Medicinal Chemistry Research (2019) 28:1661–1673
https://doi.org/10.1007/s00044-019-02402-9

MEDICINAL CHEMISTRY RESEARCH

ORIGINAL RESEARCH

Designing of novel ERRγ inverse agonists by molecular modeling studies of docking and 3D-QSAR on hydroxytamoxifen derivatives

Rui Li[1] · Yongli Du[1] · Jingkang Shen[2]

Received: 26 March 2019 / Accepted: 16 July 2019 / Published online: 13 August 2019

Abstract
ERRγinverse agonist is a powerful therapeutic target for the treatment of cancers and certain metabolic disorders. Until now, only GSK5814 was reported as selective ERRγinverse agonist. So 60 newly hydroxytamoxifen analogues were selected to perform molecular docking and 3D-QSAR study to design more selective inverse agonist of ERRγ. Both established CoMFA and CoMSIA models obtained high predictive and satisfactory value, demonstrated that bulky, hydrophobic and negative electrostatic substitutions are preferred at R_2 position, and introducing hydrophilic and H-bond donor and acceptor groups at R_1 and R_4 positions is greatly important for improving binding activities. The obtained information will be useful to provide clues for rationally designing novel and high potency ERRγinverse agonists.

Keywords ERRγinverse agonist · Molecular docking · 3D-QSAR · CoMFA · CoMSIA

Introduction

Nuclear receptors (NRs) are classified as transcription factors and control the development, homeostasis, and metabolism of the organism, which regulate the expression of specific genes only when ligands are present (Nam et al. 2002; Greschik et al. 2002; Kallen et al. 2004; Yu and Forman 2005). Orphan members belonging to nuclear receptor super family have no identified endogenous ligand and are greatly crucial in aspects of various cellular metabolism including mitochondrial energetics as well as cholesterol, bile acid and glucose metabolism (Yu et al. 2017). Estrogen-related receptors (ERRs) similar to estrogen receptor (ER) are special orphan subfamily of nuclear receptors and consist of three subtypes including ERRα,-β, and –γ, which are constitutively activated but not bind to endogenous small ligands (Giguére 2002; Horard 2003; Abad et al. 2008). Although ERRα and ERRβ are ubiquitously expressed, ERRγ is highly expressed in different fetal and adult human tissues, including the placenta, brain, skeletal muscle, heart, and liver (Lin et al. 2017).

ERRγ contains a N-terminal activation function (AF)-1 domain involved in the transcriptional regulation, a central zinc finger DNA-binding domain (DBD) highly conserved among the three members, a ligand-binding domain (LBD) as well as a C-terminal AF-2 domain interacting with co-activators and co-repressors (Wang et al. 2006; Kim et al. 2016). ERRγ plays a main role in endocrine signaling and metabolic pathways. It induces the angiogenic growth factors vascular endothelial growth factor (VEGF)-A and fibroblast growth factors (FGF) in muscle and liver (Narkar et al. 2011; Jung et al. 2016). Furthermore, ERRγ regulate several key enzymes involved in nutrient metabolism, as it induces the gluconeogenic enzymes phosphoenolpyruvate carboxykinase (PEPCK) and glucose-6-phosphatase (G6Pase) as well as lipid and alcohol metabolism (Kim et al. 2011; Zhang et al. 2015). These data indicate that targeting ERRγpathway is a powerful therapeutic strategy for the treatment of cancers and certain metabolic disorders including obesity and type-2 diabetes, cardiovascular disease and muscle atrophy. However, synthetic modulators of ERRγ containing agonists, antagonists or inverse agonists are very exiguous at present, only possessing four members. Among them 4-hydroxytamoxifen (4-OHT) (Coward et al.

✉ Yongli Du
ylduyjs@163.com

[1] School of Chemistry and Pharmaceutical Engineering, Qilu University of Technology (Shandong Academy of Sciences), 3501 Daxue Road, Jinan 250353, China

[2] State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 555 Zu Chong Zhi Road, Shanghai 201203, China

2001) and diethylstilbestrol (DES) (Nam et al. 2002) were identified as nonselective ERRγ inverse agonists, while the acyl hydrazine compound, GSK4716 (Zuercher et al. 2005), was reported as a selective ERRγ agonist, only GSK5814 (Kim et al. 2016) was developed as selective inverse agonist for ERRγ.

Therefore, molecular modeling studies were performed by comparative molecular field analysis (CoMFA) (Cramer et al. 1988) and comparative molecular similarity indices analysis (CoMSIA) (Klebe et al. 1994) modules as well as docking results, to investigate the three-dimensional quantitative structure–activity relationship (3D-QSAR) between hydroxytamoxifen analogues and ERRγ.

Materials and methods

Collection of data set

For better establishing QSAR model, 60 novel 4-hydroxytamoxifen analogs (Kim et al. 2016) as estrogen-related receptor γ (ERRγ) inverse agonists in possession of same skeleton and similar binding mode with ERRγ enzyme were picked up for this study, which their binding activities represented by IC_{50} values spread uniformly from 0.01 to 10 μM. Energy minimization module with tripos force field in Sybyl-X 2.0 (Cramer et al. 1989) was executed to optimize 60 compounds and docking suite was subsequently performed to generate three-dimensional conformations, one compound has twenty conformations, all conformations were saved as alignment file. Then IC_{50} values were converted into pIC_{50} values as key dependent variables to next perform QSAR analyses (Clark et al. 1989). 60 analogs were divided into a training set for model generation and a test set for model validation, consisting of 45 compounds and 15 molecules, respectively. During this process, it is vital that ensuring the pIC_{50} values of selected 15 compounds are evenly distributed from 2 to 5, since covering more than 3 log units is an important guarantee for a better QSAR model (Balasubramanian et al. 2014).The chemical structure, IC_{50} and pIC_{50} values of 60 compounds are exhibited in Table 1.

Preparation of protein

The high-resolution crystal structure of ERRγ (PDB ID: 2EWP) was obtained from the RCSB protein data bank (Kim et al. 2016). This crystal structure was prepared in Sybyl-X 2.0 using the protein preparation module. Ligand and water molecules were removed. Polar hydrogen atoms were added to investigate interactions between analogs and ERRγenzyme. Furthermore, the pocket of binding site was generated for the following molecular docking.

