

Structural and chemical basis for enhanced affinity and potency for a large series of estrogen receptor ligands: 2D and 3D QSAR studies

Lívia de B. Salum, Igor Polikarpov, Adriano D. Andricopulo*

Laboratório de Química Medicinal e Computacional, Centro de Biotecnologia Molecular Estrutural, Instituto de Física de São Carlos, Universidade de São Paulo, Av. Trabalhador São-Carlense 400, 13560-970 São Carlos, SP, Brazil

Received 15 December 2006; received in revised form 31 January 2007; accepted 2 February 2007

Available online 8 February 2007

Abstract

The estrogen receptor (ER) is an important drug target for the development of novel therapeutic agents for the treatment of breast cancer. Progress towards the design of more potent and selective ER modulators requires the optimization of multiple ligand-receptor interactions. Comparative molecular field analyses (CoMFA) and hologram quantitative structure–activity relationships (HQSAR) were conducted on a large set of ER α modulators. Two training sets containing either 127 or 69 compounds were used to generate QSAR models for *in vitro* binding affinity and potency, respectively. Significant correlation coefficients (affinity models, CoMFA, $r^2 = 0.93$ and $q^2 = 0.79$; HQSAR, $r^2 = 0.92$ and $q^2 = 0.71$; potency models, CoMFA, $r^2 = 0.94$ and $q^2 = 0.72$; HQSAR, $r^2 = 0.92$ and $q^2 = 0.74$) were obtained, indicating the potential of the models for untested compounds. The generated models were validated using external test sets, and the predicted values were in good agreement with the experimental results. The final QSAR models as well as the information gathered from 3D contour maps should be useful for the design of novel ER α modulators having improved affinity and potency.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Estrogen receptor; Drug design; Breast cancer; CoMFA; HQSAR

1. Introduction

Estrogens are crucial endogenous hormones that modulate the development and homeostasis of a wide range of female target tissues, including reproductive tracts, breast and skeletal system [1–3]. Their physiological effects are mediated by the estrogen receptor (ER) subtypes, ER α and ER β , which are members of the nuclear receptor superfamily of ligand-modulated transcriptional factors [4–6]. The binding of the natural ligand estradiol or other ligands to ER triggers complex signaling networks, leading to the recruitment of coregulatory complexes and to the transcription or repression of specific genes.

The high flexibility of the ER ligand-binding cavity allows the binding of a series of structurally diverse small molecules, which can exhibit agonist and antagonist effects, depending on

intrinsic cellular differences in the recruitment of coregulators [1,2,7–11].

The ligand-induced conformational changes in the ER binding site and the receptor structure in general, particularly at the N-terminal helix 12 position and its conformation, are major determinants of either agonist or antagonist properties. The understanding of the molecular recognition process responsible for the pharmacological effects of ER is very complex and has been extensively discussed in the literature [9,12,13].

Given the established role of ER α on the etiology and pathophysiology of breast cancer, the second leading cause of cancer death in women, selective ER modulation is an important strategy in the treatment of this disease [14–16]. The ER modulators tamoxifen (generic) and fulvestrant (Faslodex[®]) are well-known drugs employed in cancer therapy. Since resistance and serious side effects have been reported, including endometrial changes, hot flashes and irregular menstrual periods, there is an increasing clinical need for new therapeutic agents [2,8,9,16–20].

* Corresponding author. Tel.: +55 16 3373 8095; fax: +55 16 3373 9881.

E-mail address: aandrico@ifsc.usp.br (A.D. Andricopulo).

The key issue in the design of new selective ER modulators is to explore the properties of the chemical structure in combination with its ability of inducing a pharmacological response as a consequence of receptor-binding. As part of our ongoing research program aimed at discovering new selective and potent ER antagonists, and in order to investigate the quantitative structure–activity relationships (QSAR) of a large series of ER α ligands [21–31], we have employed the comparative molecular field analysis (CoMFA) and the hologram QSAR (HQSAR) methods to generate predictive 3D and 2D QSAR models, respectively [32,33]. The identification of key structural features responsible for binding affinity and potency should be useful for the design of novel modulators having promise of utility in clinical medicine.

2. Methodology

2.1. Data sets

The data set of 127 ER α modulators used in the QSAR analyses was selected from the literature [21–31]. Table 1 shows the general chemical structures of the series of ER α modulators employed in the QSAR studies. The complete version of Table 1, including all chemical structures and corresponding biological data, is provided as supplementary information, as well as the chemical structures in SMILES format for the complete data set.

