J Mol Model (2008) 14:149–159
DOI 10.1007/s00894-007-0264-x

ORIGINAL PAPER

3D-QSAR and docking studies of 3-arylquinazolinethione derivatives as selective estrogen receptor modulators

Aijing Xiao · Zhuoyong Zhang · Liying An · Yuhong Xiang

Received: 6 September 2007 / Accepted: 5 December 2007 / Published online: 3 January 2008

Abstract 3D-QSAR and molecular docking analysis were performed to explore the interaction of estrogen receptors (ERα and ERβ) with a series of 3-arylquinazolinethione derivatives. Using the conformations of these compounds revealed by molecular docking, CoMFA analysis resulted in the first quantitative structure-activity relationship (QSAR) and first quantitative structure-selectivity relationship (QSSR) models predicting the inhibitory activity against ERβ and the selectivity against ERá. The q^2 and R^2 values, along with further testing, indicate that the obtained 3D-QSAR and 3D-QSSR models will be valuable in predicting both the inhibitory activity and selectivity of 3-arylquinazolinethione derivatives for these protein targets. A set of 3D contour plots drawn based on the 3D-QSAR and 3D-QSSR models reveal modifications of substituents at C2 and C5 of the quinazoline which my be useful to improve both the activity and selectivity of ERβ/ ERα. Results showed that both the steric and electrostatic factors should appropriately be taken into account in future rational design and development of more active and more selective ERβ inhibitors for the therapeutic treatment of osteoporosis.

Keywords CoMFA · Dock · SERMs · 3-arylquinazolinethione · 3D-QSAR

Introduction

Estrogens are a family of naturally steroid hormones that exert a critical role in the growth, development, and sustenance of a wide range of tissues via interaction with the Estrogen receptor (ER), an important transcription factor belonging to the nuclear receptor superfamily. ER plays a key role in reproductive, cardiovascular and central nervous systems and bone tissue. Thus, ER is a key drug target for the treatment of osteoporosis and breast cancer [1]. For the past decade, the physiological effects of estrogens were attributed to a single receptor of the ligand-activated transcription factor family, now known as ERα. The discovery of the second estrogen receptor, ERβ, in 1996 [2, 3] resulted in the possibility that the tissue selectivity and function of certain estrogens and anti-estrogens were due to their specificity for either the classical ERα or the newly discovered isoform has been validated by recent studies concerning the difference in tissue distribution between the two estrogen receptor isoforms [4–8]. In addition it has been reported that the pharmacology of several classical estrogen receptor agonists and antagonists is reversed for ERβ [9, 10].

The two receptors differ in size, with ERα having 595 amino acids and ERβ having 485 amino acids. The predominant ER in the female reproductive tract and mammary glands is ERα, whereas ERβ is the primary ER in vascular endothelial cells, bone and male prostrate tissues [6]. The compounds having the potential to modulate selectivity of the different estrogen target tissues are known as selective estrogen receptor modulators (SERMs) [2].

A number of SERMs are currently used 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 and prevention or treatment for osteoporosis. However, both these agents have been linked to increased risks of thromboembolism and tamoxifen has been proved to increase the risk of

A. Xiao · Z. Zhang (✉) · L. An · Y. Xiang
Department of Chemistry, Capital Normal University,
Beijing 100037, People's Republic of China
e-mail: gusto2008@vip.sina.com

endometrial cancer. Hence the search for more tissue specific analogues continues, so as to develop distinct SERMs with fewer side effects [7]. Recently 3-Arylquinazolinethione derivatives have been reported as potent ERβ selective ligands. The present paper explores selectively requirements of 3-Arylquinazolinethione derivatives for binding with ERβ versus ERα [8].

Nowadays, three-dimensional (3D) quantitative structure-activity relationship (3D-QSAR) techniques, such as comparative molecular field analysis (CoMFA) are routinely used in modern drug design to help in understanding drug-receptor interaction. These computational techniques have been proved particularly helpful in the design of novel, more potent inhibitors by revealing the mechanism of drug-receptor interaction [9]. The applications of QSAR methodology to some SERMs have been reported [7, 10]. However, to our knowledge, there has been no report concerning the application of a QSSR methodology to the selectivity of SERMs, although the quantitative structure-selectivity relationship (QSSR) is also crucial for the development of these SERMs.

The QSAR and the QSSR studies of ERβ inhibitors, along with molecular docking modeling of the protein-inhibitor binding were performed in the present paper. 3-Arylquinazolinethione derivatives are a novel kind of SERMs with good activity and selectivity. The topographical features of ERα and ERβ active sites were discussed based on the obtained 3D-QSAR and 3D-QSSR models. The satisfactory QSAR and QSSR models obtained provide a solid basis for future rational design of more active and more selective SERMs within the family of 3-Arylquinazolinethione derivatives.

Data set and methodology

Data sets

All compounds examined in the present study were reported recently by Timur Güngör and co-workers [8]. Within a total of 54 compounds reported, nine of them were discarded, because the IC50 (i.e., the concentration causing 50% inhibitory effect) values were not available for these compounds. In the QSAR analysis, the 45 compounds with the IC50 values against ERβ were randomly divided into two sets, a training set with 38 compounds and a test set with seven. In the QSSR analysis, 32 compounds with the IC50 values against ERβ and ERα were randomly selected from the total 36 compounds as training set and the remaining four compounds were used as test set. An attractive feature of these compounds is their relative conformational rigidity, which makes them more amenable to meaningful CoMFA analysis than flexible molecules. The IC50 (ERβ) values were converted to pIC50 (i.e., -logIC50) values. The log of the reported IC50 (ERα/β) value, that is, logIC50 (ERα/β), can be used as an index for the selectivity (ERα/β). The pIC50 (ERβ) and logIC50 (ERα/β) values and the compounds were listed in Tables 1 and 2, respectively.