Molecular docking and alignment

Surflex-Dock (SFXC) module was carried out at this step used default parameters. Specially, one main parameter, maximum number of per molecular conformation, will be augmented when rational conformations are absent. Although there were plenty of clustered docking poses, only one conformation well met the requirement of low energy and low crash score as well as rational binding site (Aalizadeh et al. 2015). All rational conformations picked up from clustered docking poses were aligned based on the interactions between ligand and ERRγ enzyme binding site. After alignment in the BTK binding site, selected rational conformations were conserved as a database file which was used for 3D-QSAR study.

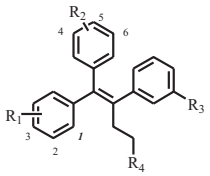
3D-QSAR analysis studies

3D-QSAR studies including CoMFA and CoMSIA analysis were performed in the form of molecular spread-sheets in SYBYL X-2.0 initiated by QSAR command, which all needful data were from the training set after molecular docking. The CoMFA (Cramer et al. 1988) and CoMSIA (Klebe et al. 1994) fields both included steric (S) and electrostatic (E) fields, the CoMSIA fields also involved hydrophobic (H), hydrogen-bond donor (D) and acceptor (A) fields. 3D-QSAR studies were calculated in standard settings with the energy cutoff values of 30 kcal/mol and attenuation factor α was defined as 0.3. After 3D-QSAR analyses, the standard contour maps of CoMFA and CoMSIA for visualizing the results were developed, which were displayed in the field type 'StDev*Coeff'.

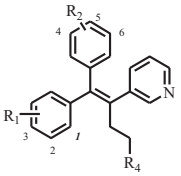
Model validation

The developed CoMFA and CoMSIA models were demanded to check the stability and robustness by external test set validations and internal validation. The external test set was applied to verify the predictive accuracy of the derived 3D-QSAR models, in which 15 analogs were not included in the model building. Internal validation aimed to inspect the predictability of the data set was carried out by means of the partial least squares (PLS) (Wold et al. 1984) approach in cross-validation method. PLS approach generated the cross-validated q^2 and optimum number of components (NOC). The final CoMFA and CoMSIA models were developed using the obtained NOC. Only the coeffcient values fall between 1.0 and 0.5 (Ron Wehrens et al. 2000), an accurate model is accepted. Moreover, non-cross-validation analysis was employed to yield the coeffcient of determination r^2 and the F-test value (F) to better evaluate the accuracy and robustness of the developed models.

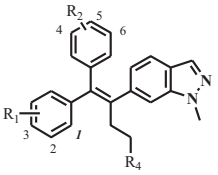
Table 1 Chemical structures of 60 compounds with their IC_{50} and pIC_{50}



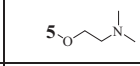
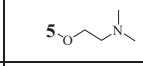
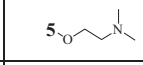
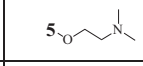
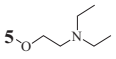
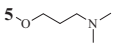
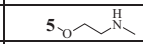
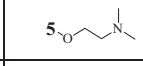
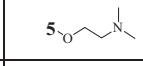
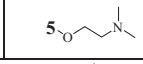
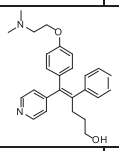
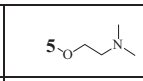
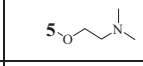
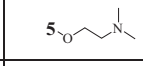
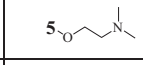
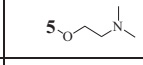
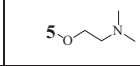
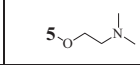
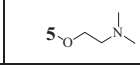
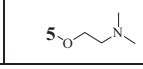
1-10,12-15,21-55,57-60

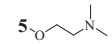
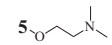
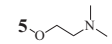
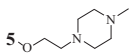
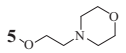
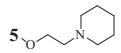
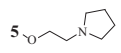
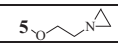
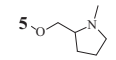
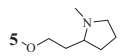
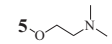
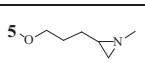
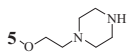
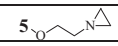
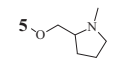
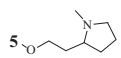
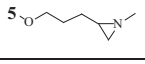
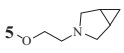
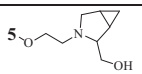
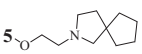
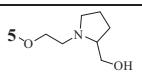
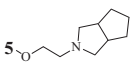
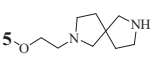
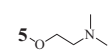


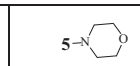
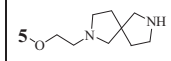
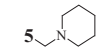
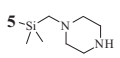
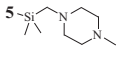
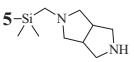
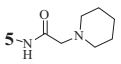
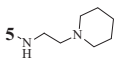
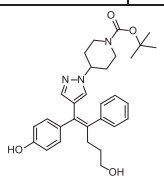
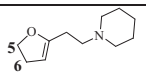
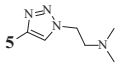
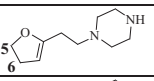
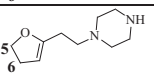
16-18



19,20

Mol.	R_1	R_2	R_3	R_4	IC_{50}(μM)	pIC_{50}
1*	**3**-F		H	CH_2OH	3.900	2.409
2*	**3**-Cl		H	CH_2OH	0.520	3.284
3	**1**-F,**3**-F		H	CH_2OH	6.700	2.174
4	**1**-F		H	CH_2OH	1.500	2.824
5	**3**-OH		H	CH_2OH	0.180	3.745
6*	**3**-OH		H	CH_2OH	0.090	4.048
7	**3**-OH		H	CH_2OH	0.190	3.721
8	**3**-OH		H	CH_2OH	0.110	3.959
9	**3**-CF_3		H	CH_2OH	1.300	2.886
10	**3**-SO_2CH_3		H	CH_2OH	10.00	2.000
11					10.00	2.000
12	**3**-CH_3		H	CH_2OH	0.270	3.569
13	**3**-OCH_3		H	CH_2OH	1.000	3.000
14	**3**-CH_3		F	CH_2OH	0.500	3.301
15	**3**-Br		F	CH_2OH	1.100	2.959
16*	**3**-Br		--	CH_2OH	10.00	2.000
17*	**3**-CH_3		--	CH_2OH	10.00	2.000
18	**3**-H		--	CH_2OH	10.00	2.000
19	**3**-OCH_3		--	CH_2OH	10.00	2.000
20	**3**-CH_3		--	CH_2OH	5.600	2.252