2.2. Computational approach

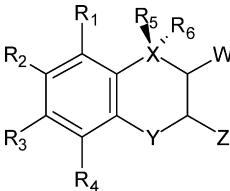
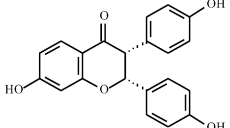
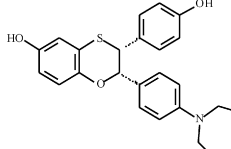
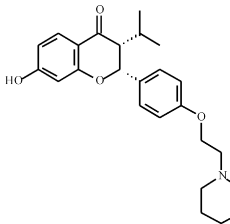
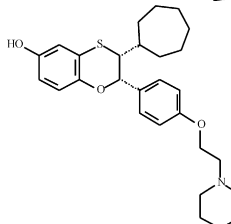
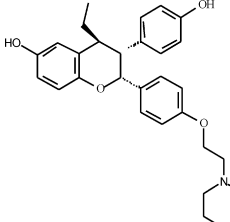
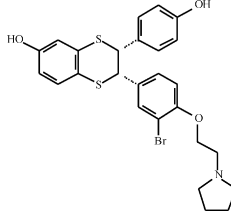
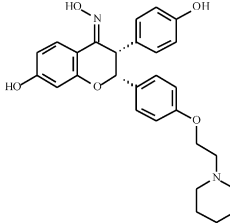
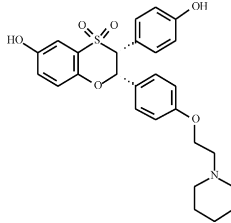
QSAR modeling analyses, calculations, and visualizations for CoMFA and HQSAR were performed using the SYBYL 7.2 package (Tripos Inc., St. Louis, USA) running on Red Hat Enterprise Linux workstations. The 3D structures of the modulators were constructed using standard geometric parameters of molecular modeling software package SYBYL 7.2. Each single optimized conformation of each molecule in the data set was energetically minimized employing Tripos force field and Gasteiger–Huckel charges. Ligands presenting tertiary amine function were modeled in the protonated state, since under physiological conditions these ligands are expected to be in this form.

2.3. Molecular modeling and structural alignment

Docking protocol as implemented in GOLD 3.0 (Cambridge Crystallographic Data Centre, Cambridge, UK) was employed to search the possible binding conformations of ligands into the ER α ligand-binding site. Crystallographic data for ER α in complex with (2*S*,3*R*)-2-(4-{2-[(3*R*,4*R*)-3,4-dimethylpyrrolidin-1-yl]ethoxy}phenyl)-3-(4-hydroxyphenyl)-2,3-dihydro-1,4-benzoxathiin-6-ol (compound **33**, Table 1) used in simulations were retrieved from the Protein Data Bank (PDB code: 1XP1). All docking simulations were performed with the X-ray coordinates of the estrogen receptor structure to which the ligand and water molecules were removed. Hydrogen atoms were added in standard geometry using the Biopolymer module

Table 1

Representative chemical structures and associated binding affinity values (pIC₅₀) of the series of ER α modulators employed in the QSAR studies

			
Compound	pIC ₅₀	Compound	pIC ₅₀
	6.31		7.54
	7.05		8.13
	7.85		8.40
	5.98		5.96

as implemented is Sybyl 7.2. His, Gln, and Asn residues were manually checked for possible flipped orientation, protonation, and tautomeric states with Pymol 0.99 (DeLano Scientific, San Carlos, USA) side-chain wizard script. The binding site was centered on C22 atom of the (2*S*,3*R*)-2-(4-{2-[(3*R*,4*R*)-3,4-dimethylpyrrolidin-1-yl]ethoxy}phenyl)-3-(4-hydroxyphenyl)-2,3-dihydro-1,4-benzoxathiin-6-ol and a radius of 10 Å was considered for the docking procedure. Default parameters and GOLDScore function were employed in all runs, and only the best ranked conformation of each modulator was considered for 3D QSAR studies [34,35,36]. The aligned data set is depicted in Fig. 1.

2.4. 3D QSAR: CoMFA models

To better understand and explore the contributions of electrostatic and steric fields in the binding affinity and potency of the ER α modulators, and in order to build predictive 3D



Fig. 1. Three-dimensional data set alignment for conformations generated by GOLD 2.1.

QSAR models, CoMFA studies were performed based on the molecular alignment described. The contributions of steric and electrostatic fields were calculated according to Lennard–Jones and Coulomb potentials, respectively [32]. The aligned training set of molecules was placed in a 3D grid box such that the entire set was included in it. CoMFA steric and electrostatic fields were generated at each grid point with Tripos force field using a sp^3 carbon atom probe carrying a + 1 net charge. The CoMFA grid spacing of 2.0 Å in the x , y , and z directions, and the grid region were generated by the CoMFA routine to encompass all molecules with an extension of 2.0 Å in each direction. CoMFA region focusing method was applied to increase the resolution of CoMFA models. The default value of 30 kcal mol⁻¹ was set as the maximum steric and electrostatic energy cutoff. All models were investigated using full cross-validated r^2 (q^2) partial least squares (PLS) leave-one-out (LOO) method, with CoMFA standard options for scaling of variables. Progressive scrambling method was applied to determine the sensitivity of the QSAR models to chance correlations.

2.5. 2D QSAR: HQSAR Models

Since the integrated information obtained from 2D and 3D methods is recognized as a valuable strategy in drug design [34,37,38], we have explored the 2D molecular features related to the biological properties of this series of ER ligands. Predictive 2D QSAR models were constructed using the HQSAR technique, which requires only 2D structures and biological activity values as input [33,34,39,40]. In this method, each molecule in the data set is broken down into several unique structural fragments, which are arranged to form a molecular hologram. HQSAR models can be affected by a number of parameters concerning hologram generation: hologram length,

fragment size, and fragment distinction. Several combinations of fragment distinction were considered during the QSAR modeling runs. Holograms were generated using distinct combinations of atoms (A), bonds (B), connections (C), hydrogen atoms (H), chirality (Ch), and donor and acceptor (DA) as fragment distinctions. The HQSAR analysis was performed by screening the 12 default series of hologram length values ranging from 53 to 401 bins using distinct fragment sizes. The patterns of fragment counts from the training set modulators were then related to the experimental biological parameters using PLS analysis. LOO cross-validation was applied to determine the number of components that yield optimally predictive model.

