Asian Journal of Chemistry Vol. 20, No. 6 (2008), 4369-4378

Rationalization of Physico-chemical Properties of 1,3,5-Trisubstituted Aryls as Highly Selective PPAR$_\delta$ Agonist

RAJARAM HEMALATHA*, PALANIVELU MANOJ KUMAR,
SURESH CHANDRA MAHAJAN and SATHISH GOPALRAO KASKHEDIKAR†
Molecular Modeling Study Group, Mahakal Institute of Pharmaceutical Studies
Behind Airstrip, Near Karcha, Dewas Road, Ujjain-456 664, India
E-mail: rhlathaster@gmail.com

To optimize the physiochemical properties of 1,3,5-trisubstituted aryls with high selective agonist activity on PPAR$_\delta$ a quantitative structural activity relationship. Hansch approach was made using combination of various thermodynamic, electronic and spatial descriptors. Several regression expressions are obtained using multiple linear regression analysis. The best QSAR is further validated by leave-one-out cross validation method. The present studies reveal that for selective PPAR$_\delta$ agonist activity, modification at R_4 and R_5 substituted positions in molecule is more favourable and also electronic parameters play a key role in activity.

Key Words: 1,3,5-Trisubstituted aryls, PPAR$_\delta$ agonist.

INTRODUCTION

The peroxisome-proliferator activated receptors (PPARs) are lipid activated transcription factors, belonging to the nuclear hormone receptor superfamily. Three different isoforms PPAR$_\alpha$, PPAR$_\gamma$, PPAR$_\delta$ of PPARs, which differs by their target tissue and physiological functions[1,2]. The hypolipidemic fibrates[3] and the insulin sensitizing thiazolidinediones[4-6] are believed to be acting through activation of the PPAR$_\alpha$ and PPAR$_\gamma$ subtypes, respectively. Obesity is a growing threat to global health by virtue of its association with insulin resistance, glucose intolerance, hypertension and dyslipidemia, collectively known as the metabolic syndrome[7,8]. PPAR$_\delta$ has emerged as a powerful metabolic regulation in diverse tissues including fat, skeletal muscle and the heart[9-11]. Its transcriptional program enhances fatty acid catabolism and energy uncoupling, resulting in decreased triglyceride stores, improved endurance performance and enhanced cardiac contractility, respectively. These suggest that high affinity PPAR$_\delta$ synthetic drug

†Molecular Modeling Laboratory, Department of Pharmacy, Shri.G.S. Institute of Technology and Science, 23, Park Road, Indore-452 003, India.

may uniquely target multiple components of the metabolic syndrome including obesity, insulin resistance, hyperglycemia, dyslipidemia and atherosclerosis.

The aim of the present work is to study the QSAR of selective $PPAR_{\delta}$ agonists and therefore to identify associated molecular properties and also optimize their agonist activity. QSAR studies have predictive ability and simultaneously provide deeper insight into the mechanism of drug-receptor interactions[12] even before their synthesis. Thus, it may be helpful in designing new potent molecules.

