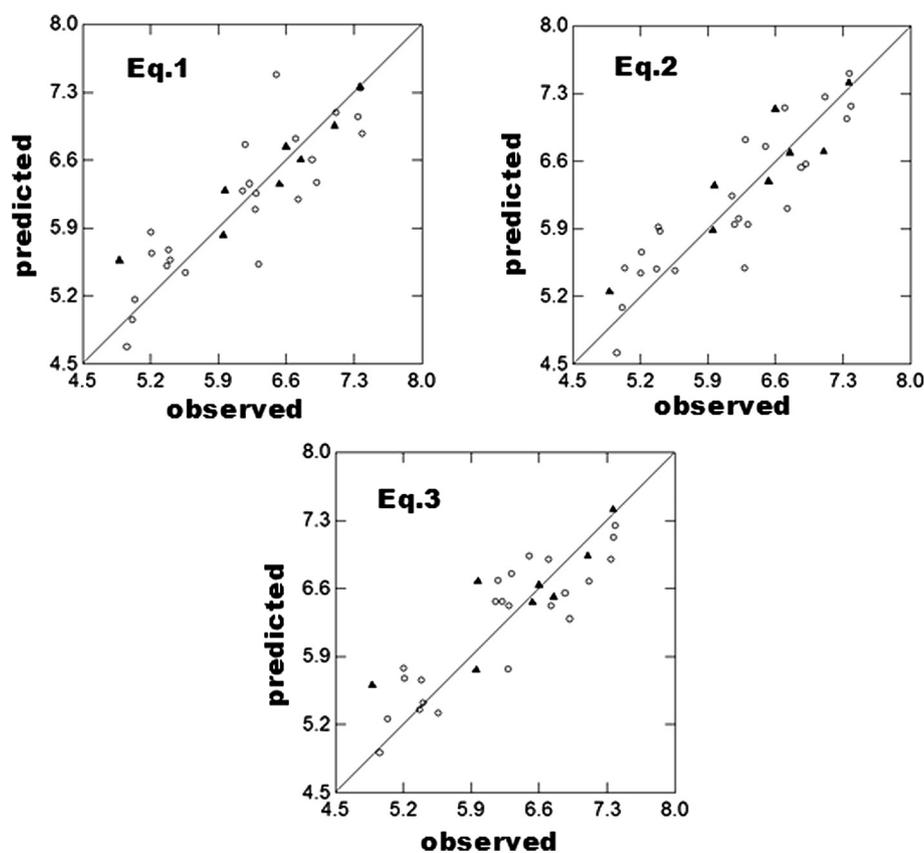
Figure 2 Plots for observed versus predicted activity of training (circle) and test (triangle) set of ADCA analogs.

volume and shape of the compounds. These descriptors are favorable for activity. The std_dim2 depend on the structure connectivity and conformation of compounds. The vsurf_CP is a critical packing descriptor which considers surface and shape of molecules. The higher surface area of molecules is associated with either hydrophobicity or important for dispersion interactions. Its positive correlation with activity indicates that the descriptor is conducive for the activity. SlogP_VSA2 and SlogP_VSA6 (Eq. (2)) are indicative of accessible van der Waals surface area with lag 2 and 6 respectively of the compounds associated with SlogP. Their positive

correlations with the activity indicate that the compounds with high accessible van der Waals surface area would have higher activity. The chloro function in compound 20 has higher accessible van der Waals surface area than fluoro containing compound 19 and exhibited better activity. However, compounds 26 (fluoro) and 27 (chloro) showed comparable activity (Table 1).

