Indian Journal of Biotechnology
Vol 4, January 2005, pp 122-126

Quantitative structure activity relationship study of pyrazole ligands binding to estrogen receptor-α-selective agonists

Soumya Srivastava[1] and S K Srivastava[2*]

[1]Department of Chemistry, University of Allahabad, Allahabad 211 002, India

[2]Department of Chemistry, Dr H S Gour University, Sagar 470 003, India

Received 16 October 2003; revised 19 February 2004; accepted 26 February 2004

A quantitative structure activity relationship (QSAR) study on tetrasubstituted pyrazoles as high affinity ligands for the estrogen receptor (both ERα and ERβ subtypes) have been performed using various combinations of hydrophobic (MlogP), steric (MR) and electronic (Xeq) descriptors. The regression analysis of the data has shown better results in multiparametric regressions upon introduction of dummy parameters (indicator variables). The results suggest that the binding affinity for ERα and ERβ subtypes in tetrasubstituted pyrazoles is largely enhanced by the negative coefficient of MlogP and positive coefficient of MR descriptors.

Keywords: quantitative structure - activity relationship, estrogen receptors, tetrasubstituted pyrazoles

IPC code: Int. Cl.[7] A 61 K 38/00; C 07 D 231/00

Introduction

The estrogen receptor has the capacity to bind a wide variety of non-steroidal ligands with high affinity[1]. This feature has stimulated the development of non-steroidal ligands that act as selective estrogen receptor modulators (SERMs, compounds that have a mixed endocrine profile that affords agonistic or antagonistic activity in a tissue specific manner). Some of this selectivity may have been derived from differential binding to two subtypes, ERα and ERβ. Such compounds are used for menopausal hormone replacement, in fertility regulation and in the prevention and treatment of breast cancer[2].

If the non-steroidal ligand bears a reasonable structural relationship with steroidal estrogens, it is generally quite easy to imagine the orientation that this ligand is likely to adopt when it is bound by ER[3]. ER mutagenesis studies[4], and the recent X-ray crystallographic structures of ER complexed with both estradiol and three non-steroidal ligands (raloxifene, hydroxytamoxifen and diethylstilbestrol) provide additional guidance in the selection of reasonable binding orientations for ligands of this type[3]. However, when the non-steroidal estrogens have structures that are more divergent from those of steroidal estrogens, it becomes a greater challenge to predict ligand binding orientation[5].

After thorough investigation of various heterocyclic diazole structures as core elements for non-steroidal estrogens of novel design[6], the tetrasubstituted pyrazole core has emerged as a target of particular interest in developing SERMs with high affinity and selectivity for ER subtypes[7]. SERMs acted as an agonist on both ER subtypes, but it was considerably more potent on ERα than on ERβ[8]. Thus, this pyrazole was termed an ERα potency-selective agonist[8].

Materials and Methods

2D-QSAR analysis was performed on the tetrasubstituted pyrazole derivatives derived by Stauffer's group[7] using Hansch (Linear Free Energy Relationship) approach keeping physicochemical parameters [hydrophobic (MlogP), steric (MR) and electronic (Xeq)] as an independent and binding affinity values (reported in purified full length human ERα and ERβ using competitive radiometric binding assay[7] as dependent parameter. QSAR has been performed separately on ERα and ERβ subtypes.

Partition coefficient (log P)

log P, one of the most widely used hydrophobic parameters, is a free energy related parameter (LFER),

*Author for correspondence:
Tel: 91-7582-221259
E-mail: professorsks@rediffmail.com

which expresses the relative free energy change occurring on moving a substituent from one phase to another.

$\log P = \sum \pi$ (additive free energy)

M stands for Moriguchi, the scientist who proposed this method.

Molar Refractivity (MR)

It is a parameter for correlation of dispersion forces involved in the binding of haptens to antibodies. It has been correlated with lipophilicity, molar volume and steric bulk as follows:

$MR = \eta^2 - 1 / \eta^2 + 1.\ MW/d$

where η is the refractive index, d is the density and MW is the molecular weight of the compound.

Equalized Electronegativity (Xeq)

Sanderson's principle of electronegativity equalization states that "when two or more elements initially different in electronegativity combine chemically, they become adjusted to the same intermediate electronegativity within the compound". It is formulated as:

$Xeq = N / \sum v / x$

where N is the total number of atoms present in the species formula, v is the number of particular elements in the species formula and x is the electronegativity of that particular atom[9].

Indicator Parameters (Ind)

Certain features, which cannot be described by continuous variables, can very well be explained by indicator descriptor or dummy variables or *de novo* constants used in multiple regression analysis. In QSAR equations, dummy variables either a substituent or a molecular fragment normally describe a certain structural element. Free Wilson analysis may be interpreted as a regression analysis using only indicator variables.