Molecular docking

The crystal structure of ERβ in complexation with genistein (PDB entry code 1QKM) [11] was extracted from Brookhaven Protein Database (PDB http://www.rcsb.org/pdb). Molecular modeling of 3-arylquinazolinethione derivatives based on the structure of genistein was performed using software Sybyl 7.1 [12].

For the purpose of tackling the interacting mode of 3-arylquinazolinethione derivatives (inhibitors) with ERβ (enzyme), the advanced docking program AutoDock 3.05 [13–15] was used to perform the automated molecular docking of the representative flexible ligands (compounds Genistein, 1aar, 1ax, and 1aag in Table 1). The Lamarckian genetic algorithm (LGA) [15] was applied to deal with the inhibitor-enzyme interactions. Briefly, the LGA described the relationship between the inhibitors and the enzymes by the translation, orientation, and conformation of the inhibitors. These so-called ‘state variables’ were the inhibitors’ genotype, and the resulting atomic coordinates together with the interaction and the intramolecular energies were the inhibitors’ phenotype. The environmental adaptation of the phenotype was reverse-transcribed into its genotype and became heritable traits. Each docking cycle, or generation, consisted of a regimen of fitness evaluation, cross-over, mutation, and selection. A Solis and Wets local search [16] performed the energy minimization on a user-specified proportion of the population. The docked structures of the inhibitors were generated after a reasonable number of evaluations. The whole docking operation could be stated as follows.

First, the ERβ molecule was checked for polar hydrogens and assigned for partial atomic charges, then the PDBQs file was created, and the atomic salvation parameters were also assigned for the macromolecules. The torsion angles of inhibitors were defined in order to explore conformations during the docking process.

Second, since the program AutoDock allows only one of the docking partners to be flexible, the receptor was kept rigid, and the ligand was allowed to be flexible. The 3D grid box ($50 \cdot 50 \cdot 50$ Å) centered on the putative active site and the lattice point distance of 0.375 Å was created by the AutoGrid algorithm to evaluate the binding energies between the 3-arylquinazolinethione derivatives and ERβ. In this stage, the ERβ was embedded in the 3D grid and a probe atom was placed at each grid point. The affinity and electrostatic potential grid were calculated for each type of atom in the

Table 1 Structures of 3-Arylquinazolinethione derivatives

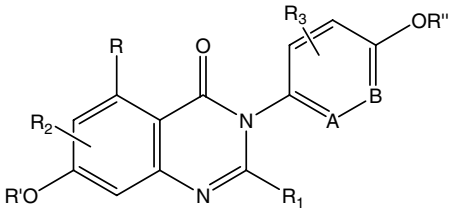
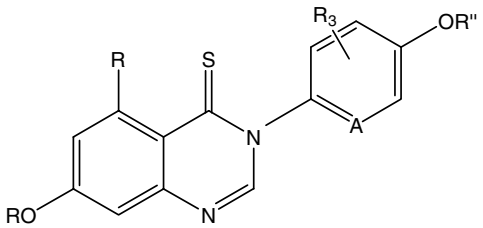
compd no.	R_1	R_2	R_3	R	R’	R”
Panel A						
1aa	H	H	H	OH	H	H
1ab	H	H	H	OH	CH_3	H
1ac	H	H	2’-CH_3	OH	H	H
1ad	H	H	3’-Cl	OH	H	H
1ae	H	H	3’-F	OH	H	H
1ag	H	H	3’-F	OH	H	H
1ah	H	H	2’-Cl	OH	H	H
1ai	H	H	3’-CH_3	OH	H	H
1ak	H	H	H	OH	H	H
1al	H	H	H	OH	H	H
1am	H	8-I	H	OH	H	H
1an	H	6,8-di-I	H	OH	H	H
1ao	H	8-CH3	H	OH	H	H
1ap	H	6-I	H	OH	H	H
1aq	H	6-nPr	H	OH	H	H
1as	H	H	H	H	H	H
1at	H	H	H	OH	H	H
1au	H	H	H	H	H	H
1av	CH_3	H	H	OH	H	H
1aw	C_2H_5	H	H	OH	H	H
1ax	nPr	H	H	OH	H	H
1ay	nBu	H	H	OH	H	H
1az	iPr	H	H	OH	H	H

Table 1 (continued)

1aab	SH	H	H	OH	CH_3	H
1aac	SH	H	H	OH	H	H
1aad	Cl	H	H	OH	H	H
1aae	OH	H	H	OH	H	H
1aag	4-OHPh	H	H	H	H	H
1aah	H	H	H	OCH_2CH_3	H	H
1aai	H	H	H	NH_2	H	H
1aaj	H	H	H	CH_3	H	H
1aal	H	H	H	CN	H	H
1aam	H	H	H	OSO_2CF_3	H	H
1aan	H	H	H	CH_2CH_3	H	H
1aao	H	H	H	OH	H	H
1aap	H	H	H	CO_2CH_3	H	H
1aaq	H	H	H	PH	H	H
1aar	H	H	H	CH_2OH	H	H
1aas	H	H	H	OH	H	H

Panel B



1ba	H	OH	H	H	H	OH
1bb	3'-F	OH	H	H	3'-F	OH
1bc	3'-CH_3	OH	H	H	3'-CH_3	OH
1bd	H	CH_3	H	H	H	CH_3
1be	H	CH_2CH_3	H	H	H	CH_2CH_3
1bf	H	OH	H	H	H	OH

inhibitors. During docking processes, the Lamarckian genetic algorithm with pseudo-Solis and Wets method was used. A series of the docking parameters were set on. The number of generations, energy evaluations, and docking runs were set to 370,000, 1,500,000, and 10, respectively. The kinds of atomic charges were taken as Kollman-united-atom [17] for ERβ and Gasteiger-Hückel charges for the 3-arylquinazolinethione derivatives.

