http://www.e-journal.net

ISSN: 0973-4945; CODEN ECJHAO
E-Journal of Chemistry
Vol. 1, No. 5, pp 243-250, October 2004

QSAR Analysis of 5-substituted-2-Benzoyl-aminobenzoic acids as PPAR Modulator

R HEMALATHA, L K SONI, A K GUPTA and S G KASKHEDIKAR*

Molecular Modelling Study Group,
CADD Laboratory, Dept. of Pharmacy,
S.G.S.I.T.S., Indore-452003, INDIA

Received 1 August 2004; Accepted 10 September 2004

Abstract: A quantitative structure activity relationship (QSAR) study on a series of analogs of 5-aryl thiazolidine-2, 4-diones with activity on PPAR-α and PPAR-γ was made using combination of various thermodynamic, electronic and spatial descriptors. Several statistical regression expressions were obtained using multiple linear regression analysis. The best QSAR model was further validated by leave one out cross validation method. The studied revealed that for dual PPAR-α/γ activity dipole-dipole energy and PMI-Z play significant role and contributed positively for PPAR-γ and PPAR-α activity respectively. Thus, QSAR brings important structural insight to aid the design of dual PPAR-α/γ receptor agonist.

Keywords: QSAR analysis, Benzoylaminobenzoic acids, Peroxisome proliferator activated receptor (PPAR-α/γ) agonist

Introduction

Type 2 diabetes is a metabolic disorder that afflicts 120 million worldwide at present and is estimated to rise to over 200 million by the year 2010[1]. In addition to the characteristic combination of insulin resistance and insulin deficiency, the type 2 diabetic often displays cardiovascular risk factors including dyslipidemia (hypertriglyceridemia, low HDL, and small dense LDL), hypertension, and obesity. The recent publication of the United Kingdom Prospective Diabetes Study (UKPDS) [2] has revealed that in Type 2 diabetes, intensive glucoselowering therapy is ineffective at reducing cardiovascular complications, despite decreasing microvascular complications such as retinopathy.

The PPARs (peroxisome proliferator activated receptors) were cloned less than a decade ago and are members of the superfamily of nuclear transcription factors that includes the receptors for steroid, retinoid, and thyroid hormones[3,4]. The PPARs form heterodimers with another nuclear receptor, the 9-*cis*-retinoic acid receptor (RXR). This heterodimer complex interacts with critical DNA response elements within promoter regions, and when activated by agonist ligand binding, it leads to gene transcription of proteins involved in control of lipid and carbohydrate metabolism. PPAR-γ agonists (*e.g.*, rosiglitazone and troglitazone) have displayed clinical utility for increasing insulin sensitivity and improving glycemic control in Type 2 diabetes. In addition, these compounds have been shown to inhibit atherosclerosis in the mice[5-7].

Studies have shown that thiazolidine-2,4-diones give highly potent in vivo antidiabetic activities[8]. Thiazolidine-2,4-diones (TZDs) are generally selective for PPARγ, although a TZD, KRP-297, with activity at PPAR-α and PPAR-γ was recently disclosed[9]. The structurally related isoxazolidinedione JTT-501 has also been recently reported with PPARγ activity similar to troglitazone accompanied by weak activity at PPAR-α[10]. PPAR-α agonists (*e.g*., gemfibrozil, fenofibric acid) produce reductions in serum triglycerides and increases in HDL cholesterol, in some cases accompanied by reductions in serum fibrinogen[11,12]. The combined profile of a dual PPAR-α/γ agonist thus appears well suited for treatment of hyperglycemia together with prevention of cardiovascular disease in Type 2 diabetes[13]. Herein we describe the quantitative structure activity relationship (QSAR) study of novel 5-subsituted 2-benzoylaminobenzoic acids as PPARα/γ modulators. This study has resulted in the identification of molecular properties, which significantly correlated with of PPAR-γ and PPAR-α agonist activity.