An indicator variable is generated for each structural feature that deviates from an arbitrarily chosen compound value one, indicating the presence of a certain substituent or structural feature, and zero, indicating its absence is correlated with the biological activity values. The numerical values of the indicator parameter only serve to identify category or class membership. They show the significance of a particular group or a substituent in a given series of drug. They account for the abrupt increase or decrease of a given pharmacological activity at any specific site in the drug molecule. In the present study, the indicator parameter, Ind is taken as unity for substituents having isobutyl group and zero in all the other cases.

Regression Analysis

Recent developments in statistics provide interesting set of measures of validity that are based on simulating the predictive power of a model. Maximum R^2 improvement method, used to identify prediction models[10], finds the 'best' model for the prediction of property/activity.

Multiple regression analysis for correlating the binding affinity values of the present set of compounds with molecular descriptors were carried out using SPSS Software version for window, 9.0.0 Vev, December, 1998. Several multiple regressions were attempted using correlation matrix in this program and the best results were considered and discussed in developing QSAR and hence, for modeling the binding affinity values of the compounds in the present study.

Computations

The values of independent parameter were automatically loaded from DRAGON Software[11] and QSAR regression analyses were executed on Compaq PC using SPSS Software. The source of molfiles was from ISIS Draw Package version 2.3 developed by MDL Information Systems, Inc. All the structures were drawn in ISIS Draw and transported to the DRAGON Software wherein their parameter values were calculated.

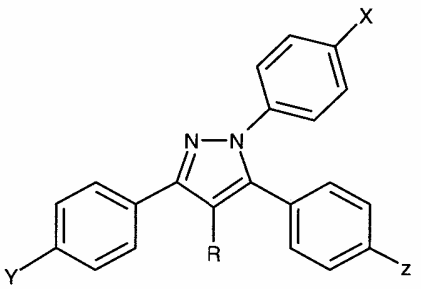
Results and Discussion

Physico-chemical data for pyrazole ligands is given in Table 1. The correlations of MlogP, MR and Xeq with RBA ERα (eqs 1-3) and RBA ERβ (eqs 4-6) gave the following simple regression (monoparametric) equations:

$$\text{Log RBA} = -0.372(\pm 0.282)\ M \log P + 0.302 \quad \ldots (1)$$
$$n = 16,\ r^2 = 0.215,\ SE = 0.512,\ F_{(1,14)} = 0.957$$

where, n is the number of data points, r is the correlation coefficient, SE is the standard error of estimation, F is the variance ratio between the calculated and

Table 1—Physico-chemical data for pyrazole ligands



Compd	Substituents				Descriptors				Observed activity		Calculated activity PreRBAERα Eq.(11)	Res-RBAERα	Observed activity		Calculated activity PreRBAERβ Eq.(13)	Res-RBAERβ
	R	X	Y	Z	MlogP	MR	Xeq	Ind	RBAERα	Log-RBAERα			RBAERβ	Log-RBAERβ		
4a	Me	H	OH	OH	2.76	0.17	26.99	0	0.76	-0.12	-0.12	-	0.28	-0.55	-0.66	0.11
4b	Et	H	OH	OH	2.61	0.15	31.32	0	31	+1.49	1.86	-0.37	1.1	+0.04	+0.17	-0.13
6a	i-Pr	H	OH	OH	3.79	0.17	26.66	0	5.6	+0.75	0.84	-0.09	0.86	-0.07	-0.22	0.15
4c	n-Pr	H	OH	OH	3.79	0.17	26.66	0	16.8	+1.23	0.84	0.46	0.52	-0.28	0.22	-0.06
4d	i-Bu	H	OH	OH	3.77	0.17	26.66	1	56	+1.75	1.21	0.54	1.4	+0.15	0.27	-0.12
4e	n-Bu	H	OH	OH	2.98	0.17	26.66	0	8.7	+0.93	0.61	0.33	0.47	-0.33	-0.28	-0.05
4f	Et	OH	OH	OH	2.91	0.23	27.99	0	36	+1.56	1.32	0.24	0.15	-0.82	-0.7	-0.12
4g	n-Pr	OH	OH	OH	2.91	0.23	27.99	0	49	+1.69	1.32	0.37	0.12	-0.92	-0.7	-0.22
4h	i-Bu	OH	H	OH	2.91	0.23	27.99	1	75	+1.88	1.62	0.26	0.89	-0.05	-0.17	0.12
4i	n-Bu	OH	OH	OH	3.08	0.26	27.99	0	14	+1.15	1.49	-0.34	0.18	-0.74	-1.07	0.33
7a	Et	OH	H	H	3.12	0.23	28.99	0	3.1	+0.49	0.63	-0.11	1.5	+0.18	-0.38	-0.2
7b	Et	H	OH	H	3.33	0.22	28.99	0	2.6	+0.41	-0.73	-0.32	0.61	-0.21	-0.05	-0.16
7c	Et	H	H	OH	3.33	0.22	29.99	0	0.04	-1.4	-0.73	-0.67	0.06	-1.22	-0.05	0.21
4b	Et	H	OH	OH	2.12	0.23	29.99	0	31	+1.49	0.86	0.63	1.1	+0.04	0.12	-0.08
8a	Et	OH	H	OH	2.33	0.22	29.32	0	7	+0.85	-0.19	1.04	0.8	-01	0.48	0.38
8b	Et	OH	OH	H	2.54	0.21	30.32	0	8.9	+0.95	1.03	-0.08	0.32	-0.49	-1.03	0.54

