MALAYSIAN JOURNAL OF BIOCHEMISTRY & MOLECULAR BIOLOGY

The Official Publication of The Malaysian Society For Biochemistry & Molecular Biology (MSBMB)
http://mjbmb.org

SHORT COMMUNICATION

TWO-DIMENSIONAL QUANTITATIVE STRUCTURE ACTIVITY RELATIONSHIP (2D-QSAR) ANALYSIS ON ESTROGEN RECEPTOR ALPHA (ERα) AND BETA (ERβ) AGONISTS

Jaslyn Chong Chai Lin, Ling Suk Jiun*, Adelina Chia Yoke Yin*

School of Biosciences, Taylor's University Lakeside Campus, No. 1, Jalan Taylors, 47500 Subang Jaya, Selangor, Malaysia

*Corresponding Author: Tel.: +603 5629 5000; Fax: +603 5629 0001; E-mail: yokeyin.chia@taylors.edu.my

HISTORY

Received: 15th September 2017
Accepted:1th December 2017

KEYWORDS

Estrogen Receptor, Agonist, Selective estrogen receptor modulator, 2D-QSAR.

ABSTRACT

Estrogen receptor has 2 subtypes – ERα and ERβ and both are important in human health via ER-mediated actions. To correlate the structural characteristics of ER agonists and their exhibited estrogenic activities, validated 2D-QSAR models were developed for ERα (r^2=0.7311, q^2=0.6772, $r^2_{(pred)}$ = 0.7047) and ERβ (r^2 = 0.7239, q^2 = 0.6573, $r^2_{(pred)}$= 0.6133). Each model is associated with ten descriptors and showed the same pair of properties with the highest positive (*Molecular_FractionalPolarSurfaceArea*) and negative (*JX*) effect on the agonistic activity for both subtypes. Using the models, predicted EC_{50} of three potential agonists corresponding with the docking analysis by Ling et al. (2017) showed a consistent trend for ERα but otherwise for ERβ, suggesting further investigations.

INTRODUCTION

Estrogen receptor (ER) has two subtypes – alpha (ERα) and beta (ERβ) which belong to a large conserved superfamily of nuclear receptors. These receptors can be triggered by agonists, antagonists, selective estrogen receptor modulators (SERM) and thus, play important roles in human physiology via ER-mediated biological actions (Zhang et al., 2013, Niinivehmas et al., 2016). In relation to that, the estrogen 17β-estradiol (E_2) is an endogenous, agonistic, steroidal hormone actively involved in the metabolic homeostasis of the breast, cardiovascular, immune systems and more (Salum et al., 2008). Estrogen deficiency is known to cause discomfort in menopausal women including hot flushes and osteoporosis (An, 2016). Hence following that, estrogen was incorporated in the hormone replacement therapy (HRT) to target the symptoms but severe side effects were observed instead (Su Wei Poh et al., 2015). Until today, there is still growing interest to develop SERMs capable to maximize the estrogenic benefits and simultaneously avoid the side effects.

Quantitative structure-activity relationship (QSAR) is a popular ligand-based method that correlate compound structures and their biological activities via equations (Abbasi et al., 2017). This computational strategy is often acquired for the reasons of high throughput, low cost and animal welfare (Wang et al., 2016). This study uses two-dimensional QSAR (2D-QSAR) to explore the relationship between the structural properties and estrogenic activity exhibited by structurally-diverse agonists with relation to ERα and ERβ.

METHODOLOGY

This analysis was mainly performed with the BIOVIA Discovery Studio Client 4.0 (Accelrys Software Inc. 2016) using an Intel® Core™ i5-3450 CPU 3.10 GHz processor with the Windows 7 Home Premium operating system. The two data sets used were downloaded separately from the BindingDB.org - containing only human ERα/β agonists with EC_{50} data obtained *via* the luciferase reporter transactivation assay. Data duplicates, compounds without well-defined EC_{50} values or/and those with EC_{50} < 10,000 nM were excluded from the analysis. Subsequently, all biological data were converted to pEC_{50} values as the dependent variable for the experiment. The prepared set for each subtype was randomly divided using 80:20 ratio to obtain the training and test set. The independent variable here are the respective molecular descriptors calculated using the genetic function approximation (GFA) algorithm. GFA automatically selects the features most statistically-significant to the ER-agonistic activity and prevent the model from overfitting (Srivastav and Tiwari, 2017). This can