3. Results and discussion

3.1. Data set characterization

3D CoMFA and 2D HQSAR models were derived for 127 ER α ligands using two distinct data sets (Table 1). The first data set (data set 1) containing the whole set of 127 ER ligands was used to create the binding affinity (as measured by IC₅₀ values) models. From the original data set, 99 compounds were selected as members of the training set for model construction while the other 28 compounds were held out as members of the test set for external validation. The second data set (data set 2) containing a subset of 69 ER ligands was employed to create the potency (*in vitro* inhibition of MCF-7 cell growth) models. Data set 2 was divided as follows: 55 compounds in the training set whereas the other 14 in the test set (Table 1).

The three major chemical classes included in the data sets are flavanoids, dihydrobenzoxathiins and dihydrobenzodithiins. These chemical classes were selected both because they provide an appropriate structural diversity for QSAR modeling and because the availability of high quality biological data. Hierarchical cluster analysis performed with Tsrar 3D (Accelrys, San Diego, USA) was used as previously described to guide an appropriate compound selection. In both cases, training and test sets were selected in such a way that structurally diverse molecules possessing activities of a wide range were included in both sets [34,39]. Thus, the data sets are appropriate for the purposes of QSAR model development. The PLS (partial least squares) method was used for all 3D and 2D QSAR analyses. The predictive ability of the models was assessed by their q^2 .

3.2. CoMFA models

In CoMFA, relationships for a set of compounds are established between the biological activities and their steric and electrostatic properties. This approach is based on the assumption that changes in ligand-binding affinities are related to changes in molecular properties represented by molecular fields. The alignment rule is a crucial variable, affecting the outcome of the 3D statistical analysis. An ideal alignment should represent the ligand-binding conformations adopted in the receptor binding site. After the alignment process, each of the employed conformations was inspected individually and no conformational problem was revealed. The molecular

Table 2
COMFA results

Data set	Num	q^2	N	r^2	S.E.E.	F	Fraction	
							S	E
1	127	0.79	7	0.93	0.26	161	0.74	0.26
2	69	0.72	7	0.94	0.24	105	0.66	0.34

Data set 1, binding affinity; data set 2, potency; Num, number of compounds used in the modelling; q^2 , leave-one-out (LOO) cross-validated correlation coefficient; N , optimum number of components; r^2 , non-cross-validated correlation coefficient; S.E.E., standard error of estimate; F , F -test value; S , steric field; E , electrostatic field.

alignment shown in Fig. 1 was submitted to the CoMFA analyses, and the statistical results are presented in Table 2. Significant correlation coefficients were obtained. As it can be seen, a cross-validated correlation coefficient q^2 of 0.79 and a

conventional non-cross-validated correlation coefficient r^2 of 0.93 with a standard error of estimate (S.E.E.) of 0.26 were obtained for the data set 1 (binding affinity), while a q^2 of 0.72 and an r^2 of 0.94 with a S.E.E. of 0.24 were obtained for the data set 2 (potency). In both cases, no significant differences were obtained using the LMO (leave-many-out) method. Progressive scrambling of the data sets was carried out to check for possible chance correlations and test the stability of the models. The results of progressive scrambling further confirmed consistency of the models as defined by the critical slope, and optimum statistics for cSDEP and Q^{*2} obtained at the end of different runs. The region focusing was weighted by StDev*Coefficient values ranging from 0.3 to 1.5, and grid spacing ranging from 1.0 to 4.0. This strategy not only increased q^2 values during the process of model generation, but also resulted in the refinement of 3D contour maps.

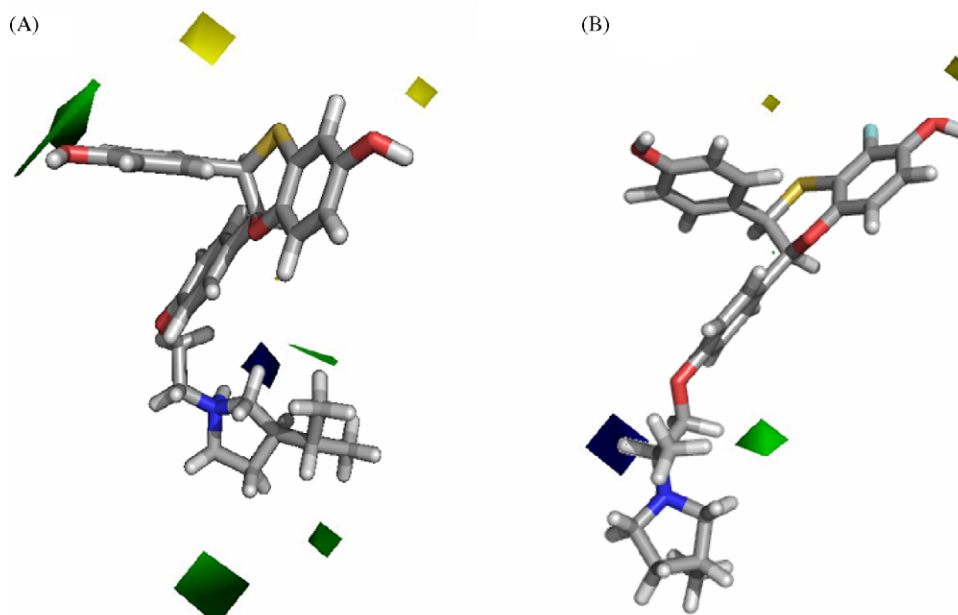
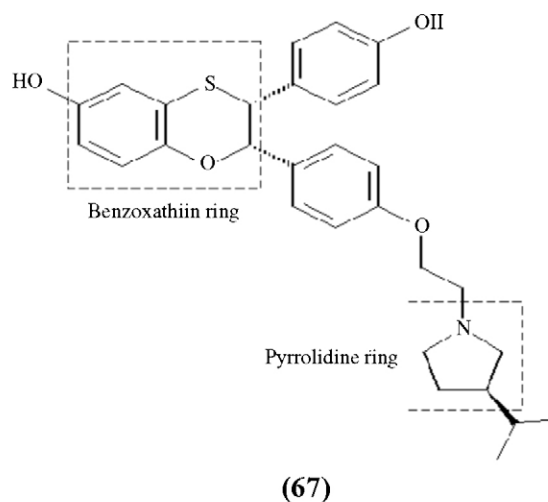