Table 2 Inhibitory activity, selectivity and predicted values

Compd no.	Meas.	3D-QSAR		Meas.	3D-QSSR	
	(pIC50(ERβ))	Calc.	Residue	(logIC50(ERα/β))	Calc.	Residue
1aa*#	6.75	5.69	1.05	1.79	1.23	0.56
1ab	5.02	4.96	0.06	0.48	0.51	-0.03
1ac	6.38	6.44	-0.06	1.00	0.94	0.06
1ad	5.14	5.34	-0.20	0.90	0.91	-0.01
1ae	6.19	5.96	0.23	1.83	1.75	0.08
1ag	6.59	6.88	-0.28	1.46	1.70	-0.24
1ah*	6.58	6.44	0.14	1.38	1.53	-0.15
1ai	5.13	5.08	0.04	0.70	0.64	0.06
1ak*	5.45	5.12	0.33	1.45	1.45	0.0047
1al	4.29	4.33	-0.05	0.00	0.15	-0.15
1am	5.39	5.45	-0.06			
1an	4.91	4.94	-0.03			
1ao*	5.89	5.61	0.30	1.08	1.12	-0.04
1ap	5.40	5.45	-0.05			
1aq	5.14	5.16	-0.03	0.18	0.22	-0.04
1as	5.07	5.26	-0.20	0.78	0.67	0.11
1at	5.42	5.46	-0.04	1.04	1.07	-0.03
1au	4.73	4.80	-0.06	1.15	1.12	0.03
1av	4.94	4.81	0.14	0.48	0.34	0.14
1aw#*	5.24	5.04	0.20	0.48	0.48	0.00
1ax	5.50	5.71	-0.22	0.30	0.17	0.13
1ay	6.60	6.67	-0.07	-0.15	-0.11	-0.04
1az	4.32	4.26	0.05	0.18	0.19	-0.01
1aab	4.14	4.07	0.07			
1aac	5.16	4.79	0.37	0.30	0.36	-0.06
1aad	4.87	5.00	-0.13			
1aae#*	5.03	5.00	0.03	0.90	0.64	0.26
1aag	3.36	3.51	-0.15	-1.00	-0.91	-0.09
1aah	4.35	4.39	-0.03			
1aai	5.08	5.15	-0.07	0.30	0.64	-0.34
1aaj*	6.50	5.86	0.64	1.48	1.17	-0.31
1aal	4.13	4.10	0.03			
1aam	4.92	4.83	0.09			
1aan	6.58	6.24	0.34	1.70	1.34	0.36
1aao	4.77	4.91	-0.14			
1aap	5.02	4.96	0.06	0.95	-0.87	0.08
1aaq	4.51	4.58	-0.07	0.48	0.37	0.11
1aar	7.62	7.49	0.13	0.70	0.85	-0.15
1aas	5.64	5.34	0.30	1.28	1.18	0.10
1ba	7.33	6.86	0.46	1.75	1.46	0.29
1bb	6.99	6.89	0.10	1.48	1.67	-0.19
1bc	6.18	6.26	-0.08	0.60	0.51	0.09
1bd#	7.10	7.00	0.10	1.34	1.35	0.01
1be	6.81	7.30	-0.49	1.18	1.41	-0.23
1bf	6.29	6.35	-0.06	1.73	1.87	-0.14

* Test samples for 3D-QSAR model validation
\# Test samples for 3D-QSSR model validation

Finally, the 10 docked conformations for each 3-arylquinazolinethione derivatives were extracted from the docking log files. The best docking conformation of each inhibitor was selected according to the criteria of interacting energy combined with geometrical matching quality for 3D-QSAR and 3D-QSSR studies. Then, other cases were docked sequentially into the binding pocket of ERβ using the parameters previously optimized.

3D-QSAR and 3D-QSSR analyses

Molecular structure building and alignment Active conformation selection is a key step for CoMFA analysis. Conformation with the lowest energy is not always the active conformation, and the proper active conformation can only be extracted from the crystal structure of the complex of the drug molecule and its binding receptor [18, 19]. In the present study, the molecular conformation of compound 1aar (the most activity compound) obtained from the molecular docking was used as a template to build molecular structures of all the compounds.

Partial atomic charges were assigned to each atom and then energy minimization of each molecule was performed using Powell method and Tripos standard force field with a distance-dependent dielectric function. The minimization was terminated when the energy gradient convergence criterion of 0.005 kcal mol^{-1} was reached or when the 2000-step minimization cycle limit was exceeded.

Molecular alignment is considered as one of the most sensitive parameters in CoMFA analysis [12]. The quality and the predictive ability of the model are directly dependent on the alignment rule. Once the active conformation was determined, pharmacophore or common substructure alignment was carried out according to some rules. In this paper, common substructure alignment was carried out using database alignment tool with compound 1aar (in the QSAR analysis) and 1ae (in the QSSR analysis) as the template molecules. Alignment of all compounds was shown in Fig. 1. It can be seen that all the compounds studied have similar active conformations.