The series of ADCA analogs were designed by Christiansen et al. to study the effect of hydrophobicity. They assumed that compounds with less hydrophobicity or high hydrophilicity would be favorable for activity. To study the influence of

hydrophobicity on the activity of ADCA analogs, the Eq. (3) was developed by replacing parameter std_dim2 in Eq. (1) with $logP_{(o/w)}$.

$$-logEC_{50} = -0.657(0.169)a\_nN + 0.352(0.139)a\_nCl \\ + 0.476(0.213)a\_nF + 0.464(0.120)logP_{(o/w)} + 4.780$$

(3)

$$n = 24, \quad r = 0.929, \quad Q^2_{(LOO)} = 0.785, \quad Q^2_{(L3O)} = 0.783,$$
$$s = 0.328, \quad F = 30.241$$

$$FIT = 3.024, \quad LOF = 0.191, \quad AIC = 0.164, \quad R^2_T = 0.758,$$
$$r_{yrand}(SD) = 0.399(0.118)$$

The $logP_{(o/w)}$ was found as the highest correlating descriptor with activity ($r = 0.857$, Table 3). Its positive correlation with activity of these analogs clearly indicates that hydrophobicity is favorable for activity and thus the ADCA analogs with hydrophobic aryl functions may exhibit superior FFAR1 agonistic activity. In all equations, the test set predictions were found to be in agreement with their observed activities (Fig. 2).

4. Concluding remarks

The QSAR study on ADCA analogs as FFAR1 agonist indicated that hydrophobicity is an important physicochemical property for their activity. Hydrophobicity of organic molecules depends upon their solvent accessible surface area and thus, substituted aryl functions (phenyl or thiophenyl) with high solvent accessible surface area would lead to better activity. In aliphatic substituents, linear one possesses more van der Waals surface area than corresponding branched substituents. Therefore, aryl function in ADCA analogs may be substituted with open chain alkyl groups to increase their hydrophobicity and activity. In addition to this, aryl groups substituted with chloro or fluoro functions at *ortho*, *meta* or *para* positions would show superior FFAR1 agonistic activity. Presence of nitrogen atom decreases the hydrophobicity and thus nitrogen containing heterocyclic ring should be avoided for better FFAR1 agonistic activity.

References

Akaike, H., 1973. Information Theory and an Extension of the Minimum Likelihood Principle. AkademiaiKiado, Budapest, pp. 267–281.

Akaike, H., 1974. A new look at the statistical identification model. IEEE Trans. Autom. Control. AC-19, 716–723.

Bharate, S.B., Nemmani, K.V., Vishwakarma, R.A., 2009. Progress in the discovery and development of small-molecule modulators of G-protein-coupled receptor 40 (GPR40/FFA1/FFAR1): an emerging target for type 2 diabetes. Expert Opin. Ther. Pat. 19, 237–264.

Briscoe, C.P., Tadayyon, M., Andrews, J.L., Benson, W.G., Chambers, J.K., Eilert, M.M., Ellis, C., Elshourbagy, N.A., Goetz, A.S., Minnick, D.T., Murdock, P.R., Sauls Jr., H.R., Shabon, U., Spinage, L.D., Strum, J.C., Szekeres, P.G., Tan, K.B., Way, J.M., Ignar, D.M., Wilson, S., Muir, A.I., 2003. The orphan G protein-coupled receptor GPR40 is activated by medium and long chain fatty acids. J. Biol. Chem. 278, 11303–11311.

Christiansen, E., Urban, C., Grundmann, M., Due-Hansen, M.E., Hagesaether, E., Schmidt, J., Pardo, L., Ullrich, S., Kostenis, E., Kassack, M., Ulven, T., 2011. Identification of a potent and selective free fatty acid receptor 1 (FFA1/GPR40) agonist with favorable physicochemical and in vitro ADME properties. J. Med. Chem. 54, 6691–6703.

Del Guerra, S., Marselli, L., Lupi, R., Boggi, U., Mosca, F., Benzi, L., Del Prato, S., Marchetti, P., 2005. Effects of prolonged in vitro exposure to sulphonylureas on the function and survival of human islets. J. Diabetes Complications 19, 60–64.

Feng, X.T., Leng, J., Xie, Z., Li, S.L., Zhao, W., Tang, Q.L., 2012. GPR40: a therapeutic target for mediating insulin secretion. Int. J. Mol. Med. 30, 1261–1266.

Friedman, J., 1990. Technical Report No. 102 Laboratory for Computational Statistics. Stanford University, Stanford.