21	3-Br		H	CF_2H	0.890	3.051
22	3-OH		H	CF_2H	0.590	3.229
23	3-Br		H	F	0.480	3.319
24*	3-OH		H	CH_2OH	2.205	2.657
25	3-OH		H	CH_2OH	0.318	3.498
26	3-OH		H	CH_2OH	0.093	4.032
27	3-OH		H	CH_2OH	0.067	4.174
28	3-OH		H	CH_2OH	0.070	4.155
29	3-OH		H	CH_2OH	0.108	3.967
30	3-OH		H	CH_2OH	0.020	4.699
31	3-OH		F	CH_2OH	0.026	4.585
32*	3-OH		H	CH_2OH	0.784	3.106
33	3-OH		H	CH_2OH	1.044	2.981
34*	3-Br		H	CH_2OH	1.522	2.818
35*	3-Br		H	CH_2OH	0.064	4.194
36	3-Br		H	CH_2OH	0.296	3.529
37	3-Br		H	CH_2OH	1.657	2.781
38*	3-OH		H	CH_2OH	0.079	4.102
39	3-OH		H	CH_2OH	0.083	4.081
40*	3-OH		H	CH_2OH	0.050	4.301
41	3-OH		H	CH_2OH	0.091	4.041
42	3-OH		H	CH_2OH	0.165	3.783
43	3-OH		H	CH_2OH	0.326	3.487
44*	3-Br		H	CH_2OH	0.110	3.959

45	3-OH		H	CH_2OH	2.230	2.652
46	3-NO_2		H	CH_2OH	8.233	2.084
47	3-Br		H	CH_2OH	2.300	2.638
48	3-OH		H	CH_2OH	1.281	2.893
49*	3-OH		H	CH_2OH	1.256	2.901
50	3-OH		H	CH_2OH	0.460	3.337
51*	3-OH		H	CH_2OH	0.248	3.606
52	3-OH		H	CH_2OH	0.813	3.090
53	3-OH		H	CH_2OH	2.780	2.556
54	3-OH		H	CH_2OH	1.546	2.811
55*	3-OH		H	CH_2OH	2.340	2.631
56					10.00	2.000
57	3-OH		H	CH_2OH	0.268	3.572
58	3-OH		H	CH_2OH	6.295	2.201
59	3-OH		H	CH_2OH	0.894	3.049
60	3-CH_3		H	CH_2OH	2.131	2.671

*Test set. IC_{50}: half-maximal inhibitory concentration. All IC_{50} values were from Kim's research group recorded in reference 11 and 20

Results and discussion

Molecular docking

The aim of the molecular docking was to predict the binding affinity and visualize interactions between ERRγ inverse agonists and ERRγ enzyme. The accuracy of the docking program was confirmed by comparing the docked conformation of ligand (green) and existed ligand (yellow) extracted from the crystal structure of ERRγ (PDB ID: 2EWP). This result was presented in Fig. 1a, revealing the excellent agreement, which certified that the selected experimental parameters and procedures of molecular docking were reasonable. As depicted in Fig. 1a, the common scaffolds of three compounds enclosed by blue aperture were at the same position and mainly interacted with residues Glu275 and Asn346.

For explanation of binding mode in the active site, Compound 30 ($IC_{50} = 0.02\,\mu M$) was selected for more detailed analysis, as it was the most representative molecule in term of high binding activity and reasonable conformation in data set.

From Fig. 1a, III ring of compound **30** interacted with −C−O of Glu275 by double hydrogen bonding action.

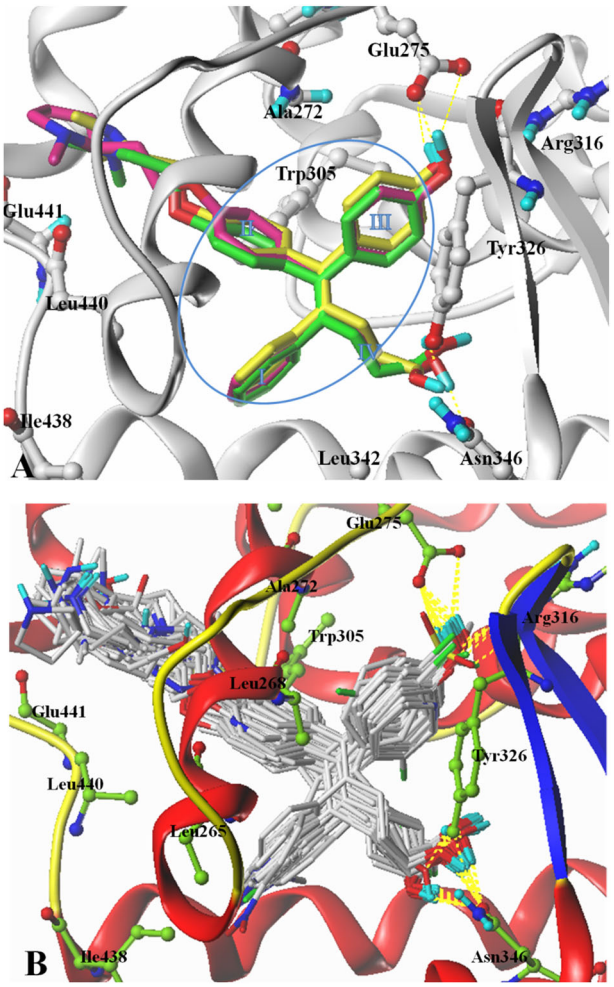


Fig. 1 **a** Predicted binding pose of ligand (green) compared with existed ligand (yellow) found in X-ray crystal structure; the position of compound 30 (red) in the active site of protein and the binding pocket of ERRγ enzyme. Hydrogen bonds are represented as yellow dotted lines and main protein residues are labeled with ball and stick form. **b** Docking-based alignment of data set molecules

Hydroxyl group in IV region had a hydrogen bond interaction with –OH of Tyr326 and another H-bonding interaction with −C=O of Asn346. Moreover, I ring interacted with the benzene ring of Trp305 through a conjugate effect, which greatly strengthened inverse modulation of **30** acted on ERRγ enzyme. Synthesized Fig. 1a, b, at the top left of the pocket, substituent group at II ring was well filled in a roomage formed by Aln272, Leu268, Trp305, Leu440, Glu441 and other resides. These action characteristics proved that compound **30** was the most active molecule in the data set.