CoMFA analysis In 3D-QSAR and 3D-QSSR analyses, all aligned molecules were put into a 3D cubic lattice that extended at least 0.4 nm beyond the volumes of all investigated molecules on all axes. The region was partitioned into hundreds of grids with certain grid spacing. In the CoMFA analysis, Lennard-Jones 6–12 and Coulomb potentials were employed to calculate the CoMFA steric and electrostatic interaction fields, respectively. A sp^3-hybrized carbon atom with a charge of +1 was used as the probe atom and the steric and electrostatic energy cutoff was 33 kcal mol^{-1} and column-filtering value set to 2.1 kcal mol^{-1}.

Partial least squares (PLS) method was carried out with the leave-one-out (LOO) cross-validation procedure to determine the optimum number of components for the final non-cross-validated 3D-QSAR and 3D-QSSR models. The optimal number of components produces the smallest root mean predictive sum of squared errors, which usually corresponds to the highest cross-validated squared coefficient (q^2) $q^2 = 1 - \frac{\sum (Y_{obs} - Y_{pre})^2}{\sum (Y_{obs} - Y_{mean})^2}$, where $Y_{predicted}$, $Y_{observed}$ and Y_{mean} are predicted, observed, and mean values of the target property (pIC50(ERβ) or logIC50(ERα/β), respectively. $\Sigma(Y_{predicted} - Y_{observed})^2$ is the predicted sum of squares (PRESS). Conventional correlation coefficient, R^2, and its standard error, s, were also computed for the final PLS models, and summarized in Table 3.

Results and discussion

Molecular docking analyses

Ligands' conformation The automated molecular docking can produce several options of binding conformations or each inhibitor. The conformation corresponding to the lowest binding energy with ERβ was selected as the possible binding conformation. The conformational superposition of Genistein from both the X-ray crystal structure and the AutoDock result was obtained. The root mean square deviation (RMSD) between these two conformations is 0.012 Å, indicating that the parameter set for the AutoDock simulation is reasonable to reproduce the X-ray structure. The AutoDock method and the parameter set could be extended to search the binding conformations of other inhibitors accordingly.

Depicted in Fig. 2 are the most stable structures for ERβ binding with compounds 1aar, 1ax, and 1aag obtained from the molecular docking. The docking with these three particular compounds was carried out because compound 1aar is associated with the highest IC50 (ERβ) value, compound 1aag is associated with the lowest IC50 (ERβ) value, and compound 1ax is associated with a middle IC50 (ERβ) value. A survey of the binding structures depicted in Fig. 2 reveals that these three compounds bind to ERβ quite similarly, providing a rational structural basis for our 3D-QSAR and 3D-QSSR analyses. A remarkable feature of the ERβ-ligand binding structures in Fig. 2 is that there are two common hydrogen bonds. One is between the nitrogen of the imidazole of the His475 or main chain nitrogen of the Leu476 and the hydroxyl hydrogen of the phenol ring of the ligand. The other is between the backbone oxygen of the Glu305 or side chain oxygen of Leu339 and the hydroxyl hydrogen of the quinazolinone of the ligand. This hydrogen bond exists in all of the docked ERβ-ligand binding structures.

Interactions of subsites To illustrate the interaction mechanism, compound 1aar, the most potent inhibitor among the 45 3-Arylquinazolinethione derivatives, was selected for more detailed analysis. In the latter discussions, all the descriptions referred to the compound 1aar unless otherwise noted. Figure 3 generally represents the interacting model of 1aar with ERβ.

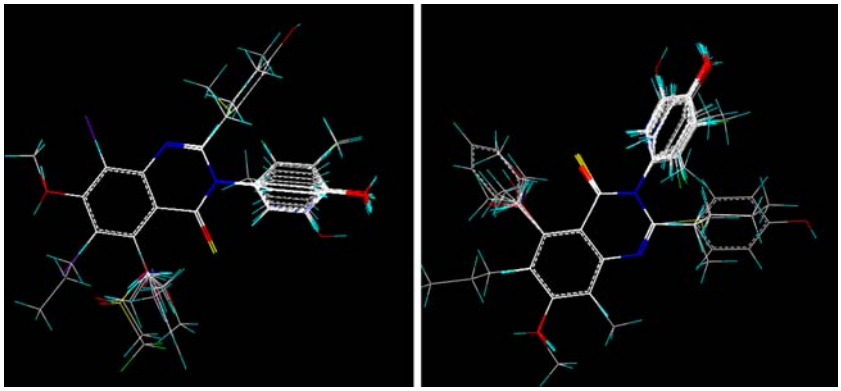


Fig. 1 Alignments of 3-Arylquinazolinethione derivatives of PIC50(ERβ) (left) and IC50(ERα/β) (right)

The quinazoline moieties of all the 45 3-Arylquinazolinethione derivatives of the training set locate much at the same site. It can be seen from Fig. 3 that the quinazoline moiety of compound 1aar is surrounded by residues Leu298, Met336, Ala302, Leu339, and Arg346, mainly through the hydrophobic interaction. In addition, the 1-position N of the compound which has bulky subsituent at the 2-position does not form a hydrogen bond with the main chain of Leu298, all of these compounds have low activity such as the compounds 1av,1aw, 1ax, 1az, 1aab, 1aad, and 1aag, which is consistent with the result of the QSAR model. Thereinto, compound 1ay (with a butyl at the 2-position) have a higher activity mainly because of the hydrophobic interaction between the butyl and main chain of Leu298, which is also consistent with the result of the QSAR model. Compared with the quinazoline moiety, the phenol ring seems much more flexible. The phenol ring is both hydrophobic and polar. Hydroxyl hydrogen of the phenol form hydrogen bond with His475 and Leu476, the benzene part of the phenol, most of which is exposed to the solvent, may form a hydrophobic interaction contact with the side chain of Met336.