Golbraikh, A., Tropsha, A., 2002. Beware of q2! J. Mol. Graph. Model. 4, 269–276.

Grant, R.W., Wexler, D.J., Watson, A.J., Lester, W.T., Cagliero, E., Campbell, E.G., Nathan, D.M., 2007. How doctors choose medications to treat type 2 diabetes: a national survey of specialists and academic generalists. Diabetes Care 30, 1448–1453.

Graveling, A.J., Frier, B.M., 2009. Hypoglycaemia: an overview. Prim Care Diabetes 3, 131–139.

Gupta, M.K., Prabhakar, Y.S., 2006. Topological descriptors in modeling the antimalarial activity of 4-(3′,5′-disubstituted anilino)quinolines. J. Chem. Inf. Model. 46, 93–102.

Gupta, M.K., Prabhakar, Y.S., 2008. QSAR study on tetrahydroquinoline analogues as plasmodium protein farnesyltransferase inhibitors: a comparison of rationales of malarial and mammalian enzyme inhibitory activities for selectivity. Eur. J. Med. Chem. 43, 2751–2767.

Gupta, M.K., Sagar, R., Shaw, A.K., Prabhakar, Y.S., 2005. CP-MLR directed QSAR studies on the antimycobacterial activity of functionalised alkenols – topological descriptors in modeling the activity. Bioorg. Med. Chem. 13, 343–351.

Hansch, C., Hoekman, D., Leo, A., Weininger, D., Selassie, C.D., 2002. Chem-bioinformatics: comparative QSAR at the interface between chemistry and biology. Chem. Rev. 102, 783–812, <http://www.who.int/mediacentre/factsheets/fs312/en/index.html> (accessed on 18.03.13).

Inzucchi, S.E., Sherwin, R.S., 2005. The prevention of type 2 diabetes mellitus. Endocrinol. Metab. Clin. North Am. 34, 199–219.

Itoh, Y., Kawamata, Y., Harada, M., Kobayashi, M., Fujii, R., Fukusumi, S., Ogi, K., Hosoya, M., Tanaka, Y., Uejima, H., Tanaka, H., Maruyama, M., Satoh, R., Okubo, S., Kizawa, H., Komatsu, H., Matsumura, F., Noguchi, Y., Shinohara, T., Hinuma, S., Fujisawa, Y., Fujino, M., 2003. Free fatty acids regulate insulin secretion from pancreatic beta cells through GPR40. Nature 422, 173–176.

Kahn, S.E., Haffner, S.M., Heise, M.A., Herman, W.H., Holman, R.R., Jones, N.P., Kravitz, B.G., Lachin, J.M., O'Neill, M.C., Zinman, B., Viberti, G., ADOPT Study Group, 2006. Glycemic durability of rosiglitazone, metformin, or glyburide monotherapy. N. Engl. J. Med. 355, 2427–2443.

Kotarsky, K., Nilsson, N.E., Flodgren, E., Owman, C., Olde, B., 2003. A human cell surface receptor activated by free fatty acids and thiazolidinedione drugs. Biochem. Biophys. Res. Commun. 301, 406–410.

Kubinyi, H., 1994a. Variable selection in QSAR studies. I. An evolutionary algorithm. Quant. Struct. Act. Relat. 13, 285–294.

Kubinyi, H., 1994b. Variable selection in QSAR studies. II. A highly efficient combination of systematic search and evolution. Quant. Struct. Act. Relat. 13, 393–401.

Kumar, V., Gupta, M.K., Singh, G., Prabhakar, Y.S., 2012. CP-MLR/PLS directed QSAR study on the glutaminylcyclase inhibitory activity of imidazoles: rationales to advance the understanding of activity profile. J. Enzyme Inhib. Med. Chem. (in press).

Maedler, K., Carr, R.D., Bosco, D., Zuellig, R.A., Berney, T., Donath, M.Y., 2005. Sulfonylurea induced beta-cell apoptosis in cultured human islets. J. Clin. Endocrinol. Metab. 90, 501–506.