Fig. 2. Data set 1 CoMFA steric and electrostatic contour maps for the derivative with highest affinity, **67** (A). Corresponding maps for the data set 2 and the most potent compound **43** (B). Greater values of the affinity and potency are correlated with: more bulk near green; less bulk near yellow and more positive charge near blue.

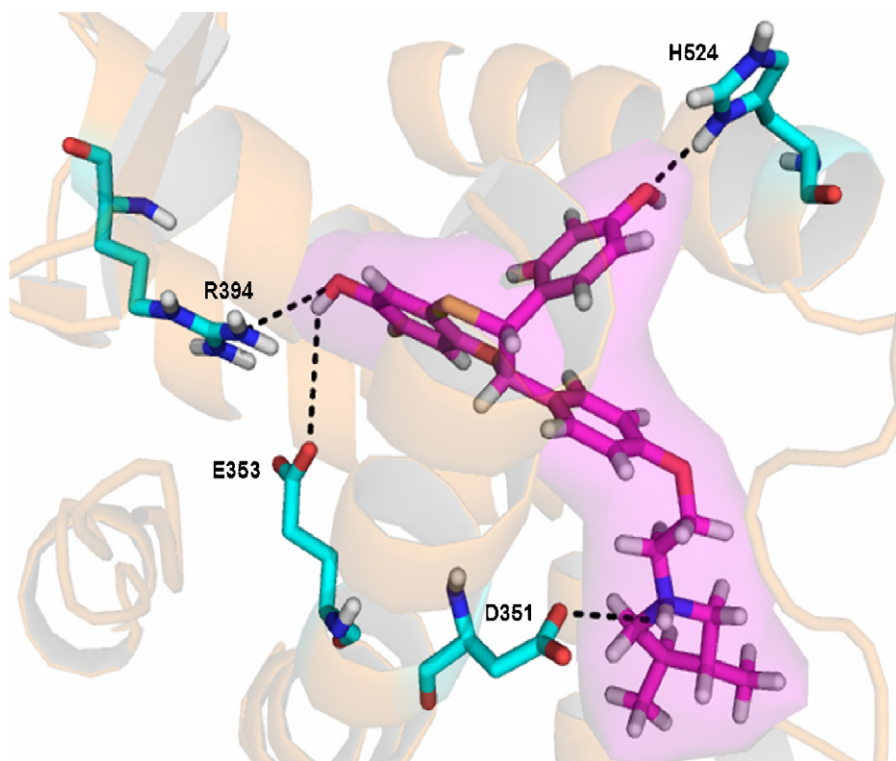


Fig. 3. Representation of the binding mode of the highest affinity modulator (**67**, colored in pink) into crystal structure of ER α . In blue, some important residues in the ligand-binding site.

CoMFA steric field descriptor for data set 1 explains 74% of the total variance, while the electrostatic descriptor accounts for the rest 26%. For data set 2, steric and electrostatic fields explain 66 and 34%, respectively. As expected, the steric field descriptors were more important to explain the total variance in both affinity and potency models [41–44], since the ER ligand-binding cavity possess hydrophobic features which must be fitted correctly by selective ligands, as confirmed by the structural hydrophobic core of estradiol, its natural endogenous ligand.

The data set 1 CoMFA steric and electrostatic fields contour maps for compound **67** (highest affinity of the series) are shown in Fig. 2A, whereas the corresponding maps for the data set 2 and compound **43** (the most potent of the series) are shown in Fig. 2B as PLS StDev*Coefficient plots. A comparison between affinity and potency electrostatic contour maps reveals that there is no significant difference, since the protonated nitrogen atom in the pyrrolidine ring is related to increasing affinity and potency. CoMFA steric fields are also similar in both 3D contour maps. Less steric bulky substituents attached to the benzoxathiin ring are positively related to the biological parameters [45,46]. The presence of either *para*- or *meta*-substituents is considered a positive contribution to affinity, but not to potency, as indicated by the green region in the affinity map around the phenyl ring. There is still room for bulkier groups in the pyrrolidine ring, which would increase both affinity and potency as suggested by the 3D contour maps. These chemical connections are especially attractive for molecular modification and further SAR studies aimed at increasing affinity and potency.

