Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 15 (2005) 957–961

QSAR of estrogen receptor modulators: exploring selectivity requirements for ER_α versus ER_β binding of tetrahydroisoquinoline derivatives using E-state and physicochemical parameters☆

Subhendu Mukherjee,[a] Achintya Saha[a] and Kunal Roy[b,*]

[a]*Department of Chemical Technology, University of Calcutta, 92 A P C Road, Kolkata 700 009, India*
[b]*Drug Theoretics and Cheminformatics Lab, Department of Pharmaceutical Technology, Jadavpur University, Kolkata 700 032, India*

Received 20 October 2004; revised 14 December 2004; accepted 16 December 2004
Available online 21 January 2005

Abstract—Considering importance of developing selective estrogen receptor modulators (SERMs), the present paper explores selectivity requirements of tetrahydroisoquinoline derivatives for binding with ER_α versus ER_β receptors using E-state index and physicochemical parameters. The best model [$n = 21$, $Q^2 = 0.512$, $R_a^2 = 0.613$, $R = 0.819$, $F = 11.6$ (df 3, 17)] for ER_α binding data obtained from radioligand binding assay showed importance of C_1, C_{15} and lipophilicity ($\log P$) while the best model [$n = 21$, $Q^2 = 0.768$, $R_a^2 = 0.796$, $R = 0.904$, $F = 40.1$ (df 2, 18)] for ER_β binding data showed importance of C_1 and molar refractivity (MR). While modeling ER_α/ER_β selectivity [$n = 21$, $Q^2 = 0.695$, $R_a^2 = 0.739$, $R = 0.882$, $F = 19.8$ (df 3, 17)], C_1, C_{15} and molar refractivity were found to be significant contributors. The data obtained from cellular transcription assay were also modeled. In case of ER_α, the best equation involving E-state values of C_1 and C_{14} and $\log P$ explained 62.1% of the variance while the best equation for ER_β involving E-state values of C_1 and C_{15} and MR explained 64.6% of the variance of the response variable. In case of ER_α/ER_β selectivity, the best equation involving E-state values of O_8, C_{14} and N_{27} showed 48.3% explained variance, which increased to 63.5% on deletion of single outlier. From the analysis it appears that the nitrogen atom of the aminoethoxyphenyl substituent and 6-hydroxy substituent of the tetrahydroisoquinoline nucleus play important roles for ER_α/ER_β selectivity in addition to R_1 and R_2 substituents.

Estrogenic effects have been primarily related with the female reproductive organs and are mediated principally through the estrogen receptor (ER). Estrogens also have direct effect on other tissues, for example, they are known to be present in specific cells of the skeletal and cardiovascular systems.[1,2] ER is a member of the nuclear hormone receptor superfamily.[3] Different ER-ligands induce distinct structural changes in the receptor that influence its ability to interact with other proteins critical for the regulation of target gene transcription. The ER selectivity reflects the diversity of estrogen receptor forms and regulators and the diversity of ER target genes. The principal endogenous ligand for ER in most species is 17-β estradiol. The biological effects of estrogens are known to be mediated by two receptors referred to as estrogen receptor-α (ER_α) and receptor-β (ER_β).[4] The existence of these two subtypes provide possible explanation for the tissue-selectivity.[5] The two receptors differ in size, with ER_α having 595 amino acids and ER_β having 485 amino acids. The predominant ER in the female reproductive tract and mammary glands is ER_α, whereas ER_β is the primary ER in vascular endothelial cells, bone and male prostrate tissues. The compounds that have the potential to modulate selectivity of the different estrogen target tissues are known as selective estrogen receptor modulators (SERMs).[4] An aryl substituted pyrazole derivative was reported to be the first agent that could discriminate between ER_α and ER_β subtypes.[6] This compound was found to have 120-fold higher potency to stimulate ER_α than ER_β.[6] The binding sites of ER_α and ER_β differ in two amino acids (Leu and Met in α subtype are replaced by Met and Ile respectively in ER_β).[7] The existence of two rather than one ER makes

Keywords: QSAR; SERM; E-state index; Physicochemical parameters; Tetrahydroisoquinoline derivatives; Selectivity.

☆ Presented in the Medicinal Chemistry section of 56th Indian Pharmaceutical Congress, Kolkata (India), December 3–5, 2004.

* Corresponding author. Tel.: +91 33 2414 6676; fax: +91 33 2414 6677; e-mail: kunalroy_in@yahoo.com

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.12.048

the mechanism of action of estrogens and antiestrogens (SERMs) more complex.[8] Estrogens, upon binding to its high-affinity receptor (or receptors), trigger expression of multiple genes involved in the regulation of cell proliferation and differentiation.[9] Unlike estrogens and antiestrogens, the SERMs exert selective agonist or antagonist effects on different estrogen target sites.[10,11] This unique effect of SERMs may be due to following three mechanisms: (i) differences in estrogen receptor expression in different tissues; (ii) differences in estrogen receptor conformation on ligand binding and (iii) differences in expression and binding to estrogen receptor of coregulatory proteins. ER_α is always activator, whereas ER_β inhibits the actions of estrogen. Therefore, the relative level of expression of these two types of receptors will affect the cellular responsiveness to SERMs.