Table 3 Summary of CoMFA results for 3D-QSAR and 3D-QSSR models

PLS statistic	3D-QSAR	3D-QSSR
q^2	0.636	0.520
N	6	6
R^2	0.967	0.958
r^2_{pred}	0.62	0.70
F	150.758	96.075
S	0.200	0.147
Field contribution		
Steric / Electrostatic	0.586/0.414	0.631/0.369

Note: q^2 is the cross-validated squared coefficient.
N is the optimal number of components.
R^2 is the non cross-validated squared coefficient.
r^2_{pred} is the external validation value
S is the standard error of estimation.
F is the F-test value.

Hydrogen-bonding interactions Following a similar binding pattern, hydrogen bonding is an important characteristic [20] of the interaction between 3-Arylquinazolinethione derivatives and ERβ. There are mainly three hydrogen bonds formed between the 3-Arylquinazolinethione derivatives and some residues in ERβ. It can be seen clearly from Fig. 4 that the hydroxyl hydrogen of the phenol ring of compound 1aar acts as a donor to form a hydrogen bond with the nitrogen of the imidazole of His475. Similarly, as a hydrogen donor, the hydroxyl hydrogen of the phenol ring also forms a hydrogen bond with the backbone carbon of the Leu476. The third hydrogen bond is formed between the hydroxyl hydrogen of the quinazolinone of compound 1aar, as a donor, and the sides chain oxygen of Leu339. Acting as an‘anchor’, the hydrogen-bonding intensely determined the 3D space position of the quinazolinone ring and the phenol ring in the binding pocket, and facilitated the hydrophobic interaction of the aromatic and heterocyclic rings with the side chain of Leu298, Met336, Arg346, Gly472, Leu339, Leu301, Met473, and Ile373.

The compounds (1ba~1bf), having thiocarbonly form additional weak hydrogen bonds with the resides of ERβ besides the three main hydrogen bonds described above. The

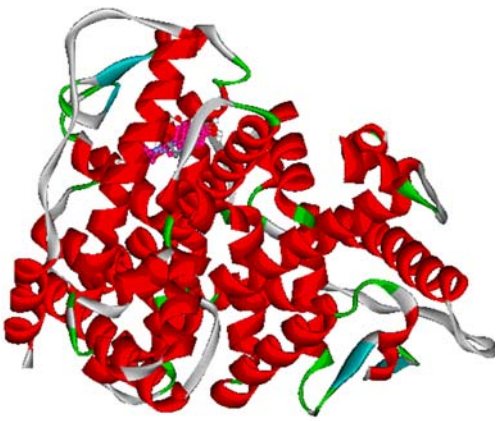


Fig. 2 Structures of ERβ binding with compounds 1aar, 1ax and 1aag obtained from molecular docking

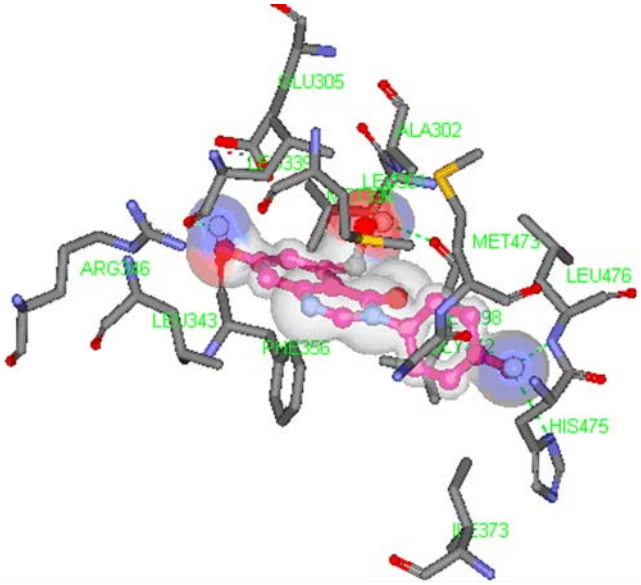


Fig. 3 The interacting mode of compound 1aar with ERβ. The inhibitors and the important residues for inhibitor-protein interaction are represented by ball-and-stick models, respectively. The green dashed lines denote the hydrogen bonds

mercapto, as an acceptor bonding with Leu298, may affect the activity and selectivity of the compounds 1ba~1bf.

3D-QSAR and -QSSR models

CoMFA model CoMFA was developed to model the protein-ligand interactions based on standard steric and electrostatic molecular fields. Despite being unable to describe all of the binding forces, CoMFA is still a widely used tool for QSAR analysis at 3D level. The major objective of CoMFA analysis is to find the best predictive model within the system. PLS analysis results for the pIC50 (ERβ) values are listed in Table 3. For convenience, the models obtained for the ERβ inhibitory activity and for the selectivity between ERβ and ERα were represented by pIC50 (ERβ) and logIC50 (ERα/β) models, respectively. As summarized in Table 3, a CoMFA model with a cross-validation q^2 value of 0.636 for six components was obtained for the inhibitory activity characterized by pIC50 (ERβ), whereas a model with the q^2 value of 0.520 for six components was obtained for the selectivity characterized by logIC50 (ERα/β). The non-cross-validated PLS analyses were repeated with the optimum number of components, as determined by the cross-validation analysis, to give R^2=0.967 and R^2=0.958 for the pIC50 (ERβ) and logIC50 (ERα/β) models, respectively. These correlation coefficients suggest that both the pIC50 (ERβ) and logIC50 (ERα/β) models are reliable and accurate. The inhibitory activity values predicted for these compounds are listed in Table 2 and depicted in Fig. 5a and b.