The 3D QSAR models generated are compatible with the 3D protein environment in the ER binding site, as shown in Fig. 3. The importance of the substituents in the benzoxathiin and phenyl rings corroborates with the estrogen receptor interaction sites HIS524, ARG394 and GLU353, as previously reported [9,17,41,47]. The maps indicate that the nitrogen pyrrolidine ring is important for the antagonist biocharacter. This parameter was similar to some crystallographic structures co-complexed with antagonist ligands and some studies, which designate the interactions between N of the heterocycle and ASP351 to be responsible for the antiestrogenic activity [48–51]. Our results also agree with previous studies, which reported that the optimal length of the link is represented by two carbons [28]. These results suggest that these models can be used in structure-based drug design and should be useful for the design of novel structurally related ER antagonists.

The predictive ability of the CoMFA models was assessed by predicting biological activities of external test sets. Prior to prediction, the test set compounds were processed identically to the training set compounds, as previously described. The external validation process can be considered the most valuable validation method, as these compounds were completely excluded during the training of the models. The results are listed in Tables 3 and 4, and the graphic results simultaneously displayed in Fig. 4. The predicted values fall close to the experimental pIC₅₀ values, not deviating by more than 0.42 log units. No outliers were detected in this series of ER ligands. The good agreement between actual and predicted pIC₅₀ values for test set compounds in both 3D QSAR models suggests that the constructed models are reliable and can be used for the design of modulators with improved properties.

Table 3

Experimental and predicted activities (pIC₅₀) for the ligand-binding assay with residual values for 28 test compounds

Test set compounds	Experimental	Predicted CoMFA	CoMFA residuals	Predicted HQSAR	HQSAR residuals
43	9.10	9.21	−0.11	9.05	0.05
44	8.77	8.60	0.17	8.62	0.15
45	8.36	8.46	−0.1	8.13	0.23
46	9.22	9.56	−0.34	9.25	−0.03
47	9.22	9.24	−0.02	9.14	0.08
48	9.30	9.41	−0.11	9.06	0.24
49	8.66	8.65	0.01	8.42	0.24
50	8.96	8.64	0.32	8.90	0.06
51	9.00	8.83	0.17	9.05	−0.05
52	8.89	8.90	−0.01	8.84	0.05
53	9.16	9.01	0.15	9.11	0.05
54	9.30	9.16	0.14	9.11	0.19
55	8.60	8.86	−0.26	8.59	0.01
56	8.85	9.14	−0.29	9.04	−0.19
57	9.16	9.19	−0.03	9.09	0.07
116	6.28	5.97	0.31	6.65	−0.37
117	6.98	6.77	0.21	6.85	0.13
118	6.83	6.99	−0.16	6.72	0.11
119	5.98	5.94	0.04	6.35	−0.37
120	8.51	8.93	−0.42	8.34	0.17
121	8.13	8.07	0.06	8.43	−0.3
122	7.46	7.29	0.17	7.33	0.13
123	8.16	8.55	−0.39	8.15	0.01
124	7.96	8.00	−0.04	8.31	−0.35
125	8.52	8.64	−0.12	8.60	−0.08
126	8.41	8.66	−0.25	8.54	−0.13
127	8.70	8.88	−0.18	8.51	0.19
128	8.27	8.59	−0.32	8.36	−0.09

CoMFA and HQSAR models.

3.3. HQSAR models

HQSAR investigations require the variation of parameters that specify the length, size and type of fragment that are to be encoded in the hologram. In the present paper, several combinations of these parameters were considered during the optimization of the models as previously described [34,39]. The generation of molecular fragments was initially carried out using the fragment size default (4–7) and several fragment distinctions, over the 12 default series of hologram lengths. The

statistical results from the PLS analyses for affinity (data set 1) and potency (data set 2) are presented in Tables 5 and 6, respectively. As it can be seen, the best statistical results concerning the affinity model were obtained using A/B/C/DA as distinction information ($q^2 = 0.71$, and $r^2 = 0.92$, with a S.E.E. of 0.27). On the other hand, the best statistical results for the potency model were obtained using A/B/DA as distinction information ($q^2 = 0.74$, and $r^2 = 0.92$, with a S.E.E. of 0.28).

The influence of different fragment sizes in the statistical parameters was further investigated for the best HQSAR

Table 4

Experimental and predicted activities (pIC₅₀) for the inhibition of MCF-7 cell growth assay with residual values for 14 test compounds

Test set compounds	Experimental	Predicted CoMFA	CoMFA residuals	Predicted HQSAR	HQSAR residuals
56	8.17	8.32	−0.15	8.33	−0.16
57	9.30	9.05	0.25	9.13	0.17
58	8.47	8.41	0.06	8.83	−0.36
59	7.55	7.64	−0.09	7.44	0.11
60	6.91	6.89	0.02	6.94	−0.03
61	6.95	6.94	0.01	6.70	0.25
62	8.59	8.50	0.09	8.71	−0.12
63	7.95	7.91	0.04	8.07	−0.12
64	9.05	8.88	0.17	8.94	0.11
65	9.00	8.84	0.16	8.94	0.06
66	8.80	8.91	−0.11	8.79	0.01
67	8.42	8.64	−0.22	8.66	−0.24
68	9.40	9.43	−0.03	9.43	−0.03
69	8.51	8.36	0.15	8.73	−0.22

CoMFA and HQSAR models.