EXPERIMENTAL

Analogs of 1,3,5-trisubstituted aryls (Fig. 1) as highly selective $PPAR_{\delta}$ agonists were taken from literature[13]. The biological activity data (EC_{50} in μM) were converted to negative logarithmic dose, thus correlating the data linear to free energy change and reducing the skewness of the data set (Table-1). The QSAR models were generated with training set of 19 molecules and the predictive ability of resulting model was evaluated with test of 5 molecules, which were selected randomly. The correlations were sought between $PPAR_{\delta}$ agonist activity as dependent variable and various physicochemical (hydrophobic, electronic and steric) parameters and structural indicator parameters as independent variable. The structural indicator variable I_U, I_V, I_1, I_2, I_3, I_X, I_Y and I_Z expresses 1 for presence of bond at position U, sulfur atom at position V, methyl group at R_1 position, hydrogen atom at R_2 position, chlorine atom at position R_3, nitrogen atom at position X, Y, Z and 0 for its absence. The values of substituents constants like hydrophobic (π_4 and π_5), steric (molar refractivity or MR_4 and MR_5) hydrogen acceptor (HA_4 and HA_5), hydrogen donor (HD_4 and HD_5) and electronic (field effect or F_4 and F_5, resonance effect or R_4 and R_5 and Hammett's constant or σp_4 and σp_5) for R_4 and R_5 substituted position taken into account from the literature, reported by Hansch *et al.*[14]. The series was further subjected to molecular modeling and 3D-QSAR studies using CS Chem-Office software version 6.0 (Cambridge soft) running on a P-IV processor[15]. Structures of all the compounds were sketched using builder module of the program. Then, the structure was subjected to energy minimization using molecular mechanics (MM2) until the root mean square (RMS) gradient value becomes smaller than 0.1 kcal/mol Å. Energy minimized molecule was subjected to re-optimization via Austin model-1 (AM1)[16] method until the root mean square (RMS) gradient attains a value smaller than 0.0001 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 the program. The various structural and physiochemical descriptors considered for 3D QSAR studies are given in Table-2.

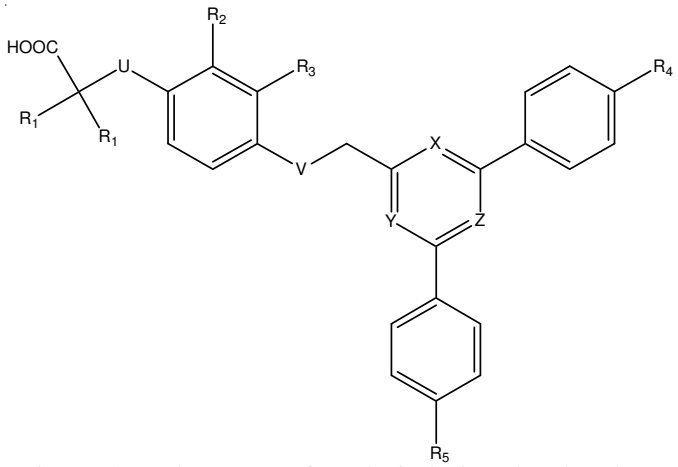

Fig. 1. General structure of 1,3,5-trisubstituted aryl analogs

TABLE-1
TRANSACTIVATION DATA FOR
1,3,5-TRISUBSTITUTED ARYL ANALOGS

Compd. No.	U	V	R_1	R_2	R_3	X	Y	Z	R_4	R_5	EC_{50} (μM)
1	O	O	H	Me	H	CH	CH	CH	CF_3	CF_3	0.01
2	O	O	H	H	H	CH	CH	CH	H	H	0.97
3	O	O	H	H	H	CH	CH	CH	CF_3	CF_3	0.11
4	O	O	Me	H	H	CH	CH	CH	CF_3	CF_3	0.05
5	O	O	H	Me	H	CH	CH	CH	H	H	0.56
6	O	O	H	H	Me	CH	CH	CH	CF_3	CF_3	0.07
7	O	S	H	H	H	CH	CH	CH	CF_3	CF_3	0.03
8	Bond	O	H	H	Cl	CH	CH	CH	CF_3	CF_3	0.54
9	Bond	S	H	H	Cl	CH	CH	CH	CF_3	CF_3	0.16
10	O	O	H	H	H	CH	CH	CH	OMe	OMe	0.35
11	O	O	H	H	H	CH	CH	CH	NMe_2	NMe_2	0.49
12	O	O	H	H	H	CH	CH	CH	Cl	Cl	0.11
13	O	O	H	H	H	CH	CH	CH	OCF_3	OCF_3	0.03
14	O	O	H	H	H	CH	CH	CH	CF_3	*m*-CF_3	0.56
15	O	O	H	H	H	CH	CH	CH	CF_3	OCF_3	0.06
16	O	O	H	H	H	CH	CH	CH	CF_3	OMe	0.07
17	O	O	H	H	H	CH	CH	CH	CF_3	Me	0.07
18	O	O	H	H	H	CH	CH	CH	CF_3	Ph	0.06
19	O	O	H	H	H	N	N	CH	OMe	OMe	0.11
20	O	O	H	H	H	N	N	CH	OCF_3	OCF_3	0.09
21	O	O	H	H	H	N	N	CH	CF_3	CF_3	0.04
22	O	O	H	H	H	CH	N	N	OMe	OMe	0.21
23	O	O	H	H	H	CH	N	N	OCF_3	OCF_3	0.04
24	O	O	H	H	H	CH	N	N	CF_3	CF_3	0.03