ER subtype	Descriptor	Elaboration
Alpha (ERα)	Molecular_FractionalPolarSurfaceArea *	The fraction of polar surface area over the total molecular surface area
	Kappa_1*	First Order Shape Index
	HBA_Count*	Total number of hydrogen bond accepting groups in the molecule
	Molecular_Mass	Molecular mass of molecule
	Molecular_PolarSASA	The polar solvent accessible surface area for each molecule using a 2D approximation.
	Num_AliphaticSingleBonds*	Total number of aliphatic single bonds present on the molecule
	Num_ChainAssemblies	Number of fragments remaining from all ring bonds when removed from a molecule
	PHI	Molecular flexibility
	Num_H_Acceptors*	Total number of hydrogen acceptors on the molecule
	JX*	Balaban's index based on relative electro-negativities
Beta (ERβ)	Molecular_FractionalPolarSurfaceArea*	The fraction of polar surface area over the total molecular surface area
	HBA_Count*	Total number of hydrogen bond accepting groups in the molecule
	Kappa_1*	Shape index of order one
	SC_3_C	Counts the number of clusters
	Num_RingAssemblies	Number of ring assembles or fragments remaining when all non-ring bonds
	E_DIST_equ	Total: Edge distance/equality
	Num_AliphaticSingleBonds*	Total number of aliphatic single bonds present on the molecule
	Num_H_Acceptors*	Total number of hydrogen acceptors on the molecule
	CHI_3_C	Kier & Hall's Third Order Cluster Molecular Connectivity Index
	JX*	Balaban's index based on relative electro-negativities

Figure 1 Ten molecular descriptors were calculated using the GFA algorithm for developed 2D-QSAR equation for ERα and ERβ agonistic activity respectively. Six molecular features (marked with "*") were found common in both receptors.

assist the development of products with enhanced or selective profiles for either or both ER.

The training set was used to develop 2D-QSAR models associated with the quality parameters such as the r^2, the coefficient of determination; $r^2_{(adj)}$, the r^2 adjusted for the number of terms in the model; $r^2_{(pred)}$, the prediction (PRESS) r^2 which is equivalent to q^2 from a LOO cross-validation; and *S.O.R. p-value*, the p-value for significance of regression. With reference to Arooj et al. (2012) and (Cramer et al., 1988), this best-fitted 2D-QSAR model is selected if the **(1)** dependent variables (pEC50) are five times as many as independent variables (descriptors); **(2)** $r^2_{\text{training set}} > 0.6$; **(3)** $r^2_{\text{test set}} > 0.5$; **(4)** $r^2_{(adj)} > 0.5$; **(5)** q^2 from LOO cross-validation/ $r^2_{(pred)} > 0.5$ and **(7)** the confidence interval of all individual regression coefficient significance ≥ 95%. Apart from these criteria, the developed models were justified of their predictive power by predicting the biological activity of the compounds in the external test set (Golbraikh and Tropsha, 2002).

Lastly, the 2D-QSAR model with the best correlation parameters was selected for ERα and ERβ respectively. An extra prediction step was performed to calculate the estrogenic activity (EC_{50}) of three potential agonists - Glabridin, Glycyrrhizic acid (GA) and Glycyrrhetinic acid (GE). The predicted data were compared with the average lowest binding energy obtained from a molecular docking study by Ling et al. (2017). The trend displayed by these two computational methods was used to elucidate the estrogen-like behavior of the three ligands as potential estrogen receptor agonists or SERM.

Equation 1 – 2D-QSAR model for ERα

pEC_{50} = 7.7282 + 0.0088667 (Molecular_Mass) + 0.52478 (HBA_Count) − 0.19861(Num_AliphaticSingleBonds) − 0.44722 (Num_ChainAssemblies) − 1.0045 (Num_H_Acceptors) + 23.59 (Molecular_FractionalPolarSurfaceArea) − 0.021255 (Molecular_PolarSASA) − 3.6929 (JX) + 0.55985 (Kappa_1) − 0.67434 (PHI)

$N_{training}$= 142, r^2 = 0.7311, $r^2_{(adj)=}$ 0.7106, q^2 = 0.6772, *S.O.R. p-value* = 1.13×10^{-32}