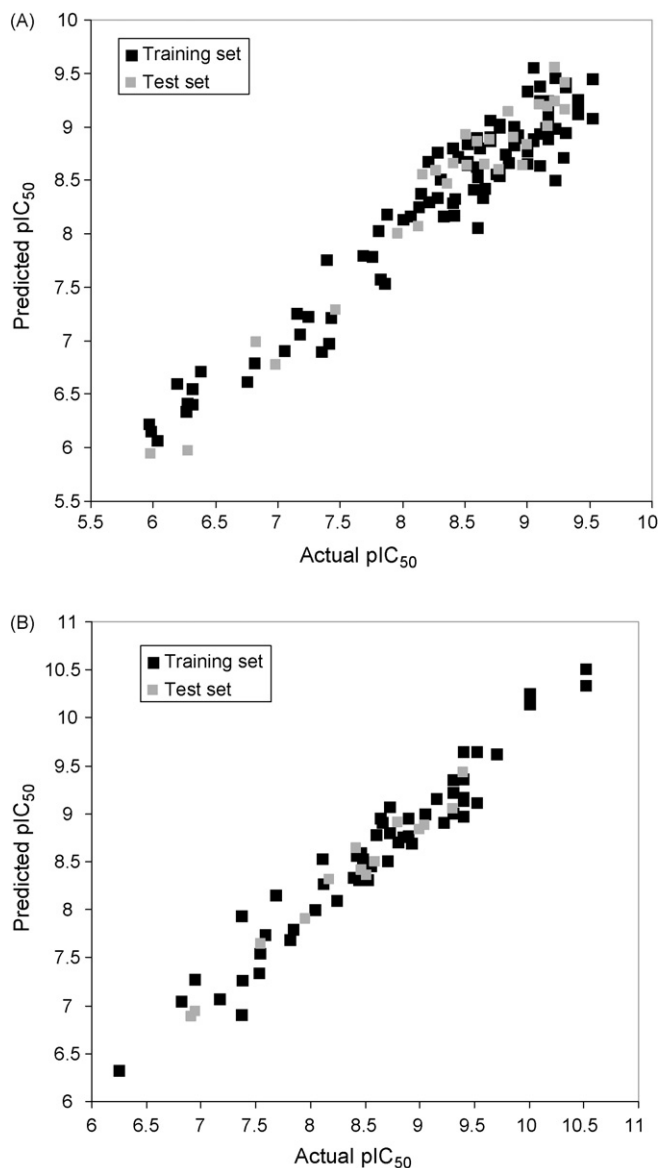


Fig. 4. Plot of predicted values of pIC_{50} vs. the corresponding experimental values for the training (black squares) and test (gray squares) set compounds for affinity (A) and potency models (B).

Table 5
HQSAR analysis for various fragment distinction on the key statistical parameters using fragment size (4–7)

Fragment distinction	Statistical parameters				
	q^2	r^2	S.E.E.	HL	N
A/B	0.70	0.89	0.32	199	6
A/B/C/H	0.63	0.78	0.44	83	6
A/B/C/H/DA	0.61	0.85	0.37	353	7
A/B/H	0.60	0.752	0.48	401	5
A/B/H/DA	0.57	0.77	0.45	307	6
A/B/C/DA	0.71	0.92	0.27	257	7

Model HQSAR for ligand-binding assay. q^2 , cross-validated correlation coefficient; r^2 , non-cross-validated correlation coefficient; S.E.E., non-cross-validated standard error; HL, hologram length; N , optimal number of components. Fragment distinction: A, atoms; B, bonds; C, connections; H, hydrogen atoms; DA, donor and acceptor.

Table 6

HQSAR analysis for various fragment distinction on the key statistical parameters using fragment size (4–7)

Fragment distinction	Statistical parameters				
	q^2	r^2	S.E.E.	HL	N
A/B	0.72	0.93	0.26	353	7
A/B/C/H	0.69	0.89	0.33	53	7
A/B/C/H/DA	0.71	0.82	0.40	53	4
A/B/H	0.71	0.91	0.29	199	7
A/B/H/DA	0.62	0.90	0.31	151	6
A/B/DA	0.74	0.92	0.28	97	7
A/B/H/DA	0.68	0.89	0.33	97	7

Model HQSAR for inhibition of MCF-7 cell growth assay. q^2 , cross-validated correlation coefficient; r^2 , non-cross-validated correlation coefficient; S.E.E., non-cross-validated standard error; HL, hologram length; N , optimal number of components. Fragment distinction: A, atoms; B, bonds; C, connections; H, hydrogen atoms; DA, donor and acceptor.