observed activities and data points with (±) sign in parentheses are standard errors of regression coefficients.

Log RBA = 0.661(± 0.314) MR + 0.817 … (2)
n = 16 r^2 = 0.183 SE = 0.625 $F_{(1,14)}$= 0.440

Log RBA = -0.768(± 0.427) Xeq + 0.378 … (3)
n = 16 r^2 = 0.369 SE = 0.509 $F_{(1,14)}$= 1.630

Log RBA = -0.424(± 0.317) MlogP + 0.243 … (4)
n = 16 r^2 = 0.281 SE = 0.510 $F_{(1,14)}$= 0.865

Log RBA = 0.480(± 0.382) MR + 0.776 … (5)
n = 16 r^2 = 0.136 SE = 0.622 $F_{(1,14)}$= 0.648

Log RBA = -0.245(± 0.118) Xeq + 0.239 … (6)
n = 16 r^2 = 0.189 SE = 0.712 $F_{(1,14)}$= 0.461

Equations (1 to 6) indicate that no statistical significant mono-parametric correlations are possible for modeling the activity. Thus, the binding affinity values could better be modeled by multivariate analysis.

The first step in analyzing multivariate correlation is to investigate auto-correlation. It has been done by obtaining correlation matrix, which is useful in determining the extent to which independent variables are correlated with one another as well as with the variables (relative binding affinity values in this case) i.e. to estimate autocorrelation/collinearity. It can be useful in determining redundant independent variables. Also in correlation matrix, the correlations close to ±1.0 are observed, since that indicates changes in the independent variables are linearly related to changes in the dependent variable. Such a correlation matrix obtained in the present case for ERα and ERβ subtypes is given in Tables 2 and 3.

A perusal of the correlation matrix indicates MlogP neither correlates with MR nor with Xeq for ERα subtype and for ERβ subtype MlogP & MR & MR and Xeq do not show any autocorrelation. So these parameters have been taken together and multiple regression analysis (MRA) performed showed significant improvement in correlation depicted from equations (7-10) as follows:

Log RBA = -0.810(±0.513) MlogP + 0.468(±0.211) MR + 0.326 … (7)
n = 16 r^2 = 0.710 SE = 0.443 $F_{(2,13)}$ = 6.94

Log RBA = -0.783(±0.443) MlogP - 0.660(±0.431) Xeq + 0.684 … (8)
n = 16 r^2 = 0.768 SE = 0.410 $F_{(2,13)}$ = 8.39

Log RBA = -0.830(±0.511) MlogP + 0.660(±0.511) MR + 0. 261 … (9)
n = 16 r^2 = 0.588 SE = 0.416 $F_{(2,13)}$ = 0.87

Log RBA = 0.512(±0.396) MR - 0.410(±0.245) Xeq + 0.211 → n = 16 r^2 = 0.402 SE = 0.511 $F_{(2,13)}$ = 4.662 … (10)

Equations 7, 8, 9 and 10 account for 71% (R=0.84), 77%(R=0.88), 51%(R=0.77) and 40% (R=0.63) variances, respectively and their F-value is also significant at 95% confidence interval. Now, to improve the degree of correlation, an indicator parameter (Ind) for substituents having isobutyl group was introduced in equations as follows:

Log RBA = -0.932(± 0.476) MlogP + 0.511 (±0.310) MR + 0.174 (±0.043) Ind + 0.528 … (11)
n = 16 r^2 = 0.796 SE = 0.462 $F_{(1,14)}$= 7.886

Log RBA = -0.872 (± 0.342) MlogP - 0.830 (±0.410) Xeq + 0.233 (±0.008) Ind + 0.382 … (12)
n = 16 r^2 = 0.796 SE = 0.462 $F_{(1,14)}$= 7.886

Log RBA = -0.741(± 0.450) MlogP + 0.762(±0.410) MR + 0.410 (±0.233) Ind + 0.723 … (13)
n = 16 r^2 = 0.786 SE = 0.376 $F_{(1,14)}$= 9.832