Experimental

Analogs of 5-substituted 2-benzoylaminobenzoic acids (table 1), as PPAR α/γ modulators were taken from the reported work of Thor *et al*[14] (excluding compounds with biological activities numerically not well defined). The biological activity data K_i values (ligand binding affinity in μM) were converted to negative logarithmic dose in mole (pK_i) for QSAR analysis. The series was subjected to molecular modeling and 3D-QSAR studies using CS Chem-Office Software version 6.0 (Cambridge soft)[15] running on a P-III processor. Structures of all the compounds (table 1) were sketched using builder module of the program. These structure were then subjected to energy minimization using molecular mechanics (MM2) until the root mean square (RMS) gradient value became smaller than 0.1kcal/mol. Å. Minimized molecules were subjected to re-optimization via Austin model-1 (AM1) method until the root mean square (RMS) gradient attained a value smaller than 0.01 kcal/mol. Å using MOPAC. The geometry optimization of the lowest energy structure was carried out using Eigenvector following (EF) routine. The descriptor values for all the molecules were calculated using "compute properties" module of program.

Calculated thermodynamic descriptors included critical temperature (T_c), ideal gas thermal capacity (C_p), critical pressure (P_c), boiling point (BP), Henry's law constant (H), bend energy (E_b), stretch bend energy (SBE), heat of formation (H_f), total energy (TE) and logarithm of partition coefficient (logP).

Steric descriptors derived were connolly accessible area (CAA), connolly molecular area (CMA), connolly solvent excluded volume (CSEV), exact mass (EM), molecular weight (MW), principal moment of inertia-X component (PMI-X), principal moment of inertia-Y component (PMI-Y) and principal moment of inertia-Z component (PMI-Z), molar refractivity (MR) and ovality (OVAL).

Electronic descriptors such as X-component of dipole moment (DPL_1), Y-component of dipole moment (DPL_2), Z-component of dipole moment (DPL_3), dipole moment (DPL_4), dipole-dipole energy (DDE) electronic energy (ElcE), highest occupied molecular orbital energy (HOMO), lowest unoccupied molecular orbital energy (LUMO), repulsion energy (NRE), VDW-1,4- energy (E14), Non-1, 4-VDW energy (E_v) and total energy were calculated.

Multiple linear regression analysis method was used to perform QSAR analysis employing in-house VALSTAT[16] program. The best model was selected on the basis of various statistical parameters such as correlation coefficient (r), standard error of estimation (SE), sequential Fischer test (F). The model was further validated on various statistical parameters like leave one out cross validated square correlation coefficient (Q^2) using cross validation method[17], boot- strapping square correlation coefficient (r^2_{bs}), randomize

biological activity data test (chance) and test for outliers (Z-score value) which confirm the robustness and applicability of QSAR equation on the structural analogs.

Results and Discussion

When data set was subjected to multiple linear regression analysis, in order to ascertain QSAR between binding affinity at PPAR-γ receptor as dependent variables and physiochemical descriptor as independent variable, several multivariant equations were obtained. The statistically significant equation with coefficient of correlation (r) =0.899 was considered as model for PPAR-γ binding affinity (table 2 and figure 1), The model showed overall internal statistical significance level better than 99.9% as it exceeded the tabulated $F_{(3,12\ \alpha\ 0.001)} = 12.7$. The inter-correlation within the parameters is significantly low (less than 0.435) suggested the non-dependency of the parameters on each other.

$$pKi = 0.187(\pm 0.040)*DDE - 0.313(\pm 0.061)*DPL_1 - 0.114(\pm 0.026)*E14 + 7.487$$
$$n=16,\ r=0.899,\ r^2=0.808,\ SE=0.222,\ F=16.850 \qquad \text{(eqn.1)}$$

The model was further subjected for leave one out (LOO) cross validation method (table 2 & figure 2), the value of $Q^2 \geq 0.3$ in cross validation method corresponds to a confidence limit greater than 95%, which minimized the risk of finding significant explanatory equation for the biological activity just by mere opportunity. The value of cross-validated squared correlation co-efficient (Q^2=0.632), predictive residual sum of square (S_{PRESS}=0.308) and standard error of predictivity (S_{DEP}=0.267) suggested good predictive ability of the biological activity with low S_{DEP}. The r^2_{bs}=0.807 is at par with the conventional squared correlation coefficient (r^2), indicating that no single compound much more/less contributed to the model. Randomize biological activity data test (Chance < 0.001) revealed that the result was not based on chance correlation (table 3). The model was further tested for outlier by Z-score method no compound was found to be outlier suggested that the model is able to explain the structural diversify analogs, which is helpful in designing of more potent compounds using physiochemical parameters.