TABLE-2
DESCRIPTORS USED IN PRESENT QSAR STUDY

S. No.	Descriptors	Type	Descriptions
1	BP	Thermodynamic	Boiling point
2	CP	Thermodynamic	Critical pressure
3	CT	Thermodynamic	Critical temperature
4	HF	Thermodynamic	Heat of formation
5	HLC	Thermodynamic	Henry's law constant
6	IGTC	Thermodynamic	Ideal gas thermal capacity
7	log P	Thermodynamic	Logarithmic partition coefficient
8	MP	Thermodynamic	Melting point
9	MR	Thermodynamic	Molar refractivity
10	SGP	Thermodynamic	Standard Gibb's free energy
11	VDW-1,4	Thermodynamic	Van der Waals force
12	PARTCOFF	Thermodynamic	Partition coefficient for water/octanol
13	N-1,4-VDW	Thermodynamic	Non 1,4-Van der Waals force
14	STERG	Thermodynamic	Stretch energy
15	STBERG	Thermodynamic	Stretch bend energy
16	TORERG	Thermodynamic	Torsion energy
17	TOTERG	Thermodynamic	Total energy
18	CAA	Steric	Connolly accessible surface area
19	CMA	Steric	Connolly molecular surface area
20	CSEV	Steric	Connolly solvent-excluded volume
21	EM	Steric	Exact mass
22	MW	Steric	Molecular weight
23	OVAL	Steric	Ovality
24	PMI-X	Steric	Principal moments of inertia - X axis
25	PMI-Y	Steric	Principal moments of inertia - Y axis
26	PMI-Z	Steric	Principal moments of inertia - Z axis
27	D_1	Electronic	Dipole moment - X axis
28	D_2	Electronic	Dipole moment - Y axis
29	D_3	Electronic	Dipole moment - Z axis
30	D_4	Electronic	Resultant Dipole moment
31	EERG	Electronic	Electronic energy
32	HOMO	Electronic	Energy of highest occupied molecular orbital
33	LUMO	Electronic	Energy of lowest unoccupied molecular orbital
34	REPLERG	Electronic	Repulsion energy
35	BENDERG	Electronic	Bending energy
36	DDERG	Electronic	Dipole-dipole energy

Sequential multiple regression analysis method was used to perform QSAR analysis employing in-house VALSTAT[17] program. The ± data within the parentheses are associated with t-value at 95 % confidence interval of

coefficient of the descriptors in regression equation. The equations were selected on the basis of various statistical parameter (Table-3) such as correlation coefficient (r), standard error of estimate (s), sequential Fisher test (F). The robustness and applicability of QSAR equation as best model, on the structural analogs was further confirmed, using various QSAR validation technique like leave-one-out (LOO) validated square correlation coefficient (r^2_{cv}) using cross validation method[18,19], boot strapping square correlation coefficient (r^2_{bs}) randomize biological activity data (chance) and test for outliers (z-score value). Use of more than one variable in the multivariate equation was justified by autocorrelation study.

TABLE-3
SUMMARY OF MULTIPLE LINEAR REGRESSION (MLR) ANALYSIS WITH VALIDATION

Eqn. No.	n	r	q^2	Std.	F	r^2_{bs}	S_{PRESS}	S_{DEP}	r^2_{pred}	ICWP
1	17	0.92	0.76	0.19	23.57	0.87	0.233	0.203	0.69	< 0.66
2	17	0.91	0.72	0.20	20.25	0.79	0.250	0.218	0.63	< 0.77
3	19	0.85	0.54	0.25	12.70	0.78	0.319	0.284	0.57	< 0.32
4	19	0.85	0.50	0.30	12.68	0.75	0.406	0.360	0.54	< 0.15

n = number of compounds, r = correlation coefficient, q^2 = cross-validated correlation coefficient, Std. = Standard error of estimate, F = Variance ratio at specified degree of freedom (df), r^2_{bs} = bootstrapping r^2, S_{PRESS} and S_{DEP} = predicted residual sum of squares and standard deviation error of prediction, r^2_{pred} = predictive r^2.