Log RBA = 0.632 (± 0.307) MR - 0.397 (± 0.197) Xeq + 0.331 (± 0.201) Ind 0.510 … (14)
n = 16 r^2 = 0.713 SE = 0.504 $F_{(1,14)}$= 7.314

From the regression analysis data of equations 11, 12, 13 and 14, it can be inferred that hydrophobic parameter MlogP and MR plays a crucial role in enhancing the RBA values of the estrogen receptor. Besides, electronic parameter, Xeq, also affects the binding

Table 2—Correlation matrices for ERα subtype

	MLog P	MR	Xeq
MLog P	1.00	0.23	0.14
MR		1.00	0.63
Xeq			1.00

Table 3—Correlation matrices for ERβ subtype

	MLog P	MR	Xeq
MLog P	1.00	0.23	0.71
MR		1.00	0.14
Xeq			1.00

affinity to some extent. The large negative coefficient of MlogP indicates that highly hydrophobic susbstituents would lower the binding affinity in this class of Ers, however, positive coefficient of MR indicates that sterically bulky susbstituents would raise the binding affinity. Since the relationship with MR is linear, this implies that the receptor has some flexibility at this site. It can very well be inferred that some component of hydrophobicity is embodied in MR term.

The negative coefficient of electronic parameter, Xeq, suggests that the electron donating groups are favorable for relative binding affinity values. The X-ray crystallographic structure of the receptor shows that there is some very hydrophobhic space above the B-ring of the ligand at position 11, presumably sufficient to accommodate groups of moderate size without larger substituents. This would require some movement of ligand or receptor for a complex to form. It can be generalized from the set of congeners studied in the tetrasubstituted pyrazole series that the negative hydrophobic and positive steric effect is operative here. However, the fact that correlation taken with indicator parameter here is small and positive, which indicates that substituents chosen as indicator parameter would increase the relative binding affinities markedly. This unique combination of parameters and substituent may be retained in future drug designing for enhancing the affinity values of tetrasubstituted pyrazole ligands.

Acknowledgement

The authors are grateful to Heads, Chemistry Department, Allahabad University, Allahabad and Dr H S Gour University, Sagar for providing the facilities to carry out the work. One of us (SS) is also thankful to the Director CDRI, Lucknow for providing the facility for training programme under Dr A K Saxena, Scientist, Medicinal Chemistry Division to learn 2D and 3D QSAR studies.

References

1 Gao H, Katzenellenbogen J A, Garg R & Hansch C, Comparative QSAR analysis of estrogen receptor ligands, *Chem Rev,* 99 (1999) 723-744.

2 Katzenellenbogen B S & Katzenellenbogen J A, Estrogen receptor transcription and transactivation: Estrogen receptor-alpha and estrogen receptor-beta regulation by selective estrogen receptor modulators and importance in breast cancer, *Breast Cancer Res,* 2 (2000) 335-344.

3 Anstead G M, Carlson K E & Katzenellenbogen J A, The estradiol pharmacophore: Ligand structure-estrogen receptor binding site, *Steroids*, 62 (1997) 268-303.

4 Ekena K, Weis K E, Katzenellenbogen J A & Katzenellenbogen B S, Different residues of the human estrogen receptor are involved in the recognition of structurally diverse estrogens and antiestrogens, *J Biol Chem,* 272 (1997) 5069-5075.

5 Anstead G M, Peterson C S & Katzenellenbogen J A, Synthesis, estrogen receptor binding affinity and binding orientation considerations, *J Steroid Biochem,* 33 (1989) 877-887.

6 Katzenellenbogen J A, Katzenellenbogen B S, Fink B S, Stauffer S R, Mortensen D S *et al*, Estrogen Receptor Ligands, Patent WO 00/19994, April 13, 2000.

7 Stauffer S R, Coletta C J, Tedesco R, Sun J, Katzenellenbogen J A *et al*, Pyrazoles ligands: Structure-affinity/activity relationships and estrogen receptor-α-selective agonists, *J Med Chem*, 43(2000) 4934-4947.

8 Sun J, Meyers M J, Fink B E, Rajendran R, Katzenellenbogen B S *et al*, Novel ligands that function as selective estrogens or antiestrogens for estrogen receptor-α or estrogen receptor-β *Endocrinology*, 140 (1999) 800-804.

9 Hansch C & Leo A, *Substitution constants for correlation analysis in chemistry and biology* (Wiley, New York) 1970.

10 Chatterjee S, *Regression analysis by examples*, 3rd ed (Wiley & Sons, New York) 2000.

11 Todeschini R & Consonni V, *Various Molecular Descriptors, A Handbook of Molecular Descriptors* (Wiley–VCH Weinheim, Germany) 2000.