As shown in Fig. 1b, the selected 60 molecules are aligned at same position after molecular docking program and possess similar features in that masses of they interact with Glu275 and Asn346 by hydrogen bond actions. The various binding activities result from different factors: groups in the III domain interacted with Glu275 or Arg316 or both together and groups at IV space militate with Tyr326 or Asn346. Substituents at I position are different small groups. The active site substitutional groups at the extension of II section occupied are also dissimilar. These diverse elements resulted in the selected 60 molecules used to perform the molecular modeling studies possessing multiple IC_{50} values.

3D-QSAR analysis studies

The obtained training set containing rational docking pose were subjected to establish 3D-QSAR model using partial least squares (PLS) statistics with different field contribution values. To pick up the best field combination model and avoiding the over-fitting problem, various stability statistics were taken into consideration including optimum NOC, cross-validated correlation coeffcient (q^2), coeffcient of determination (r^2) and standard error of estimate (SEE) as well as F statistical values. In principle, if q^2 and r^2 have higher values and SEE have smaller value, the built 3D-QSAR model is more accurate. In general, q^2 and r^2 are two important indicators that the model is highly predictive when q^2 is more than 0.5 and r^2 is greater than 0.8 (Cleves and Jain 2006). Therefore, a reasonable CoMFA model ($q^2 = 0.644$, NOC = 6, $r^2 = 0.976$) and CoMSIA model ($q^2 = 0.693$, NOC = 8, $r^2 = 0.992$) were developed for the selected training set. The detailed statistical summary of the CoMFA and CoMSIA analysis is shown in Table 2.

A reasonable CoMFA model, embodied steric and electrostatic fields, was accepted on the foundation of satisfactory q^2, r^2 and SEE values, presenting 0.644, 0.976 and 0.123 respectively. Likewise, q^2, r^2, and SEE values of CoMSIA model also acquired good results (0.693, 0.992, and 0.072, respectively) at the condition of steric, electrostatic, hydrophobic, H-bond acceptor and donor fields were all employed. Quantitative data in Table 2 indicated both CoMFA and CoMSIA model were reliable and reasonable.

Contour map analysis

Contour maps were generated from CoMFA and CoMSIA models directed towards changing achieved results into visual pictures, which could provide a comprehensive understanding of key structural requirements responsible for the biological activity. The diverse maps are described in the following.

CoMFA contour map analysis

CoMFA contour maps including steric and electrostatic contour maps are vividly displayed in Figs. 2a, b aimed to

Table 2 Detailed statistical summary of CoMFA and CoMSIA models

CoMFA	NOC	q^2	r^2	SEE	F value	Field contributions				
						S	E	H	D	A
S + E	6	0.644	0.976	0.123	357.048	0.396	0.604	–	–	–
S + E	7	0.678	0.953	0.176	146.317	0.226	0.774	–	–	–
S + E + H	8	0.621	0.986	0.096	373.875	0.150	0.532	0.318	–	–
S + E + D	7	0.687	0.967	0.147	214.861	0.148	0.463	–	0.389	–
S + E + A	7	0.661	0.960	0.161	175.732	0.139	0.447	–	–	0.414
S + E + H + D	7	0.632	0.978	0.121	318.431	0.106	0.343	0.236	0.315	–
S + E + H + A	7	0.647	0.977	0.123	309.324	0.111	0.348	0.224	–	0.317
S + E + D + A	7	0.694	0.969	0.142	229.929	0.101	0.296	–	0.304	0.299
S + E + H + D + A	8	0.693	0.992	0.072	634.037	0.080	0.243	0.168	0.265	0.244

q^2 cross-validated correlation coefficient, *NOC* optimumnumber of components, r^2 non cross-validated correlation coefficient, *SEE* standard error of estimation, *F* value *F*-test value, *S* steric, *E* electrostatic, *H* hydrophobic, *A* acceptor, *D* donor

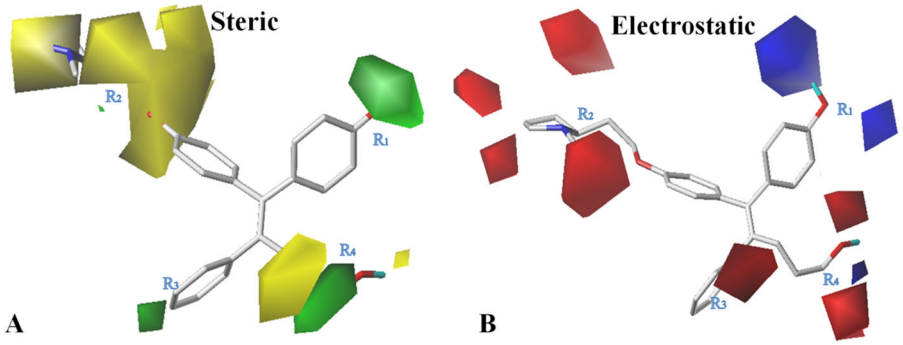


Fig. 2 CoMFA StDev*Coeff contour maps **a** Steric contour map (green: favored; yellow: disfavored); **b** Electrostatic contour map (blue: favored; red: disfavored). Compound 30 is shown as capped sticks model

illustrate which substituted groups are reasonable. Compound 30 is selected as template molecule.

In the CoMFA steric contour map (Fig. 2a), green fields favor bulky groups and red maps represent the opposite. Green contour maps appeared at R_1 and R_4 positions, indicating that more bulky groups in these regions could improve activity. This possibly explained that the binding activity of compounds **2** ($IC_{50} = 0.52\ \mu M$), **8** ($IC_{50} = 0.11\ \mu M$) and compound **9** ($IC_{50} = 1.3\ \mu M$) with chlorine, hydroxyl and trifluoromethyl groups at the R_1 position was threefold even fortyfold more potent than compound **1** ($IC_{50} = 3.9\ \mu M$) with only a fluorine atom at R_1 position. In addition, yellow contours occurred at the R_2 and R_4 position suggests that adding a bulky substitution in this region could decrease activity, this may explain why the activity of compound **56** ($IC_{50} = 10\ \mu M$) with an bulky hexatomic ring group at R_2 position dropped sharply compared with other compounds.

In the Fig. 2b displaying electrostatic contour maps, blue contours located near R_1 and R_4 position imply that positive substitutions in these regions could increase the physiological action. This was in accordance with the fact that compound **43** ($IC_{50} = 0.326\ \mu M$) with a hydrogen atom at the extremity of R_1 position was more potent than compound **46** ($IC_{50} = 8.233\ \mu M$) with a nitro here. Meanwhile, the red contours appeared at three places except R_1 position interpreted favorable influence of negative compositions. Compound **53**, **54**, **55** and **58** without oxygen atoms linking to benzene ring displayed bad IC_{50} values.