Predictive r^2 was based only molecules not included in the training set and is defined as r^2_{pred} = (SD-PRESS)/SD, where SD is the sum of the squared deviation between biological activity of the molecules in the test set and mean biological activity of the training set molecules and PRESS is the sum of the squared deviation between the predictive and actual activity values for every molecules in the test set. Like r^2_{cv} (q^2) the predictive r^2 can assume a negative value reflecting a complete lack of predictive ability of training set for the molecules, which were included in the test set[20,21].

RESULTS AND DISCUSSION

In the present study, an attempt has been made to find structural requirement for selective $PPAR_\delta$ agonist activity using QSAR Hanch approach on trisubstituted aryls. Among several models generated eqns. 1 and 2 were selected for 2D QSAR discussion after removal of few compounds as outlier. The reason for outlier may be absence of phenoxy group and presence of double bond in compound **8** and *meta* substitution in the tail phenyl ring of compound **14**. Both the models explain for more than 82 % of the variance

in binding affinity. But eqn. 1 having good internal (r = 0.92 and q^2 = 0.76) and external predictivity (r^2_{pred} = 0.69) was selected as best model. The model showed overall internal statistical significance level better than 99 % as it exceeded the tabulated $F_{(3,13,\alpha 0.01)}$ = 5.74. The intercorrelation within the parameter (ICWP) is significantly low (less than 0.67) suggested the non-dependence of the parameters on each other. The model was subjected for leave-one-out cross validation method (Table-4 and Fig. 2), the value of q^2 = 0.3 in cross validation method corresponds to a confident limit greater than 95 %, which minimized the risk of finding significant explanatory for the biological activity just by mere opportunity. The predictive residual sum of square ($S_{PRESS} \geq 0.233$) and standard error of predictivity (S_{DEP} = 0.204) suggested good predictive ability of the biological activity of diversified structure. The boot strapping r^2 (r^2_{bs} = 0.87) is at par with conventional squared correlation coefficient indicating that no single compound much more/less contributed to the model.

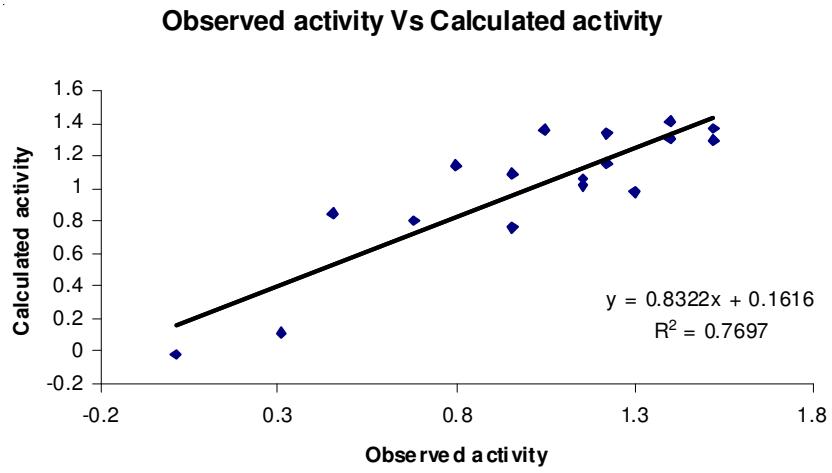

Fig. 2. A plot of observed *vs.* calculated pEC_{50} values of $PPAR_{\delta}$ activity using eqn. 1 of 2D-QSAR model

$$pEC_{50} = [0.36 (\pm 0.48)] - I_2[0.34 (\pm 0.29)] + F_4[2.12 (\pm 1.49) + \sigma_{p5} [0.45(\pm 0.45)]$$
$$n = 17, r = 0.92, S = 0.19, F = 23.57, ICWP < 0.66 \quad (1)$$

$$pEC_{50} = [0.44 (\pm 0.68)] - I_2[0.43 (\pm 0.35)] + F_4[1.68 (\pm 2.32) + \sigma_{p4} [0.46 (\pm 0.65)]$$
$$n = 17, r = 0.91, S = 0.20, F = 20.25, ICWP < 0.77 \quad (2)$$