The study revealed that for PPAR-γ binding affinity, dipole-dipole energy of the compound contributed positively which is responsible for the interaction of molecular polar portion with receptor. Substitution of group that is favorable for DDE may be enhances binding affinity of molecule with receptor. While vander Waals -1,4- energy contributed negatively to the activity, explains the depth of the attraction potential and how easy it is to push atoms together to the pocket of PPAR-γ receptor at helix-3. Dipole moment at X-component contributed negatively to the PPAR-γ receptor binding affinity.

The correlation was also established between PPAR-γ transactivation potency (EC_{50}) of fifteen compounds of the series with physiochemical properties to explore the parameters, which play significant role in the agonist activity. Various correlated equations were obtained which account for more than 75% variance in activity, the equation 2 was considered as model for transactivation potency.

$$pEC_{50} = 0.017(\pm 0.003)MW + 0.120(\pm 0.032)E_v - 8.145(\pm 2.389)OVAL + 12.924$$
$$n=15,\ r=0.866,\ r^2=0.750,\ SE=0.167,\ F=11.020 \qquad \text{(eqn.2)}$$

Model having coefficient of correlation (r) =0.866 with low error of estimation (0.167) (table 2 & figure 3). The model showed overall internal statistical significance level better than 99.5% as it exceeded the tabulated $F_{(3,11\ \alpha\ 0.005)} = 9.17$. The value of cross-validated squared correlation coefficient (Q^2=0.542), predictive residual sum of square (S_{PRESS}=0.226) and standard error of predictivity (S_{DEP}=0.193) suggested good predictive ability of the potency with low S_{DEP} (table 2

& figure 4). The r^2_{bs}=0.775 is at par with the conventional squared correlation coefficient (r^2), randomize biological activity data test (Chance < 0.001) and absent of outlier suggested robustness of the model. Equation 2 suggested that Non-1, 4-VDW energy and molecular weight contributed positively while ovality contributed negatively to the EC_{50} potency data.

For PPAR-α only eight compounds having numerically well defined binding affinity are subjected to regression analysis in order to explore the physiochemical properties responsible for the activity. Eqn. 2 was considered as statistical significant model for PPAR-α, which account for more than 73.8% of the variance in the binding affinity (table 4 & figure 5).

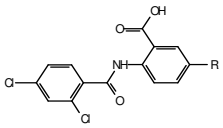
$$pKi = 3.573e\text{-}005(\pm 8.682e\text{-}006)* PMI\text{-}Z + 4.281 \quad \text{(eqn.3)}$$

$$n=8,\ r=0.859,\ r^2=0.738,\ SE=0.098,\ F=16.935$$

For PPAR-α binding affinity model having correlation coefficient value ($r \geq 0.859$) and significantly low standard error of estimation (SE = 0.098). The data showed better statistical significance >98% with $F_{(1,6)}$ = 16.935 against the tabulated value for sequential Fischer test at 98% ($F_{1,6\ \alpha\ 0.02}$= 13.7). The model was further subjected for leave one out cross validation method, the value of (Q^2=0.243), (S_{PRESS}=0.166) and (S_{DEP}=0.144) suggested moderate predictive ability of the biological activity (table 4 & figure 6). Randomize biological activity data test is less than 0.001. The model also shows that no compound is outlier. Equation 3 revealed that for PPAR-α binding affinity, principal moment of inertia of Z component contributed positively suggested that substitution of bulkier group is favorable for the activity.

The models suggested that for dual PPAR-α/γ binding affinity dipole-dipole energy and PMI-Z are essential parameter, which is contributed positively for PPAR-γ and PPAR-α respectively.