CoMSIA contour map analysis

CoMSIA contour maps analysis including steric, electrostatic, hydrophobic, H-bond donor and H-bond acceptor maps are shown in the following, accompanied with compound **30** as the template molecule in the detailed interpretation.

CoMSIA steric/electrostatic contour maps had some similar and different features compared with CoMFA steric/electrostatic contour maps. Green map in Fig. 3a emerged at R_2 position manifested long straight chains here instead of branched chains were completely necessary. Blue maps in Fig. 3b occurred at three areas however red map only located at one position, which suggested positive groups at

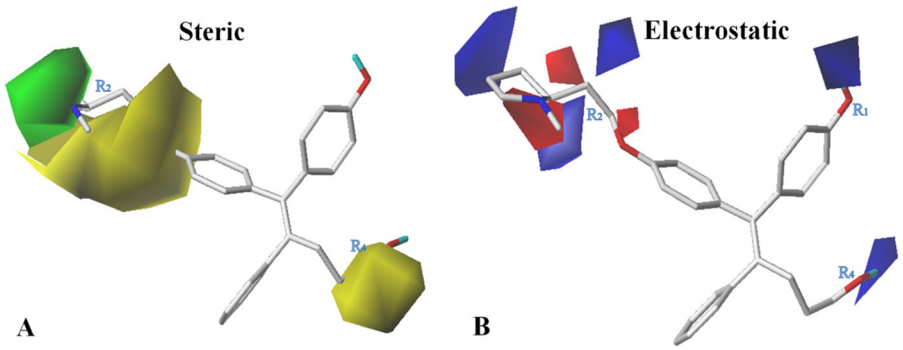


Fig. 3 CoMSIA StDev*Coeff contour maps **a** Steric contour map (green: favored; yellow: disfavored); **b** Electrostatic contour map (blue: favored; red: disfavored). Compound 30 is shown as capped sticks model

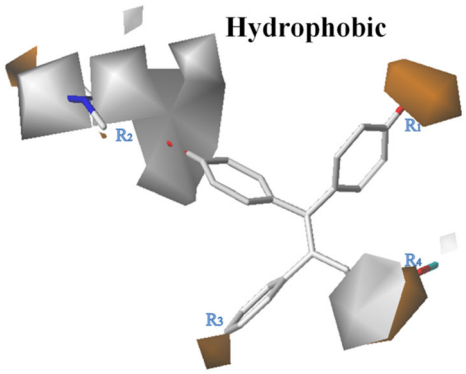


Fig. 4 CoMSIA StDev*Coeff contour maps: Hydrophobic contour map (white: favored; orange: disfavored). Compound 30 is shown as capped sticks model

R_2 position also had advantageous influence in improving activity. This could explain the fact that compound **26** with a hydrogen atom at the top of R_2 position displayed the good bioactivity than **25** with an oxygen atom here.

The hydrophobic contour map of CoMSIA model is shown in Fig. 4. White contours at the R_2 and top of R_4 position indicate that hydrophobic groups here are beneficial for binding activities. This is consistent with the fact mostly compounds possessed straight hydrocarbyl substituents in these regions. Orange contours mainly around R_1 and R_4 positions revealed hydrophilic groups here are helpful to enhance the activity. Hence compound **7** ($IC_{50} = 0.19\ \mu M$) and **8** ($IC_{50} = 0.11\ \mu M$), embracing hydroxyl group at R_1 and R_4 place, showed higher activity levels than compound **21** ($IC_{50} = 0.89\ \mu M$), **22** ($IC_{50} = 0.59\ \mu M$) and **23** ($IC_{50} = 0.48\ \mu M$) with fluorine substitutions.

The H-bond donor and acceptor of CoMSIA contour map are shown in Fig. 5. The remarkable cyan contours implied the admirable presence of hydrogen bond donor groups were come out at R_1 and R_4 position and purple maps largely appeared at R_2 position. This could be validated by compound **30** ($IC_{50} = 0.02\ \mu M$) and **36** ($IC_{50} = 0.784\ \mu M$) possessing hydroxyl group at R_1 position compared respectively with compound **32** ($IC_{50} = 0.298\ \mu M$) and **37** ($IC_{50} = 1.657\ \mu M$) occupied by bromine atom here. The enormous magenta contour around the R_4 position showed that H-bond acceptor groups in this position could increase the activity of molecules. This may explain why compounds **7** ($IC_{50} = 0.19\ \mu M$) and **8** ($IC_{50} = 0.11\ \mu M$) with hydroxyl group at R_4 position displayed better IC_{50} values than compound **22** ($IC_{50} = 0.59\ \mu M$).

Model validation of CoMFA and CoMSIA models

The experimental and predicted pIC_{50} values of the developed QSAR models are listed in Table 3. The corresponding scatter plots according to Table 3 are shown in Fig. 6. As shown in Fig. 6, the coefficient value of CoMFA/ CoMSIA regression model were respectively 0.9899 and 0.9965, constant value of the both were 0.01556 and 0.0091, which could further validate the developed CoMFA and CoMSIA models are reliable and accurate and used to predict the activities of new designed molecules.

Design and validation of novel derivatives

According to foregoing QSAR study, summarized work (shown in Fig. 7) directed the design and prediction of ten novel 2-(5-Phenylamino-biphenyl-2-yl)-ethanol derivatives. The ten derivatives were docked into active site and then were aligned. Their activities were not only calculated to be better than ligand ($pIC_{50} = 3.9586$), but also docking scores of ten newly designed compounds were higher compared with ligand. The detailed structures, predicted pIC_{50} values and docking score are displayed in Table 4.E1 was selected for detailed comparison of newly-designed molecules and ligand as it possessed the highest docking score. As shown in Fig. 8a, it strongly displays the similarities between E1 and ligand such as the occupied space and interaction. E1 and ligand both