The best model shows that the electronic effect (Field effect-F_4 and Hammet-σ_{p5} constant) contributed positively and indicator variable (I_2) contributed negatively. The field effect depends upon the intrinsic tendency

TABLE-4
OBSERVED AND PREDICTED VALUES FOR
1,3,5-TRISUBSTITUTED ARYL DERIVATIVES

S. No.	Compd. No.	Pred.[a]	Obs.[b]	Compd. No.	Pred.[c]	Obs.[d]
1	13	1.289	1.5229	24	1.255	1.5229
2	6	1.022	1.1549	20	1.587	1.0458
3	21	1.404	1.3979	14	0.764	0.2518
4	19	0.762	0.9586	2	1.125	0.9586
5	24	1.364	1.5229	23	1.349	1.3979
6	1	-0.013	0.0132	16	1.038	1.1549
7	10	0.839	0.4559	7	1.212	1.5229
8	23	1.306	1.3979	12	0.722	0.9586
9	15	1.332	1.2218	4	0.096	0.2518
10	22	0.805	0.6778	6	1.393	1.1549
11	2	1.084	0.9586	10	0.663	0.4559
12	3	0.977	1.3010	15	1.113	1.2218
13	20	1.358	1.0458	17	0.951	1.1549
14	17	1.063	1.1549	13	1.312	1.5229
15	18	1.147	1.2218	8	0.477	0.2676
16	11	0.113	0.3098	3	1.505	1.3010
17	9	1.135	0.7959	19	0.583	0.9586
18	8	Outlier	0.2676	22	0.917	0.6778
19	14	Outlier	0.2518	18	0.804	1.2218
20	7	1.403	1.5229	5	1.447	2.0000
21	16	1.042	1.1549	1	-0.042	0.0132
22	4	0.357	0.2518	9	0.711	0.7959
23	12	1.328	0.9586	21	1.036	1.3979
24	5	1.402	2.0000	11	1.126	0.3098

[a]Predicted pEC_{50} values (μM) of PPARδ for 2D QSAR model using leave one out method; [b]Observed pEC_{50} values (μM) of $PPAR_{\delta}$ for 2D QSAR model; [c]Predicted pEC_{50} values (μM)of $PPAR_{\delta}$ for 3D QSAR model using leave one out method; [d]Observed pEC_{50} values (μM) of $PPAR_{\delta}$ for 3D QSAR model; S. No. 1-19 training compounds and S.No. 20-24 test compounds for 2D and 3D QSAR.

of a substituent to release or withdraw electrons. The positive contributions of field effect (F_4) suggest that at R_4, an organic molecule or group which possesses positive field effect may increase the activity. σ is a descriptor of the substituent. The magnitude of σ gives the relative strength of the electron withdrawing or donating properties of the substituent. The positive contribution of σ-*para* constant (σ_{p5}) inferred that at R_5 it can be substituted with electron withdrawing groups, which increase the receptor activation. The negative contribution of indicator variable (I_2) reveals that in R_2 position hydrogen substitution is not favourable for the activity.