Table 1 Structure, in-vitro binding affinity in μM (K_i) and transactivation potency in μM (EC_{50}) of 5-substituted-2- benzoylaminobenzoic acid analogs



C. No.	Substitution (R)	PPAR-γ		PPAR-α
		K_i (μM)	EC_{50} (μM)	K_i (μM)
1	Methyl	2.20	1.6	30.0
2	2-Thinylmethoxy	0.40	0.3	25.0
3	2-(3-Thinyl)ethoxy	0.60	0.8	20.0
4	2-(2-Thinyl)ethoxy	1.00	0.7	[↑]NWD
5	$-OCH_2CH_2SCH_3$	1.60	0.6	19.0
6	4-Ethoxybenzyloxy	3.60	0.7	NWD
7	2-(2-Pyridinyl)ethoxy	1.50	3.2	NWD
8	3-Thienyl	0.90	0.7	28.0
9	2-Furyl	0.50	1.0	22.0
10	3-Ethoxyphenyl	1.20	0.3	NWD
11	8-Quinolinyl	0.70	0.3	NWD
12	3-Carboxyphenyl	1.40	NWD	67.0
13	5-Nitro-2-pyridinyloxy	0.18	0.4	NWD
14	3-Nitro-2-pyridinyloxy	0.17	0.4	NWD
15	2-Pyrimidinyloxy	0.10	1.3	NWD
16	6-Chloro-2-pyrazinyloxy	0.16	0.45	21.0

[↑]*Data are not numerically well defined.*

Table 2. PPAR-γ agonist activity of 5-substituted-2- benzoylaminobenzoic acid analogs using eqn 1 and 2 respectively

C. No	pKi						EC50					
	†Obs	‡Cal	C_{res}	††Z	↑Pred	‡‡P_{res}	†Obs	‡Cal	C_{res}	††Z	↑Pred	‡‡P_{res}
1	5.66	5.90	0.24	-1.20	5.98	0.33	5.80	5.73	-0.07	0.47	5.62	-0.18
2	6.40	6.45	0.05	-0.25	6.46	0.06	6.52	6.19	-0.33	2.23	6.17	-0.35
3	6.22	6.21	-0.01	0.06	6.21	-0.01	6.10	6.03	-0.06	0.42	6.02	-0.07
4	6.00	6.04	0.04	-0.21	6.05	0.05	6.15	6.24	0.09	-0.60	6.25	0.10
5	5.80	5.89	0.10	-0.49	5.90	0.11	6.22	6.27	0.05	-0.34	6.28	0.06
6	5.44	5.52	0.07	-0.37	5.59	0.14	6.15	6.04	-0.11	0.78	5.94	-0.22
7	5.82	5.98	0.15	-0.77	6.00	0.17	5.49	5.79	0.30	-2.02	5.96	0.46
8	6.05	5.92	-0.13	0.66	5.87	-0.18	6.15	6.35	0.19	-1.31	6.40	0.25
9	6.30	6.01	-0.29	1.44	5.92	-0.38	6.00	6.02	0.02	-0.14	6.03	0.03
10	5.92	5.55	-0.37	1.88	5.36	-0.56	6.52	6.47	-0.06	0.37	6.43	-0.09
11	6.15	6.39	0.23	-1.16	6.47	0.31	6.52	6.41	-0.12	0.80	6.39	-0.14
12	5.85	6.21	0.35	-1.78	6.26	0.40	-	-	-	-	-	-
13	6.74	6.66	-0.08	0.41	6.58	-0.16	6.40	6.53	0.14	-0.92	6.57	0.17
14	6.77	6.54	-0.23	1.18	6.37	-0.40	6.40	6.40	0.01	-0.04	6.41	0.01
15	7.00	6.88	-0.12	0.59	6.82	-0.18	5.89	5.84	-0.05	0.32	5.82	-0.07
16	6.80	6.79	-0.01	0.04	6.79	-0.01	6.35	6.35	0.00	-0.01	6.35	0.00

†Negative logarithm of observed activity in mole taken from reference 14. Residual of observed and calculated activity data. ††Z-score data for screening of outliers. ↑Leave one out predicted activity data. ‡‡ Residual of observed and leave one out predicted activity data.