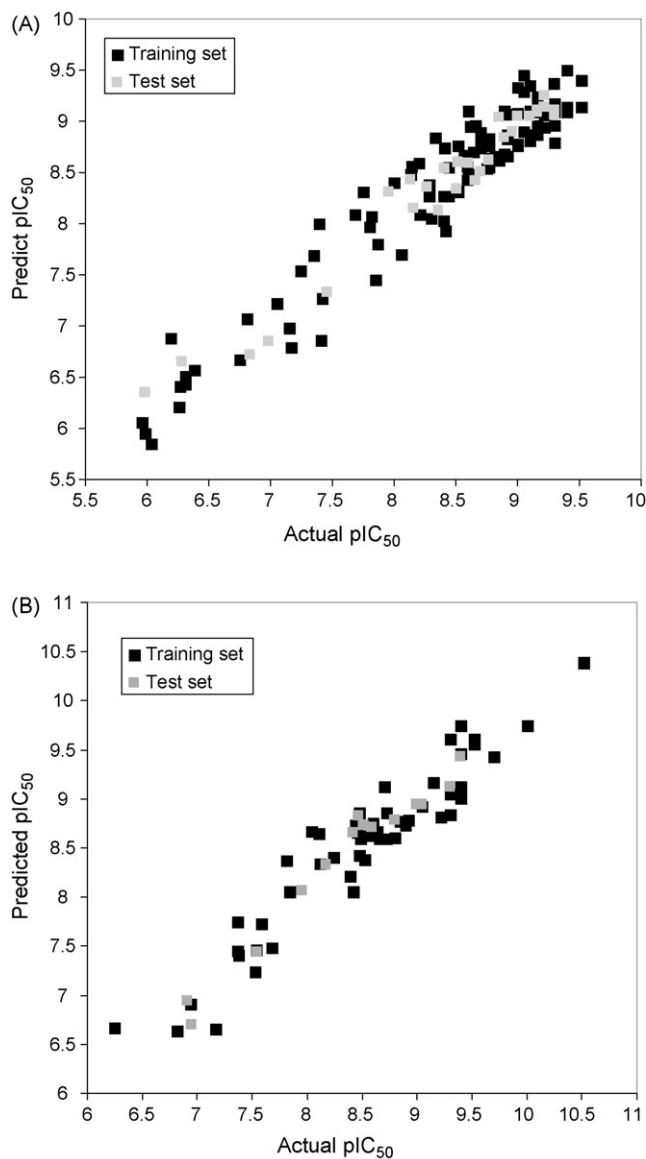


Fig. 5. Plot of predicted values of pIC_{50} vs. the corresponding experimental values for the training (black squares) and test (gray squares) set compounds for affinity (A) and potency models (B).

models generated previously. Fragment size parameters control the minimum and maximum lengths of fragments to be included in the hologram fingerprint. The statistical results for the different fragment sizes evaluated (2–5, 3–6, 4–7, 5–8 and 6–9) show no improvements in the HQSAR models. A measure of internal consistency is available in the form of q^2 . However, the most important test of a QSAR model is its ability to predict the property value for new compounds, not included in the training set. As the structure encoded in a 2D fingerprint is directly related to the biological activity of molecules within training set, HQSAR models are also able to predict the activity of new structurally related modulators from its fingerprint. Thus, the predictive power of the best affinity and potency HQSAR models derived using the training sets was assessed by predicting pIC_{50} values for test sets. The results of the external validation are listed in Tables 3 and 4. Again, the good agreement between experimental and predicted pIC_{50} values for the test set compounds indicates the robustness of the HQSAR models. The predicted values fall close to the experimental pIC_{50} values, deviating by no more than 0.37 log units in both models. The graphic results for the experimental versus predicted pIC_{50} values of both training and test sets are displayed in Fig. 5. From the low residual values, it can be seen that the affinity and potency HQSAR models obtained are highly reliable and can be used to predict the biological activity of novel compounds within these structural class. There are no outliers in both sets.

4. Conclusions

The QSAR models described herein exhibited both good internal and external consistency, showing substantial predictive power. The use of 3D structural information is useful in drug design and the CoMFA models generated in our studies are clearly stable and robust. Moreover, the 3D QSAR models are compatible with the 3D protein environment in the ER binding site, and should be useful for the design of novel structurally related ER α modulators. On the other hand, a powerful strategy is to study the integration of 2D and 3D methods in the development of predictive QSAR models. The HQSAR technique requires only 2D structures and biological activity as input, removing the necessity of conformer generation and molecular alignment which makes use of QSAR non-trivial in a number of settings. It is worth noting that we have employed the same training and test sets for all 3D and 2D QSAR analyses. The highly consistent results confirmed that investigations can be carried out concomitantly to search for synergies between 3D and 2D QSAR technologies. These important tools should be useful for the design of novel structurally related ER α antagonists with improved affinity and potency.

Acknowledgments

We gratefully acknowledge financial support from the Brazilian grant agencies FINEP (Research and Projects Financing), CNPq (The National Council for Scientific and

Technological Development) and FAPESP (The State of São Paulo Research Foundation).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2007.02.001.