For 3D QSAR studies, eqns. 3 and 4 were selected to explain their statistical significance. Both the models explain for more than 72 % of the variance in binding affinity. But eqn. 3 having good internal (r = 0.85 and $q^2 = 0.54$) and external predictivity ($r^2_{pred} = 0.57$) was selected as best model. The model showed overall internal statistical significance level better than 99 % as it exceeded the tabulated $F_{(3,15,\alpha\,0.001)} = 9.34$. The model was further subjected to leave-one-out cross validation method (Table-4 and Fig. 3), the value of square correlation coefficient ($q^2 = 0.54$), predictive residual sum of square ($S_{PRESS} = 0.319$) and standard error of predictivity ($S_{DEP} = 0.284$) suggested good predictive ability of the biological activity. The boot strapping r^2 ($r^2_{bs} = 0.78$) is at par with conventional squared correlation coefficient. The intercorrelation within the parameter is significantly low (less than 0.32). The model also shows that no compound is outlier.

$$pEC_{50} = [-3.97 (\pm 2.66)] + CSEV [0.01 (\pm 0.01)] - D_3 [0.12 (\pm 0.07)] - D_4 [0.10 (\pm 0.07)]$$
$$n = 19,\ r = 0.85,\ S = 0.25,\ F = 12.70,\ ICWP < 0.32 \quad (3)$$

$$pEC_{50} = [-3.23 (\pm 2.33)] + CSEV [0.01 (\pm 0.01)] - D_3 [0.16 (\pm 0.09)] - D_4 [0.15 (\pm 0.09)]$$
$$n = 19,\ r = 0.85,\ S = 0.30,\ F = 12.68,\ ICWP < 0.15 \quad (4)$$

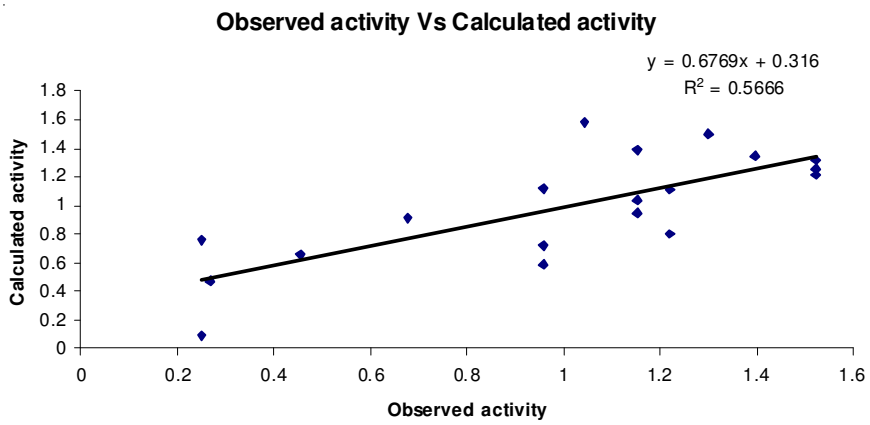

Fig. 3. A plot of observed *vs.* calculated pEC_{50} values of $PPAR_\delta$ activity using eqn.3 of 3D-QSAR model

The eqn. 3 reveals that for selective $PPAR_\delta$ agonist activity, CSEV a steric parameter contributed positively while dipole moment along the Z-axis (D_3) and resultant dipole (D_4) contributed negatively. The Connolly solvent-excluded volume (CSEV) is the volume contained within the contact

molecule surface. Positive contribution indicates that bulky substituent can contact with large volume and it may facilitate drug receptor interaction. Dipole is 3D electronic descriptor that indicates the strength and orientation behaviour of a molecule in an electro static field. Dipole properties have been correlated to long range ligand-receptor recognition and subsequent binding. The negative contribution indicates that the compounds having dipole moment in Z-axis may show less activity. The negative contribution of D_4 also suggests that the resultant dipole of overall molecule would be reduced for increasing activity. The above fact underlines the importance of electron rich functional groups and their orientation.

The study concluded that strong electronic influence (F_4, σ_{p5}, D_3 and D_4) of the substituent in core and tail phenyl ring is important for the selective $PPAR_\delta$ agonist activity. The R_2 of the core phenyl ring, R_4 and R_5 of the tail phenyl ring (supported by 2D QSAR) are more important as compared with other substituted position like U, V, R_1, R_3, X, Y and Z. If it is modified with a susbstituent so that it increases the bulkiness to facilitate drug receptor interaction and decreasing the resultant dipole of the overall molecule will be helpful in the designing of selective $PPAR_\delta$ agonist.