Validation of the 3D-QSAR and -QSSR models Consideration of q^2 is the traditional way of validating PLS models, but recent findings indicate that the use of q^2 alone may be advantageous in cases where data sets are small as a result of the loss of information associated with removing compounds to form a test set [21]. However, other authors have shown that q^2 is a useful but not sufficient criterion for model validation and have recommended the use of external test sets (r^2_{pred})for the estimation of predictive ability [22]. Predictive r^2_{pred} values were calculated using the following equation:

$$r^2_{pred} = 1 - (PRSS/SD).$$

Where PRESS is the sum of the squared differences between the observed and predicted activities and SD is the sum of squared differences between the measured activities of the test set and the average measured activity of the training set. Therefore, to find highly predictive models, r^2_{pred} and q^2 were used as the primary criteria for selecting optimal models, with r^2_{pred} given preference. So these statistical results confirmed the predictive capacity of the resultant CoMFA models. The predictive values of the holdout test compounds were listed in Table 2 and depicted in Fig. 3, which shows the correlation of experimental values versus the predicted values for compounds both in

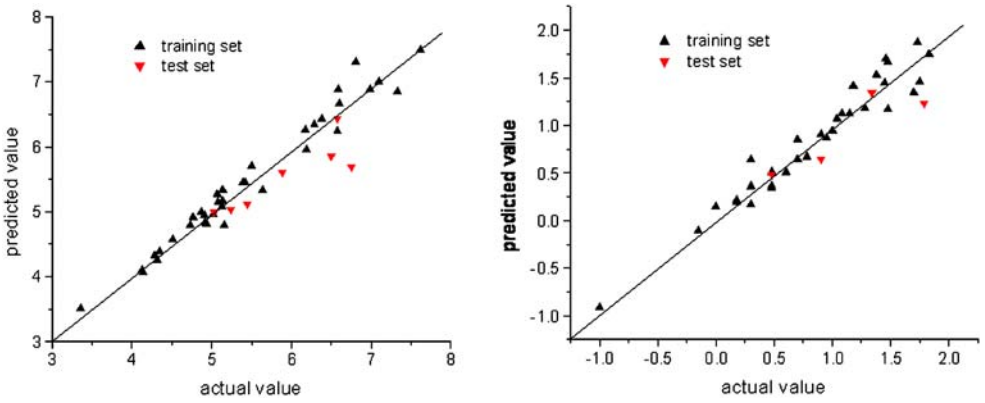


Fig. 4 Plots of the predicted pIC50(ERα/β) versus observed pIC50(ERα/β) (left) and predicted logIC50(ERβ) versus observed logIC50(ERβ) (right)

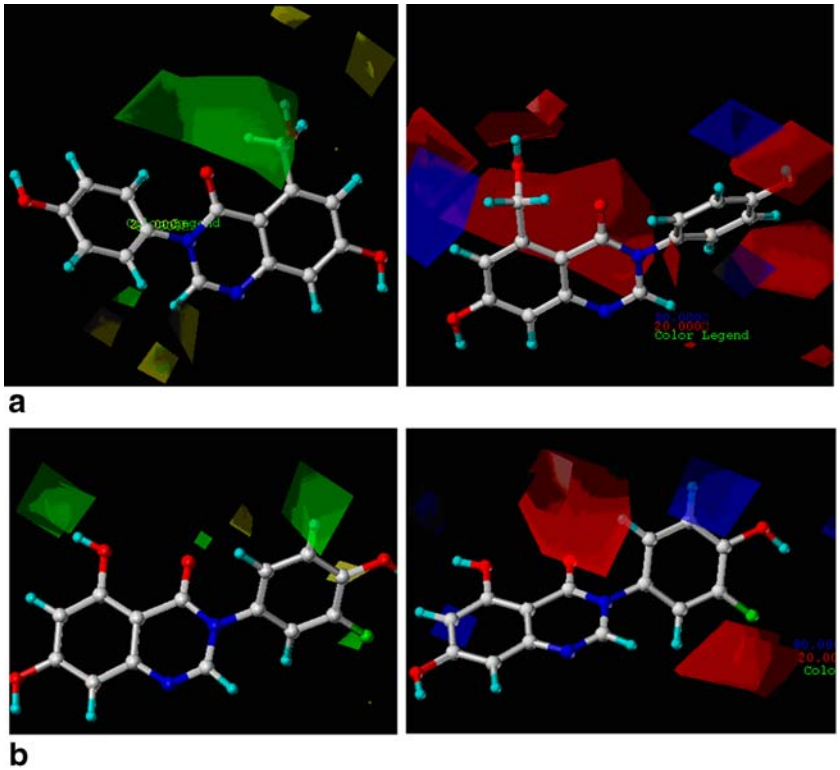


Fig. 5 Contour maps of the steric and electrostatic field of 2 models. **a)** Contour maps of pIC50(ERβ) models for steric field (left) and electrostatic field (right) as obtained by CoMFA. Green isopleths enclose areas where steric bulk will enhance affinity. Yellow contours highlighted areas that should be kept unoccupied, otherwise affinity will decrease. Red isopleths enclose areas where an increase of negative charge will enhance affinity, whereas in blue-contoured areas increase of positive charge is favorable for binding properties. These maps are demonstrated by the highly active compound 1aar. **b)** Contour maps of logIC50(ERα/β) models for steric field (left) and electrostatic field (right) as obtained by CoMFA. Green isopleths enclose areas where steric bulk will enhance selectivity. Yellow contours highlighted areas that should be kept unoccupied, otherwise selectivity will decrease. Red isopleths enclose areas where an increase of negative charge will enhance selectivity, whereas in blue-contoured areas increase of positive charge is favorable for binding properties. These maps are demonstrated by the highly active compound 1aar