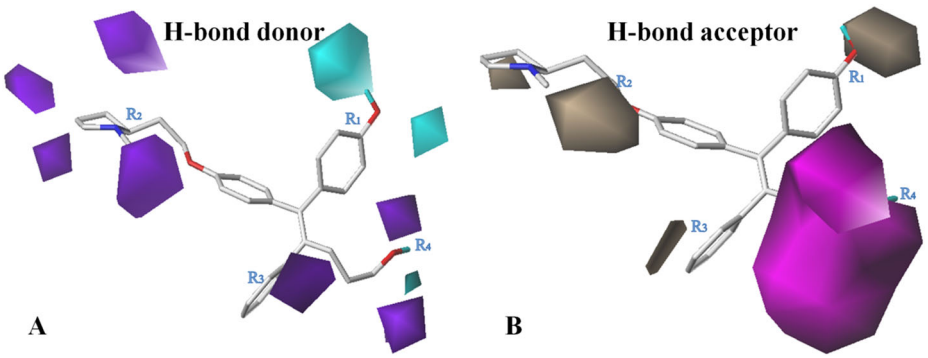


Fig. 5 CoMSIA StDev*Coeff contour maps: **a** H-bond donor map (cyan: favoured; purple: disfavoured). **b** H-bond acceptor map (magenta: favoured; brown: disfavoured). Compound 30 is shown as capped sticks model

Table 3 Experimental and predicted activity values for the developed models

Training Set Compounds	pIC_{50}	CoMFA		CoMSIA	
		Predicted	Residuals	Predicted	Residuals
3	2.174	2.113	0.061	2.106	0.068
4	2.824	2.802	0.022	2.875	−0.052
5	3.745	3.722	0.023	3.776	−0.032
7	3.721	3.760	−0.038	3.704	0.017
8	3.959	4.039	−0.081	4.111	−0.152
9	2.886	2.975	−0.089	2.860	0.027
10	2.000	1.958	0.042	1.921	0.079
11	2.000	2.018	−0.018	1.887	0.113
12	3.569	3.576	−0.008	3.599	−0.030
13	3.000	2.980	0.020	2.893	0.107
14	3.301	3.311	−0.010	3.249	0.052
15	2.959	2.949	0.010	2.906	0.052
18	2.000	2.006	−0.006	2.082	−0.082
19	2.000	1.984	0.016	2.055	−0.055
20	2.252	2.246	0.006	2.348	−0.097
21	3.051	3.093	−0.043	2.998	0.053
22	3.229	3.171	0.059	3.289	−0.060
23	3.319	3.345	−0.026	3.318	0.000
25	3.498	3.508	−0.011	3.352	0.146
26	4.032	4.041	−0.010	4.061	−0.029
27	4.174	4.167	0.007	4.256	−0.082
28	4.155	4.118	0.037	4.136	0.019
29	3.967	4.009	−0.043	3.982	−0.016
30	4.699	4.731	−0.032	4.707	−0.008
31	4.585	4.590	−0.006	4.508	0.077
33	2.981	3.056	−0.074	3.086	−0.104
36	3.529	3.541	−0.012	3.562	−0.034
37	2.781	2.730	0.050	2.754	0.027
39	4.081	4.043	0.038	4.085	−0.004
41	4.041	4.043	−0.002	4.043	−0.003
42	3.783	3.765	0.018	3.747	0.036
43	3.487	3.426	0.061	3.550	−0.063
45	2.652	2.708	−0.057	2.647	0.004
46	2.084	2.043	0.042	2.103	0.010
47	2.638	2.664	−0.026	2.601	−0.019
48	2.893	2.878	0.015	2.858	0.034
50	3.337	3.342	−0.005	3.288	0.049
52	3.090	3.178	−0.088	3.276	−0.186

Table 3 (continued)

Training Set Compounds	pIC_{50}	CoMFA		CoMSIA	
		Predicted	Residuals	Predicted	Residuals
53	2.556	2.632	−0.076	2.560	−0.004
54	2.811	2.766	0.045	2.745	0.066
57	3.572	3.490	0.082	3.599	−0.041
58	2.201	2.150	0.051	2.079	0.122
59	3.049	3.000	0.049	3.189	−0.140
60	2.671	2.682	−0.011	2.766	−0.095
Test Set Compounds	**pIC50**	**CoMFA**		**CoMSIA**	
		Predicted	**Residuals**	**Predicted**	**Residuals**
1*	2.409	2.568	−0.160	2.634	−0.225
2*	3.284	3.158	0.126	3.086	0.198
6*	4.048	4.062	−0.015	4.060	−0.012
16*	2.000	1.988	0.012	1.991	0.009
17*	2.000	1.976	0.024	2.048	−0.048
24*	2.657	2.634	0.022	2.650	0.007
32*	3.106	3.061	0.045	3.241	−0.135
34*	2.818	2.956	−0.139	2.902	−0.085
35*	4.194	4.200	−0.006	4.188	0.006
38*	4.102	4.009	0.093	3.957	0.145
40*	4.301	4.310	−0.009	4.342	−0.041
44*	3.959	3.847	0.112	3.952	0.007
49*	2.901	3.128	−0.227	2.969	−0.068
51*	3.606	3.413	0.192	3.460	0.145
55*	2.631	2.643	−0.012	2.624	0.006
56*	2.000	1.973	0.027	2.041	−0.027

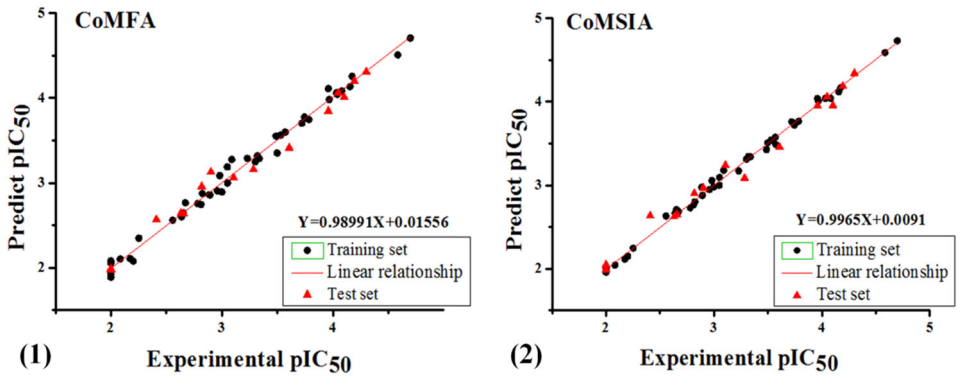


Fig. 6 Correlation between the predicted and experimental pIC_{50} of the training and test set compounds: (**1**) scatter plot of CoMFA. (**2**) scatter plot of CoMSIA. Black-color circles represent training set compounds and red-color triangles are test set compounds