ACKNOWLEDGEMENTS

The authors are grateful to the Management, Mahakal Institute of Technology, Ujjain and Director Shri G.S. Institute of Technology and Science for providing necessary facilities and infrastructure to complete this work successfully.

REFERENCES

1. C.H. Lee, P. Olson and R.M. Evans, *Endocrinology*, **144**, 2201 (2003).
2. B. Desvergne and W. Wahli, *Endocr. Rev.*, **20**, 649 (1999).
3. M. Nomura, S. Kinoshita, H. Satoh, T. Maeda, K. Murakami, M. Tsunoda, H. Miyachi and K. Awano, *Bioorg. Med. Chem. Lett.*, **9**, 533 (1999).
4. L.K. Soni, P.H. Rao, S.V. Sambasivarao, A.K. Gupta, M.A. Babu and S.G. Kaskhedikar, *Indian Drugs*, **40**, 627 (2003).
5. K. Murakami, K. Tobe, T. Ide, T. Mochizuki, M. Ohashi, Y. Akanuma, Y. Yazaki and T. Kadowaki, *Diabetes*, **47**, 1841 (1998).
6. D.A. Brooks, G.J. Etgen, C.J. Rito, A. Shuker, S.J. Dominianni, A.M. Warshawasky, R. Ardecky, J.R. Paternity, J. Tyhonas, D.S. Karanewsky, R.F. Kauffman, C.L. Brodrick, B.A. Oldham, C. Monrtose-Rafizade, L.L. Winneroski, M.M. Faul and J.R. McCarthy, *J. Med. Chem.*, **44**, 2061 (2001).
7. O. Braissant, F. Fofelle, C. Scotto, M. Dauca and W. Wahli, *Endocrinology*, **137**, 354 (1996).
8. Expert Panel on Detection, *J. Am. Med. Assoc.*, **285**, 2486 (2001).
9. P. Ordentlich, M. Downes and R.M. Evans, *Curr. Top. Microbiol. Immunol.*, **254**, 101 (2001).
10. K. Japen and M.G. Rosenfeld, *J. Cell Sci.*, **115**, 689 (2002).

11. M.L. Privalsky, *Annu. Rev. Physiol.*, **66**, 315 (2004).
12. C. Hansch and A. Leo, in ed.: S.R. Heller, Exploring QSAR-Fundamentals and Application in Chemistry and Biology, American Chemical Society, Washington DC (1995).
13. R. Epple, M. Azimioara, R. Russo, B. Bursulaya, S.-S. Tian, A. Gerken and M. Iskandar, *Bioorg. Med. Chem. Lett.*, **16**, 2969 (2006).
14. C. Hansch and A. Leo, Susbstituent Constants for Correlation Analysis in Chemistry and Biology, John Willey & Sons, New York, p. 48 (1979).
15. CS Chem Office, Version 6.0, Cambridge Soft Corporation, Software Publishers Association, 1730 M Street, NW, Suite 700, Washington DC (20036).
16. L.B. Keir, Molecular Orbital Theory in Drug Research, Academic Press, New York, p. 62 (1971).
17. A.K. Gupta, M.A. Babu and S.G. Kaskhedikar, *Indian J. Pharm. Sci.*, **66**, 396 (2004).
18. K.J. Scaper, *Quant. Struct. Act. Relat.*, **18**, 354 (1999).
19. S. Wold and L. Eriksson, in ed.: H. Van De Waterbeemd, Chemometric Methods in Molecular Design, VCH, Weinheim, p. 321 (1995).
20. C.L. Waller, T.I. Oprea, A. Giolitti and G.R. Marshall, *J. Med. Chem.*, **36**, 4152 (1993).
21. R.D. Cramer III, J.D. Bunce, D.E. Patterson and I.E. Frank, *Quant. Struct.-Act. Relat.*, **7**, 18 (1998)

(*Received*: 7 May 2007; *Accepted*: 1 March 2008) AJC-6395