the training and test sets. Results showed that predicted values of the test compounds are well in agreement with experimental values, as shown in Fig. 4.

Analysis of the 3D contour maps of the best model

The isocontour diagrams of the field contributions ('stdev*-coeff') of different properties obtained by CoMFA analyses are illustrated together with exemplary ligands.

Model for pIC50 (ERβ) Figure 5a shows the steric and electrostatic fields of compound 1aar based on the CoMFA model of pIC50. The green contours characterize the regions where bulky substiuents would increase the biological activity, whereas yellow contours indicate regions where steric bulk would not be tolerated. The green contours located at the 5-position of the compound 1aar, indicated that the bulky subsituents would be favorable, this can explain the activities of compound 1aan (a methoxy at the 5-position) and 1aaj (a methyl at the 5-position) are about 200 times higher than compound 1aal (a nitril at the 5-position). However, there is a bulky substituent (a phenyl) located at the 5-position of compound 1aaq with lower activity, the reason is that the space of binding site for the group at 5-position is limited, which was observed from the crystal structure. That is to say, adding much more bulky groups at the 5-position may bring steric clash of these inhibitors with Ala302 and Leu301 of ERβ, or the substituted more bulky groups may be exposed to solvent. At the 2-position, there are many relatively small yellow regions. So 1aag (a phenol at 2-position), 1aad (a chlorine at 2-position), and 1aab (a mercapto at 2-position) have lower activities. Nevertheless, it is interesting to note that the binding affinity improved as the length of the alkyl chain increased in size (compare 1ax with 1ay). However, this trend was not observed in the transactivation assay [8]. Results showed that there might exist other factors such as the space of the activity site and solvent effect influencing the overall activities. The above phenomenon has been explained by the all docking result above.

The blue and red contours depict the favorable sites for electropositive and electronegative groups, respectively. The large red contour encircling the quinazolinone ring

and located at 2', 3' and 4'-positions indicate positively charged substituent for these positions were unfavorable, which appears the reason of lower activity of compound 1ai (a methyl at 3'-position) comparing with the compounds 1ae (a fluoro group at the 3'-position), 1ag (fluoro group at the 2'-position), and 1ah (chlorine at 2'-position). The blue contours located at the 5-position. So the compound 1aam (a trifluoromethanesulfonate at the 5-position), 1aap (a methyloxy carbonyl at the 5-position) and 1aaq (a phenyl at the 5-position) have lower activities. It can be noticed that compounds 1ba~1bf with thiocarbonyl have higher activity than the other compounds. It is postulated that this improvement in binding affinity is due to the increased van der Waals contact between the protein and the sulfur. The most potent and selective compounds from the previous SAR studies were thionylated [8]. This "thio-effect" may also result from the change of the orbital hybridization of the heteroatom and therefore a change of the availability of the molecule metabolism [23].

Model for logIC50 (ERα/β) Figure 5b clearly indicates that the steric and electrostatic fields properties of 3-Arylquinazolinethione derivatives based on the CoMFA model of pIC50. Compound 1ae has a highly selectivity of ERβ against ERα, inserted its fluoro group at 3'-position into a green area. It suggests that this moiety is an important recognition element for binding with ERβ whereas the corresponding pocket in ERα is relatively limited. Furthermore, the combined information obtained from both Fig. 5a and b suggests that the compounds at 5-position with bulky group have high selectivity and activity. The result explains why the compound 1aan (an enthanel at the 5-position), 1aaj (a methyl at the 5-position), and 1aas (a prop-1-en-2-yl group at the 5-position) have high selectivity and activity.

The combined information obtained from both Fig. 5a and b suggests that the compounds having electronegative group at the 2' and 3'-positions increase not only the activity, but also the selectivity. So the compounds 1ah (a chlorine group at the 2'-position) and 1bb (a fluoro group at the 3'-position) have relatively high activity and selectivity. Groups with partial negative charge at 4-position could also help to increase the activity and selectivity, which can explain partially the relatively higher activity and selectivity of compounds 1ba~1bf compared to others. However, the lower selectivity of compound 1bc may be due to the electropositive methyl group located at the 3'-position. The blue contours mainly located at the 5-position, so the high selectivities of compounds 1aan (a methoxy at the 5-position), 1aaj (a methyl at the 5-position), and 1aas (a prop-1-en-2-yl group at the 5-position) were observed.

QSAR models and the X-ray crystal structure of ERβ

Docking results in this work indicated that the 3-Arylquinazolinethione derivatives are of an ideal length for forming tight hydrogen bonds between the 4'-hydroxyl and Glu305 and Arg346 at one end of the pocket and between the 7-hydroxyl and His475 at the other end [8]. Two common hydrogen bonds were formed. One is between the nitrogen of the imidazole of the His475 or main chain nitrogen of the Leu476 and the hydroxyl hydrogen of the phenol ring of the ligand, and the other is between the backbone oxygen of the Glu305 or sides chain oxygen of Leu339 and the hydroxyl hydrogen of the ligand revealed by the X-ray crystal structures. This result is consistent with the docking structure obtained in this work.