Table 3. QSAR statistics of significant equations

Equations No.	r	SE	F	r^2_{bs}	*Chance*	Q^2	S_{PRESS}	S_{DEP}
1	0.899	0.222	16.850	0.808	0.001	0.632	0.308	0.267
2	0.866	0.167	11.020	0.776	0.001	0.542	0.226	0.193
3	0.859	0.098	16.945	0.648	0.010	0.243	0.166	0.144

Table 4. PPAR-α agonist activity of 5-substituted-2-benzoylamino benzoic acid analogs using eqn. 3

C. No.	pK_i					
	†Obs	‡Cal	C_{res}	††Z	↑Pred	‡‡P_{res}
1	4.52	4.45	-0.08	0.85	4.42	-0.10
2	4.60	4.69	0.09	-1.01	4.72	0.12
3	4.70	4.75	0.05	-0.57	4.77	0.07
5	4.72	4.60	-0.12	1.35	4.58	-0.14
8	4.55	4.59	0.03	-0.36	4.59	0.04
9	4.66	4.55	-0.11	1.20	4.53	-0.12
12	4.17	4.29	0.11	-1.23	4.49	0.32
16	4.68	4.70	0.02	-0.22	4.70	0.03

†Negative logarithm of observed activity in mole taken from reference 14. ‡ Calculated activity data. Residual of observed and calculated activity data. †† Z-score data for screening of outliers. ↑ Leave one out predicted activity data. ‡‡ Residual of observed and leave one out predicted activity data.

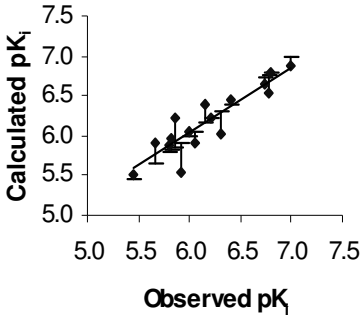


Figure 1. Plot between observed pK_i and calculated pK_i using eqn 1 for PPAR-γ

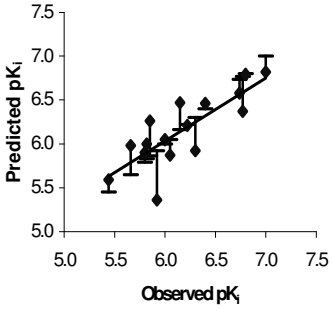


Figure 2. Plot between observed pK_i and predicted pK_i using eqn 1 for PPAR-γ

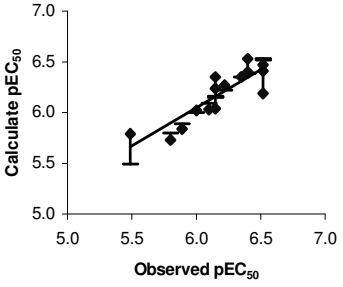


Figure 3. Plot between observed pEC_{50} and calculate pEC_{50} using eqn. 2 for PPAR-γ

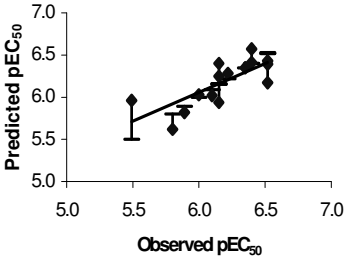


Figure 4. Plot between observed pEC_{50} and predicted pEC_{50} using eqn. 2 for PPAR-γ

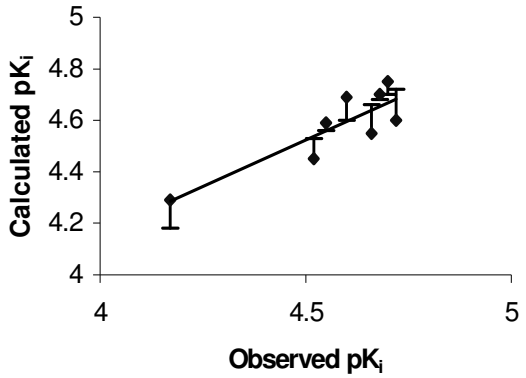


Figure 5. Plot between observed pK_i and calculated pK_i using eqn 3 for PPAR-α

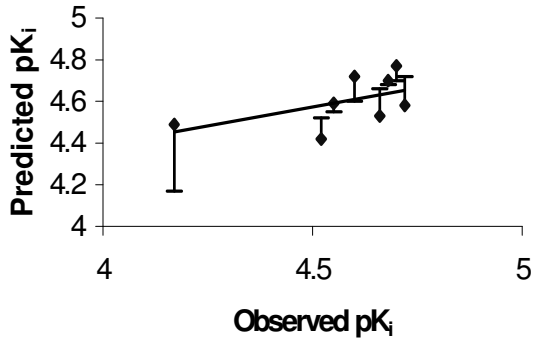


Figure 6. Plot between observed pK_i and predicted pK_i using eqn 3 for PPAR-α

Acknowledgement

Authors are thankful to the Director, Shri .G. S. Institute of Technology and Science, Indore for providing facilities for this work. One of the authors AKG is thankful to CSIR, New Delhi, for providing Senior Research Fellowship respectively.