MOE, 2011. The Molecular Operating Environment from Chemical Computing Group Inc., 1255 University Street, Suite 1600, Montreal, Quebec, Canada H3B 3X3. < www.chemcomp.com >.

Ojha, P.K., Mitra, I., Das, R., Roy, K., 2011. Further exploring r_m^2 metrics for validation of QSPR models. Chemom. Intell. Lab. Syst. 107, 194–205.

Prabhakar, Y.S., 2003. A combinatorial approach to the variable selection in multiple linear regression analysis of Selwood et al. data set – a case study. QSAR Comb. Sci. 22, 583–595.

Prabhakar, Y.S., Rawal, R.K., Gupta, M.K., Katti, S.B., 2004. CP-MLR/PLS directed structure–activity modeling of the HIV-1RT inhibitory activity of 2,3-diaryl-1,3-thiazolidin-4-ones. QSAR Comb. Sci. 23, 234–244.

Prabhakar, Y.S., Gupta, M.K., Roy, N., Venkateswarlu, Y., 2006. A high dimensional QSAR study on the aldose reductase inhibitory activity of some flavones: topological descriptors in modeling the activity. J. Chem. Inf. Model. 46, 86–92.

Rayasam, G.V., Tulasi, V.K., Davis, J.A., Bansal, V.S., 2007. Fatty acid receptors as new therapeutic targets for diabetes. Expert Opin. Ther. Targets 11, 661–671.

Rizos, C.V., Elisaf, M.S., Mikhailidis, D.P., Liberopoulos, E.N., 2009. How safe is the use of thiazolidinediones in clinical practice? Expert Opin. Drug Saf. 8, 15–32.

Roy, K., Mitra, I., Kar, S., Ojha, P.K., Das, R.N., Kabir, H., 2012. Comparative studies on some metrics for external validation of QSPR models. J. Chem. Inf. Model. 52, 396–408.

Saquib, M., Gupta, M.K., Sagar, R., Prabhakar, Y.S., Shaw, A.K., Kumar, R., Maulik, P.R., Gaikwad, A.N., Sinha, S., Srivastava, A.K., Chaturvedi, V., Srivastava, R., Srivastava, B.S., 2007. C-3 alkyl/arylalkyl-2,3-dideoxy hex-2-enopyranosides as antitubercular agents: synthesis, biological evaluation, and QSAR study. J. Med. Chem. 50, 2942–2950.

Sharma, B.K., Pilania, P., Singh, P., 2009. Modeling of cyclooxygenase-2 and 5-lipooxygenase inhibitory activity of apoptosis-inducing agents potentially useful in prostate cancer chemotherapy: derivatives of diarylpyrazole. J. Enzyme Inhib. Med. Chem. 24, 607–615.

Sharma, B.K., Singh, P., Pilania, P., Sarbhai, K., Prabhakar, Y.S., 2011. CP-MLR/PLS directed QSAR study on apical sodium-codependent bile acid transporter inhibition activity of benzothiepines. Mol. Divers 15, 135–147.

Stein, S.A., Lamos, E.M., Davis, S.N., 2013. A review of the efficacy and safety of oral antidiabetic drugs. Expert Opin. Drug Saf. 12, 153–175.

Tan, C.P., Feng, Y., Zhou, Y.P., Eiermann, G.J., Petrov, A., Zhou, C., Lin, S., Salituro, G., Meinke, P., Mosley, R., Akiyama, T.E., Einstein, M., Kumar, S., Berger, J.P., Mills, S.G., Thornberry, N.A., Yang, L., Howard, A.D., 2008. Selective small-molecule agonists of G protein-coupled receptor 40 promote glucose-dependent insulin secretion and reduce blood glucose in mice. Diabetes 57, 2211–2219.

Waugh, N., Cummins, E., Royle, P., Clar, C., Marien, M., Richter, B., Philip, S., 2010. Newer agents for blood glucose control in type 2 diabetes: systematic review and economic evaluation. Health Technol. Assess. 36, 1–248.

WHO, 2014. < http://www.who.int/mediacentre/factsheets/fs312/en/ > (accessed on 12.12.14).