A number of SERMs are currently in clinical trials and two compounds of this category, tamoxifen and raloxifene, are presently in the market for the treatment of hormone-dependent breast cancer[11,12] and prevention and treatment for osteoporosis.[13] However, both these agents have been linked to increased risks of thromboembolism and tamoxifen has been shown to increase the risk of endometrial cancer.[13–15] Hence the search for more tissue specific analogues continues, so as to develop distinct SERMs with lesser side effects. Recently tetrahydroisoquinolines[16] have been reported as potent ER_α selective ligands. The present paper explores selectively requirements of tetrahydroisoquinoline derivatives[16] for binding with ER_α versus ER_β receptors using atom level E-state index and physicochemical parameters (hydrophobicity $\log P$ and molar refractivity MR). Both radioligand binding (RLB) assay and estrogen response element (ERE) assay data were modeled in the present analysis. The biological activity values [IC_{50} (nM)] were first converted to logarithmic scale [pIC_{50} (μM)] and then used for the QSAR modeling.

Topological models directly give structural information to guide design of new molecules.[17] The electrotopological state (E-state) of atoms has been reported to be of importance in elucidating the important atoms or substructure in drug–receptor interactions.[18–23] An atom in a molecule is part of a field of information with regard to electronic influences and topological surroundings.[18,24] Quantification of influence of this field on any atom can correlate to the biological performance of a molecule. The contribution of an atom can be expressed as the electrotopological state (E-state),[25] mathematically defined as

$$S_i = I_i + \Delta I_i \quad \text{where,} \quad I = [(2/N)^2\delta^v + 1]/\delta \quad \text{and}$$
$$\Delta I_i = \sum (I_i - I_j)/r_{ij}^2$$

I is the intrinsic state of an atom, ΔI_i is the perturbation effect, N is the principal quantum number, δ is the number of sigma electrons on the atom (excluding those bonding to hydrogen), δ^v is the number of valence electrons (excluding those bonding to hydrogen), i and j are serial numbers of atoms and r_{ij} is the shortest graph distance between two atoms i and j plus one.

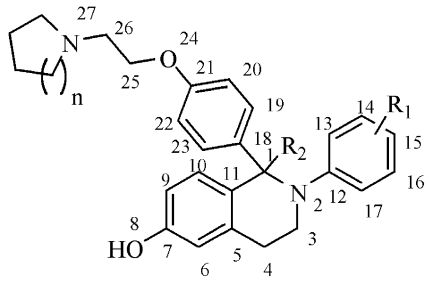

Figure 1. General structure of tetrahydroisoquinoline derivatives: common atoms are numbered 1–27.

In the present work, the atoms of the molecules were numbered keeping serial numbers of the common atoms same in all the compounds (as shown in Fig. 1). The E-state index values (S_X) were calculated using ELECTRO1 program.[26] The physicochemical parameter values ($\log P$ and MR) were calculated by *Chem Draw Ultra 5.0* software[27] using Crippen's fragmentation method.[28] All compounds considered in the present study contain 27 common atoms (excluding hydrogens). Using the program AUTOREG,[26] all possible combinations of predictor variables were tried (all-possible-subsets regression) with a restriction that predictor variables used in an equation are not much intercorrelated ($|r| < 0.5$). Using the program RRR98,[26] regression coefficients with corresponding standard errors and various statistical parameters reflecting quality[29] (like explained variance R_a^2, correlation coefficient R, standard error of estimate s, variance ratio F and average of absolute values of the residuals AVRES) of the equations were found out. Leave-one-out (LOO) cross-validation[30] was done using the programs KRPRES1 and KRPRES2,[26] which generate predicted variance (Q^2), predicted residual sum of squares (PRESS), standard deviation based on PRESS (S_{PRESS}), standard deviation of error of prediction (SDEP) and average of absolute values of predicted residuals ($Pres_{av}$).

While modeling radioligand binding assay data, the best model for ER_α binding data shows 51.2% leave-one-out predicted variance while explained variance of the equation is 61.3%.

$$[pIC_{50}]_\alpha^{RLB} = 0.586(\pm 0.369)S_1 - 0.096(\pm 0.079)S_{15} - 0.158(\pm 0.091)\log P + 2.764$$
$$n = 21, \quad R_a^2 = 0.613, \quad R^2 = 0.671,$$
$$R = 0.819, \quad F = 11.6 \text{ (df 3, 17)},$$
$$s = 0.112, \quad \text{AVRES} = 0.087,$$
$$Q^2 = 0.512, \quad \text{SDEP} = 0.122,$$
$$S_{PRESS} = 0.136, \quad Pres_{av} = 0.107 \qquad (1)$$

Eq. 1 shows importance of C_1, C_{15} and lipophilicity ($\log P$) of the molecules. The positive coefficient of S_1 and negative coefficient of S_{15} in Eq. 1 indicates that the ER_α binding affinity increases with increase of E-state value of C_1 (mainly influenced by R_2 substituent) and decrease of E-state value of C_{15} (mainly influenced by R_1 substituent). Again, the negative coefficient of