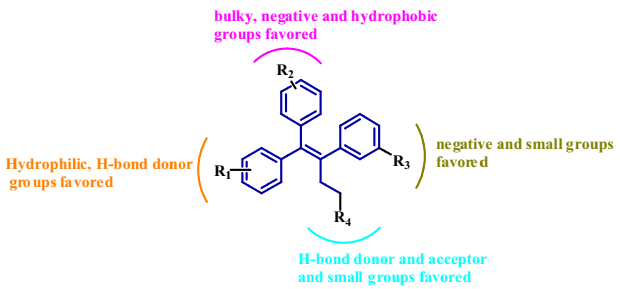
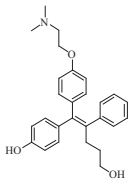
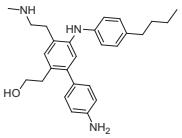
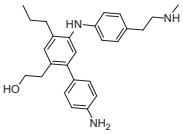
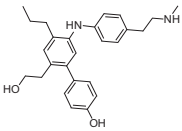
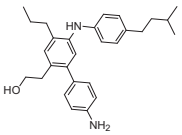
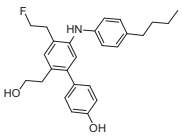
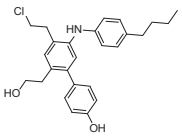
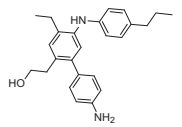
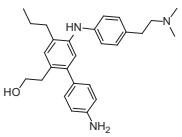
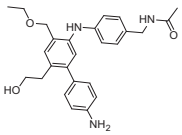
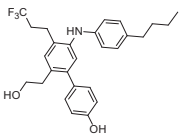


Fig. 7 SAR summarized results based on our work

have H-bond action with Glu275, Tyr326 and Asn346. The difference is ligand conjugated with Trp305 but E1 interacts with Leu309 and Ala431 by hydrogen action. As depicted in Fig. 8b, ten novel molecules were aligned well based on interactions between molecule and enzyme in active site, more importantly, they were at same position in binding pocket of ERRγ enzyme and similar interaction compared with hydroxytamoxifen derivatives and ligand. In addition, ten novel analogs were provided with stable physical and chemical properties and simple synthesis. Such results indicate that these 3D-QSAR

Table 4 The detailed structures, predicted pIC_{50} values and docking score of novel 10 compounds

Number	Structure	Docking scorce	Predicted pIC_{50}	
			CoMFA	CoMSIA
2EWP-ligand		10.2296	3.850	3.918
E1		11.6962	4.377	4.432
E2		11.5840	4.462	4.639
E3		11.3018	4.459	4.075
E4		11.2465	3.862	4.894
E5		11.0804	4.100	4.501
E6		10.9861	4.196	4.488
E7		10.8829	4.512	4.289
E8		10.8636	4.729	4.825
E9		10.5908	4.495	4.971
E10		10.5841	4.484	4.210

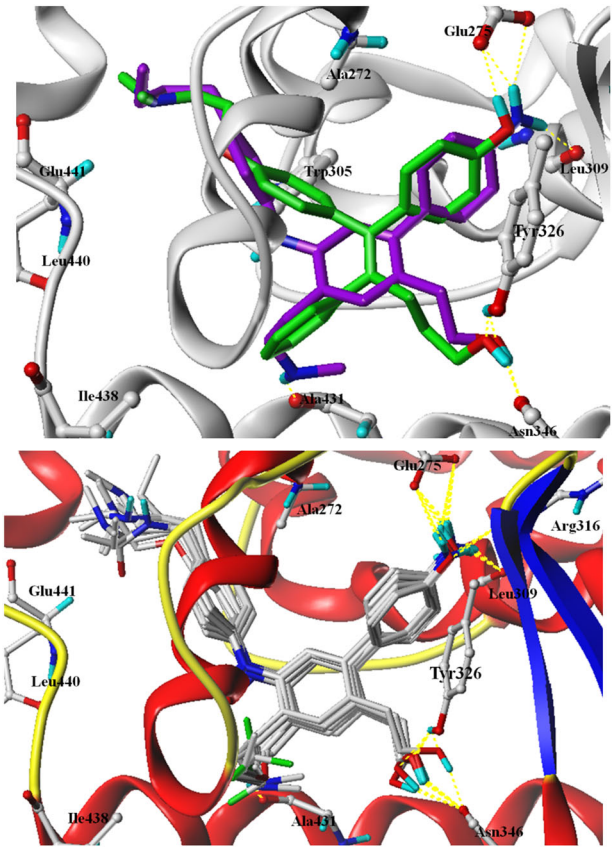


Fig. 8 **a** Docked conformation of E1 (purple) compared with reasonable binding pose of ligand (green) found in X-ray crystal structure; the position of E1 in the active site of protein and the binding pocket of ERRγ enzyme. Hydrogen bonds are represented as yellow dotted lines and main protein residues are labelled with ball and stick form. **b** Docking-based alignment and the position in the binding pocket of ten 2-(5-Phenylamino-biphenyl-2-yl)-ethanol derivatives

models with potential predictive ability could be prospectively used in structure modification and optimization.

Conclusion

A desirable 3D-QSAR study on hydroxytamoxifen derivatives was generated based on CoMFA/CoMSIA models and rational docking conformations. The reasonable CoMFA ($q^2 = 0.644$, $r^2 = 0.976$) and CoMSIA models ($q^2 = 0.693$, $r^2 = 0.992$) displayed satisfactory predictive abilities validated by predicting the activities of test set and new designed compounds. Moreover, 3D-QSAR models also presented admirable correlations indicated the stability quality of CoMFA/CoMSIA models. According to the useful information provided by CoMFA and CoMSIA contour maps, ten 2-(5-Phenylamino-biphenyl-2-yl)-ethanol derivatives possessed rational docked pose and higher pIC_{50} values as well as steady binding position in active site of ERRγ enzyme were designed. These results give evidence that this 3D-QSAR models and later devised compounds could further provide a valuable clue in designing novel ERRγ inverse agonists with high modulation activity and reliability safety.