The QSAR results revealed that at the 5-position, bulky group was favorable for the activity and selectivity, and at 2-position bulky group was unfavorable. The docking results demonstrated that 5-position was a little far away from Ala302 and Leu298, so, only the larger group could form hydrogen bonds with Ala302 or Leu298. And bulky groups at the 2-position could alter the preferred dihedral angle of the distal phenyl moiety and thus affect the binding affinity. This phenomenon observed from compound 1aag, was consistent with the QSAR results and docking results. Both of the QSAR models and docking results in this work suggested that positions C2 and C5 of the quinazoline were oriented to regions where ERα and ERβ differ in the ligand binding domain. Modifications of substituents at C2 and C5 of the quinazoline may offer the opportunity to improve ERβ/ ERα selectivity.

Thus, both the QSAR analysis and molecular docking consistently suggested that the introduction of a bulky group into the proper site increase the ERβ inhibitory activity and reduce the ERα inhibitory activity. In short, our QSAR results were in agreement with the X-ray crystal structures of PDE5 reported in literature.

Conclusions

Understanding intermolecular interactions of 3-Arylquinazolinethione derivatives with ERβ and ERα was achieved by performing molecular docking, 3D-QSAR, and 3D-QSSR analyses. The use of molecular conformations of the compounds derived from molecular docking led to satisfactory 3D-QSAR and 3D-QSSR models (with high cross-validation correlation coefficient q^2 and conventional correlation coefficient R^2 values) for predicting the inhibitory activity against ERβ and the selectivity against ERα. The high q^2 and R^2 values, along with further testing, indicated that the obtained QSAR and QSSR models were

valuable in predicting both the inhibitory activity and selectivity of 3-Arylquinazolinethione derivatives against these protein targets. A set of 3D contour plots drawn based on the 3D-QSAR and QSSR models revealed moderate bulky groups with positive-rich charges in 5-position and small bulk groups in 2-position can improve the inhibitory activity and selectivity by modifying structures of the compounds. Comparison of QSAR models with the X-ray crystal structures of ERβ indicated that 3-Arylquinazoline-thione derivatives have a similar binding mode as sildenafil and a binding site for ERβ and ERα selectivity was revealed by QSSR analysis. It can be concluded that both the steric and electrostatic factors should be considered appropriately for designing novel ERβ inhibitors with higher inhibitory activity and selectivity.

Acknowledgements This work is supported by the Scientific Research Common Program of Beijing Municipal Commission of Education. The authors are grateful to the kind help of Dr. Yuandong Hu and Miss Mei Zhou, Beijing Hongcam Software Technologies Co. Ltd.

References

1. Robinson-Rechavi M, Escriva GH, Laudet V (2003) J Cell Sci 116:585–586
2. Kuiper GGJM, Enmark E, Pelto-Huikko M, Nilsson S, Gustafsson J-A (1996) Proc Natl Acad Sci 93:5925–5930
3. Mosselman S, Polman J, Dijkema R (1996) FEBS Lett 392:49–53
4. Kuiper GGJM, Gustafsson J-A (1997) FEBS Lett 410:87–90
5. Kuiper GGJM, Carlsson B, Grandien K, Enmark E, Haeggblad J (1997) 138:863–870
6. Dechering K, Boersma Ch, Mosselman S (2000) Curr Med Chem 7:561–576
7. Subhendu M, Achintya S, Kunal R (2005) BMC Lett 15:957–961
8. Timur G, Chen Y, Rajasree G, Ma ZP, James RC (2006) J Med Chem 49:2440–2455
9. Yang GF, Lu HT, Xiong Y, Zhan CG (2006) BMC 14:1462–1473
10. Irwin RA, Menezes AT, Carlos AM (2006) STEROIDS 71:417–428
11. Pike AC, Brzozowski AM, Hubbard RE, Bonn T, Thorsell AG, Engstrom O, Ljunggren J, Gustafsson JA, Carlquist M (1999) Embo J. 18:4608–4618
12. Manual S, Louis MO (2005) Tripos Inc Sybyl 7.1
13. Goodsell DS, Olson AJ (1990) 8:195–202
14. Morris GM, Goodsell DS, Huey R, Olson AJ (1996) 10:293–304
15. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) J Comp Chem 19:1639–1662
16. Solis FJ, Wets RJ (1981) B. Maths Pera Res 6:19–30
17. Weiner SJ, Kollman PA, Case DA, Singh C, Ghio G, Alagona S, Profeta P, Weiner PJ (1984) Am Chem Soc 106:765–771
18. Hu WX, Yun LH (1992) Chin Chem Lett 3:271–280
19. Pauling P, Datta N, Ramsay W, Forster R (1980) Proc Natl Acad Sci USA 77:708–721
20. Liu GX, Zhang ZS, Luo XM, Shen JH, Liu H, Shen X, Chen KX, Jiang HL (2004) BMC 12:1447–1457
21. Hawkins DM, Basak SC, Mills D (2003) J Chem Inf Comput Sci 43:579–586
22. Golbraikh A, Tropsha A (2002) J Mol Graph Model 20:269–276
23. Parker MA, Lewicka DM, Lucaites VL, Nelson DL, Nichols DE (1998) J Med Chem 41:5148–5162