Reference

1. Amos A F, McCarty D J and Zimmet P, *Diabetic Med.,* 1997, **14** (Suppl 5), S1-S85.
2. UK Prospective Diabetes Study Group, (UKPDS 33). *Lancet***,** 1999**, 352**, 837.
3. (a) Issemann I and Green S, *Nature,* 1990, **347**, 645. (b) Mangelsdorf D J and Evans R M, *Cell*, 1995, **83**, 841.
4. Willson T M, Brown P J, Sterbach D D and Henke B R, *J. Med. Chem.,* 2000, **43**, 527.
5. Li A C, Brown K K, Silvestre M J, Willson T M, Palinski W and Glass C K, *J. Clin. Invest.,* 2000, **106,** 523.
6. Chen Z, Ishibashi S, Perrey S, Osuga J I, Gotoda T, Kitamine T, Tamura Y, Okazaki H, Yahagi N, Iizuka Y, Shionoiri F, Ohashi K Harada K, Shimano H, Nagai R and Yamada N, *Arterioscler., Thromb., Vasc. Biol.,* 2001, **21**, 372.
7. Collins A R, Meehan W P, Kintscher U, Jackson C Wakino S Noh G, Palinski W, Hsueh W A and Law R E, *Arterioscler., Thromb., Vasc. Biol.,* 2001, **21**, 365.

8. (a) Sohda T, Mizuno K, Momose Y, Ikeda H, Fujita T and Meguro K, *J. Med. Chem.,* 1992, **35**, 2617 (b) Hulin B, Clark D A, Goldstein S W, McDermott R E, Dambek P J, Kappeler W H, Lamphere C H, Lewis D M and Rizzi J P, *J. Med. Chem.***,** 1992, **35**, 1853.
9. Nomura M, Kinoshita S, Satoh H, Maeda T, Murakami K, Tsunoda M, Miyachi H and Awano K, *Bioorg. Med. Chem. Lett.*, 1999, **9**, 533.
10. (a) Shibata T, Matsui K, Nagao K, Shinkai H, Yonemori F and Wakitani K, *Eur. J. Pharmacol.***,** 1999, **364**, 211. (b) Shinkai H, *Drugs Future,* 1999, **24** (8), 893. (c) Shinkai H, Onogi S, Tanaka M, Shibata T, Iwao M, Wakitani K and Uchida I, *J. Med. Chem.***,** 1998, **41**, 1927.
11. Fruchart J C, Duriez P and Staels B, *Curr. Opin. Lipidol.*, 1999, **10**, 245.
12. Kocky M, Gerrois P, Poulain P, Derudas B, Peters J M and Gonzalez F J et al. *Blood,* 1999, **93**, 2991.
13. Auwerx J, *Metab. Dis. Diabetic Complications*, 1999, 167.
14. Thor M, Beierlein K, Dykes G, Gustavsson A, Heidrich J, Jendeberg L, Lindqvist B, Pegurier C, Roussel P, Slater S, Backman M S, Thornstrom U and Uppenberg J, *Bioorg. Med. Chem. Lett.***,** 2002, **12**, 3565.
15. C S Chem Office, Version 6.0, Cambridge Soft Corporation, Software Publishers Association, 1730 M Street, NW, Suite 700, Washington D.C. 20036(202) 452-1600, USA.
16. Gupta A K, Arockia B M and Kaskhedikar S G, VALSTAT *I.J.P.S***.**, 2004, **66**,396.
17. Kubinyi H, In QSAR**:** Hansch Analysis & Related Approaches, Weinheim, New York, 1993, 91.