N_{test} = 36, $r^2_{(pred)}$ = 0.7047

Equation 2 – 2D-QSAR model for ERβ

pEC_{50} = 3.6956 + 0.89063 (HBA_Count) − 0.29444 (Num_AliphaticSingleBonds) − 0.89907 (Num_H_Acceptors) + 0.32919 (Num_RingAssemblies) + 9.691 (Molecular_FractionalPolarSurfaceArea) − 1.5038 (CHI_3_C) − 0.0021575 (E_DIST_equ) − 4.8345 (JX) + 0.67903 (Kappa_1) + 0.61424 (SC_3_C)

$N_{training}$ =123, r^2 = 0.7239, $r^2_{(adj)}$= 0.6993, q^2 = 0.6573, *S.O.R. p-value* = 6.86×10^{-27}

N_{test} = 31, $r^2_{(pred)}$= 0.6133

RESULTS AND DISCUSSION

Equation 1 and 2 present the 2D-QSAR models selected for ERα and ERβ. Each model is associated with ten descriptors (elaborations in **Figure 1**) and the corresponding validation statistics. In Equation 1, the descriptors Molecular_FractionalPolarSurfaceArea, Kappa_1, HBA_Count, Molecular_Mass have positive signs. This suggests that ligands with high values for these features are more likely to have a higher potency with ERα. On the other hand, the other descriptors have negative coefficients - implying a detrimental effect on the ERα potency. In Equation 2, Molecular_FractionalPolarSurfaceArea, HBA_Count, Kappa_1, SC_3_C, Num_RingAssemblies are positive factors to a higher ERβ potency whereas E_DIST_equ, Num_AliphaticSingleBonds, Num_H_Acceptors, CHI_3_C, JX are inversely correlated with the ERβ-estrogenic activity. Interestingly, six common descriptors were noted between the models for both ER subtypes (**Figure 1**). However, this result is expected as many literatures have stated the highly-conserved structure for these two receptors. Another noteworthy finding was that the descriptors Molecular_FractionalPolarSurfaceArea and JX are each the positive and negative descriptor with the largest coefficient values in both the model equations. This inferred them to be the main factors to determine the ERα- and ERβ- agonistic activity. This similarity also indicated that a large fraction of polar surface area over the total molecular surface was favourable for a higher potency against both ER subtypes whereas JX implied the opposite.

As mentioned above, the models were used to predict the estrogenic activity of GLA, GA and GE. **Table 1** presents the trend comparison among the three agents to compare their binding potency (EC_{50}) and affinity (average lowest binding energy) for ERα and ERβ. The trend is compared between the QSAR-predicted EC_{50} (nm) and docking-derived average lowest binding energy, E (kcal/mol) among three experimental ligands - Glabridin, Glycyrrhizic acid (GA) and Glycyrrhetinic acid (GE). The result for ERα is consistent whereas the result for ERβ displayed some discrepancy. For both receptors, the ligand GA was noted to be least responsive among the three ligands tested. The docking study by Ling et al. (2017) suggested that Glabridin is most likely to bind to both the ER subtypes due to the relatively low average binding energy exerted, followed by GE and lastly GA. This is possibly due to the structural similarity and lipophilicity between the phytoestrogen Glabridin and the endogenous E_2 (Simmler et al., 2013). For ERα, this trend is consistent with the increasing predicted-EC_{50} from Glabridin < GE < GA. This outcome is desirable and justified the robustness of the 2D-QSAR generated for ERα with reference to the docking study. However, the ERβ QSAR results showed discrepancy as the equation predicted GE to be most potent with the lowest EC_{50} when compared to Glabridin and GA. Through this pattern, it is observed that GA was least effective in showing agonistic properties for both ER subtypes as compared to the other two ligands. This corresponded to past literatures that reported similar findings regarding GA (Ling et al., 2017).

The 2D-QSAR here was primarily conducted over structurally-diverse ligands responsive to ER. As this virtual method depends heavily on the training set for model development, structural analogues of the compounds to be analyzed may improve the model prediction quality (Vilar and Costanzi, 2012). In addition, the two-dimensional approach here is limited to the topological properties of the data set. Thus, future works can be targeted on multi-dimensions of this relationship to extract more information regarding the structural contributions to the responses with ER receptors.

Table 1. Comparison of three agents for their binding potency (EC_{50}) and affinity (average lowest binding energy) for ERα and ERβ.