Acknowledgements This work were supported by a grant from the National Natural Science Foundation of China (81872744) and Shandong Natural Science Foundation of China (ZR2019MH046).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Aalizadeh R, Pourbasheer E, Ganjali MR (2015) Analysis of B-Raf inhibitors using 2D and 3D-QSAR, molecular docking and pharmacophore studies. Mol Divers 19:915–930

Abad MC, Askari H, O'Neill J, Klinger AL, Milligan C, Lewandowski F, Springer B, Spurlino J, Rentzeperis D (2008) Structural determination of estrogen-related receptor γ in the presence of phenol derivative compounds. J Steroid Biochem 108:44–54

Balasubramanian PK, Balupuri A, Gadhe CG, Cho SJ (2014) 3D QSAR modeling study on 7-aminofuro [2, 3-c] pyridine derivatives as TAK1 inhibitors using CoMFA and COMSIA. Med Chem Res 24:2347–2365

Clark M, Cramer RD, Van Opdenbosch N (1989) Validation of the general purpose tripos 5.2 force field. J Comput Chem 10:982–1012

Cleves AE, Jain AN (2006) Robust ligand-based modeling of the biological targets of known drugs. J Med Chem 49:2921–2938

Coward P, Lee D, Hull MV, Lehmann JM (2001) 4-Hydroxytamoxifen binds to and deactivates the estrogen-related receptor gamma. Proc Natl Acad Sci USA 98:8880–8884

Cramer RD, Patterson DE, Bunce JD (1988) Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc 110:5959–5967

Cramer RD, Patterson DE, Bunce JD (1989) Recent advances in comparative molecular field analysis (CoMFA). Progre Clini Biolo Res 291:161–165.

Giguére V (2002) To ERR in the estrogen pathway. Trends Endocrin Met 13:220–225

Greschik H, Wurtz JM, Sanglier S, Bourguet W, van Dorsselaer A, Moras D, Renaud JP (2002) Structural and functional evidence for ligand-independent transcriptional activation by the estrogen-related receptor 3. Mol Cell 9:303–313

Horard B (2003) Estrogen receptor-related receptors: orphan receptors desperately seeking a ligand. J Mol Endocrinol 31:349–357

Jung YS, Lee JM, Kim DK, Lee YS, Kim KS, Kim YH, Kim J, Lee MS, Lee IK, Kim SH, Cho SJ, Jeong WI, Lee CH, Harris RA, Choi HS (2016) The orphan nuclear receptor ERRgamma regulates hepatic cb1 receptor-mediated fibroblast growth factor 21 gene expression. PloS one 11:e0159425

Kallen J, Schlaeppi JM, Bitsch F, Filipuzzi I, Schilb A, Riou V, Graham A, Strauss A, Geiser M, Fournier B, Fournier B (2004) Evidence for ligand-independent Transcriptional Activation of the Human Estrogen-related Receptor α (ERRα). J BIOL CHEM 279:49330–49337

Kim DK, Kim JR, Koh M, Kim YD, Lee JM, Chanda D, Park SB, Min JJ, Lee CH, Park TS, Choi HS (2011) Estrogen-related receptor gamma (ERRgamma) is a novel transcriptional regulator of phosphatidic acid phosphatase, LIPIN1, and inhibits hepatic insulin signaling. J Biol Chem 286: 38035–38042

Kim J, Chin J, Im CY, Yoo EK, Woo S, Hwang HJ, Jh Cho, Ka Seo, Song J, Hwang H, Kim KH, Kim ND, Yoon SK, Jeon JH, Yoon SY, Jeon YH, Choi HS, Lee IK, Kim SH, Cho SJ (2016) Synthesis and biological evaluation of novel 4-hydroxytamoxifen analogs as estrogen-related receptor gamma inverse agonists. Eur J Med Chem 120:338–352

Kim J, Woo SY, Im CY, Yoo EK, Lee S, Kim H-J, Cho JH, Lee WS, Yoon H, Kim S, Kwon OB, Hwang H, Kim KH, Jeon JH, Singh TD, Kim SW, Hwang SY, Choi HS, Lee IK, Kim SH, Jeon YH, Chin J, Cho SJ (2016) Insights of a lead optimization study and biological evaluation of novel 4-hydroxytamoxifen analogs as estrogen-related receptor γ (ERRγ) Inverse Agonists. J Med Chem 59:10209–10227

Klebe G, Abraham U, Mietzner T (1994) Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. J Med Chem 37: 4130–4146

Lin H, Doebelin C, Patouret R, Garcia-Ordonez RD, Chang MR, Dharmarajan V, Bayona CR, Cameron MD, Griffin PR, Kamenecka TM (2017) Design, synthesis, and evaluation of simple phenol amides as ERRgamma agonists. Bioorg Med Chem Lett 28:1313–1319

Nam K, Marshall PM, Wolf R, Cornell W (2002) Simulation of the different biological activities of diethylstilbestrol (DES) on estrogen receptor α and estrogen-related receptor γ. Biopolymers 68:130–138

Narkar VA, Fan W, Downes M, Yu RT, Jonker JW, Alaynick WA, Banayo E, Karunasiri MS, Lorca S, Evans RM (2011) Exercise and PGC-1alpha-independent synchronization of type I muscle metabolism and vasculature by ERRgamma. Cell Metab 13:283–293

Ron Wehrens HP, Lutgarde MC, Buydens (2000) The bootstrap: a tutorial. Chemom Intell Lab 54:35–52

Wang L, Zuercher WJ, Consler TG, Lambert MH, Miller AB, Orband-Miller LA, McKee DD, Willson TM, Nolte RT (2006) X-ray crystal structures of the estrogen-related receptor-gamma ligand binding domain in three functional states reveal the molecular basis of small molecule regulation. J Biol Chem 281:37773–37781

Wold S, Ruhe A, Wold H, Dunn III WJ (1984) The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. Siam J Sci Stat Comput 5:735–743

Yu DD, Forman BM (2005) Identification of an agonist ligand for estrogen-related receptors ERRβ/γ. BIOORG MED CHEM LETT 15:1311–1313

Yu DD, Huss JM, Li H, Forman BM (2017) Identification of novel inverse agonists of estrogen-related receptors ERRγ and ERRβ. Bioorg Med Chem Lett 25:1585–1599

Zhang Y, Kim DK, Lee JM, Park SB, Jeong WI, Kim SH, Lee IK, Lee CH, Chiang JY, Choi HS (2015) Orphan nuclear receptor oestrogen-related receptor gamma (ERRgamma) plays a key role in hepatic cannabinoid receptor type 1-mediated induction of CYP7A1 gene expression. Biochem J 470:181–193

Zuercher WJ, Gaillard S, Orband-Miller LA, Chao EYH, Shearer BG, Jones DG, Miller AB, Collins JL, McDonnell DP, Willson TM (2005) Identification and structure−activity relationship of phenolic acyl hydrazones as selective agonists for the estrogen-related orphan nuclear receptors ERRβ and ERRγ. J Med Chem 48:3107–3109