	ERα		ERβ	
Ligands	EC_{50} (nm)	E (kcal/mol)	EC_{50} (nm)	E (kcal/mol)
17β-estradiol (E_2)	N/A	-10.470 ± 0.000	N/A	-10.920 ± 0.000
Glabridin (GLA)	0.0027	-3.210 ± 0.016	0.006	-3.083 ± 0.006
Glycyrrhizic acid (GA)	1.07×10^9	332.300 ± 1.235	1.98×10^{13}	330.550 ± 6.765
Glycyrrhetinic acid (GE)	146.4941	14.570 ± 0.008	0.0008	24.147 ± 0.015

ACKNOWLEDGEMENTS

This study was supported by Taylor's University as a dissertation in partial fulfilment of the requirements for the Degree of Bachelor of Biomedical Science (Hons.)

CONFLICT OF INTEREST

The authors declare that there is no conflict of interests regarding the publication of this manuscript.

REFERENCES

Abbasi, M., Sadeghi-Aliabadi, H. & Amanlou, M. 2017. Prediction of New Hsp90 Inhibitors Based on 3,4-Isoxazolediamide Scaffold Using Qsar Study, Molecular Docking and Molecular Dynamic Simulation. *Daru,* 25**,** 17.

An, K.C. 2016. Selective Estrogen Receptor Modulators. *Asian Spine J,* 10**,** 787-91.

Arooj, M., Thangapandian, S., John, S., Hwang, S., Park, J.K. & Lee, K.W. 2012. Computational Studies of Novel Chymase Inhibitors against Cardiovascular and Allergic Diseases: Mechanism and Inhibition. *Chem Biol Drug Des,* 80**,** 862-75.

Cramer, R.D., Patterson, D.E. & Bunce, J.D. 1988. Comparative Molecular Field Analysis (Comfa). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J Am Chem Soc,* 110**,** 5959-67.

Golbraikh, A. & Tropsha, A. 2002. *Journal of Computer-Aided Molecular Design,* 16**,** 357-369.

Ling, J., L. Oh, D. & Y. Y. Chia, A. 2017. Molecular Docking Studies of Glycyrrhizic Acid (Ga), Glycyrrhetic Acid (Ge) and Glabridin (Gla) with Estrogen Receptors (Ers). *Biosciences, Biotechnology Research Asia,* 14**,** 01-11.

Niinivehmas, S.P., Manivannan, E., Rauhamaki, S., Huuskonen, J. & Pentikainen, O.T. 2016. Identification of Estrogen Receptor Alpha Ligands with Virtual Screening Techniques. *J Mol Graph Model,* 64**,** 30-39.

Salum, L.B., Polikarpov, I. & Andricopulo, A.D. 2008. Structure-Based Approach for the Study of Estrogen Receptor Binding Affinity and Subtype Selectivity. *J Chem Inf Model,* 48**,** 2243-53.

Simmler, C., Pauli, G.F. & Chen, S.N. 2013. Phytochemistry and Biological Properties of Glabridin. *Fitoterapia,* 90**,** 160-84.

Srivastav, V.K. & Tiwari, M. 2017. Qsar and Docking Studies of Coumarin Derivatives as Potent Hiv-1 Integrase Inhibitors. *Arabian Journal of Chemistry,* 10**,** S1081-S1094.

Su Wei Poh, M., Voon Chen Yong, P., Viseswaran, N. & Chia, Y.Y. 2015. Estrogenicity of Glabridin in Ishikawa Cells. *PLoS One,* 10**,** e0121382.

Vilar, S. & Costanzi, S. 2012. Predicting the Biological Activities through Qsar Analysis and Docking-Based Scoring. *Methods Mol Biol,* 914**,** 271-84.

Wang, P., Dang, L. & Zhu, B.T. 2016. Use of Computational Modeling Approaches in Studying the Binding Interactions of Compounds with Human Estrogen Receptors. *Steroids,* 105**,** 26-41.

Zhang, L., Sedykh, A., Tripathi, A., Zhu, H., Afantitis, A., Mouchlis, V.D., Melagraki, G., Rusyn, I. & Tropsha, A. 2013. Identification of Putative Estrogen Receptor-Mediated Endocrine Disrupting Chemicals Using Qsar- and Structure-Based Virtual Screening Approaches. *Toxicol Appl Pharmacol,* 272**,** 67-76.