References

- [1] S. Nilsson, J.-Å. Gustafsson, Biological role of estrogen and estrogen receptors, *Crit. Rev. Biochem. Mol. Biol.* 37 (2002) 1–28.
- [2] C.K. Osborne, R. Schiff, Estrogen-receptor biology: continuing progress and therapeutic implications, *J. Clin. Oncol.* 23 (2005) 1616–1622.
- [3] C.J. Gruber, W. Tschugguel, C. Schneeberger, J.C. Huber, Production and actions of estrogens, *N. Engl. J. Med.* 346 (2002) 340–352.
- [4] S. Green, P. Walter, V. Kumar, A. Krust, J.M. Bornert, P. Argos, P. Chambon, Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A, *Nature* 320 (1986) 134–139.
- [5] S. Mosselman, J. Polman, R. Dijkema, ER β : identification and characterization of a novel human estrogen receptor, *FEBS Lett.* 392 (1996) 49–53.
- [6] J.-Å. Gustafsson, What pharmacologists can learn from recent advances in oestrogen signaling, *Trends Pharmacol. Sci.* 24 (2003) 479–485.
- [7] H. Gronemeyer, J.-Å. Gustafsson, V. Laudet, Principles for modulation of the nuclear receptor superfamily, *Nat. Rev. Drug Discov.* 3 (2004) 950–964.
- [8] J.I. MacGregor, V.C. Jordan, Basic guide to the mechanisms of antiestrogen action, *Pharmacol. Rev.* 50 (1998) 151–196.
- [9] V.C. Jordan, Antiestrogens and selective estrogen receptor modulators as multifunctional medicines. 1. Receptor interactions, *J. Med. Chem.* 46 (2003) 883–908.
- [10] A.M. Brzozowski, A.C. Pike, Z. Dauter, R.E. Hubbard, T. Bonn, O. Engstrom, L. Ohman, G.L. Greene, J.-Å. Gustafsson, M. Carlquist, Molecular basis of agonism and antagonism in the oestrogen receptor, *Nature* 389 (1997) 753–758.
- [11] B.S. Katzenellenbogen, B. Bhardwaj, H. Fang, B.A. Ince, F. Pakdel, J.C. Reese, D.J. Schodin, C.K. Wrenn, Hormone binding and transcription activation by estrogen receptors: analyses using mammalian and yeast systems, *J. Steroid. Biochem. Mol. Biol.* 47 (1993) 39–48.
- [12] J.A. Katzenellenbogen, B.W. O'Malley, B.S. Katzenellenbogen, Tripartite steroid hormone receptor pharmacology: interaction with multiple effector sites as a basis for the cell- and promoter-specific action of these hormones, *Mol. Endocrinol.* 10 (1996) 119–131.
- [13] M. Dutertre, C.L. Smith, Molecular mechanisms of selective estrogen receptor modulators (SERM) action, *J. Pharmacol. Exp. Ther.* 295 (2000) 431–437.
- [14] S.S. Legha, S.K. Carter, Antiestrogens in the treatment of breast cancer, *Cancer Treat. Rev.* 3 (1976) 205–216.
- [15] L.J. Lerner, V.C. Jordan, The development of antiestrogens for the treatment of breast cancer, *Cancer Res.* 50 (1990) 4177–4189.
- [16] E.A. Ariazi, J.L. Ariazi, F. Cordera, V.C. Jordan, Estrogen receptors as therapeutic targets in breast cancer, *Curr. Top. Med. Chem.* 6 (2006) 181–202.
- [17] V.C. Jordan, Antiestrogens and selective estrogen receptor modulators as multifunctional medicines. 2. Clinical considerations and new agents, *J. Med. Chem.* 46 (2003) 1081–1111.
- [18] H.J. Tagnon, Antiestrogens in treatment of breast cancer, *Cancer* 39 (1977) 2959–2964.
- [19] A.E. Wakeling, M. Dukes, J. Bowler, A potent specific pure antiestrogen with clinical potential, *Cancer Res.* 51 (1991) 3867–3873.
- [20] Y. Shang, Molecular mechanisms of oestrogen and SERMs in endometrial carcinogenesis, *Nat. Rev. Cancer* 6 (2006) 360–368.
- [21] H.Y. Chen, K.D. Dykstra, E.T. Birzin, K. Frisch, W. Chan, Y.T. Yang, R.T. Mosley, F. DiNinno, S.P. Rohrer, J.M. Schaeffer, M.L. Hammond, Estrogen receptor ligands. Part 1. The discovery of flavanoids with subtype selectivity, *Bioorg. Med. Chem. Lett.* 14 (2004) 1417–1421.

- [22] S. Kim, J.Y. Wu, E.T. Birzin, K. Frisch, W. Chan, L.Y. Pai, Y.T. Yang, R.T. Mosley, P.M. Fitzgerald, N. Sharma, J. Dahllund, A.G. Thorsell, F. DiNinno, S.P. Rohrer, J.M. Schaeffer, M.L. Hammond, Estrogen Receptor Ligands. II. Discovery of benzoxathiins as potent, selective estrogen receptor α modulators, *J. Med. Chem.* 47 (2004) 2171–2175.
- [23] H.Y. Chen, S. Kim, J.Y. Wu, E.T. Birzin, W. Chan, Y.T. Yang, J. Dahllund, F. DiNinno, S.P. Rohrer, J.M. Schaeffer, M.L. Hammond, Estrogen receptor ligands. Part 3. The SAR of dihydrobenzoxathiin SERMs, *Bioorg. Med. Chem. Lett.* 14 (2004) 2551–2554.
- [24] S. Kim, J. Wu, H.Y. Chen, E.T. Birzin, W. Chan, Y.T. Yang, L. Colwell, S. Li, J. Dahllund, F. DiNinno, S.P. Rohrer, J.M. Schaeffer, M.L. Hammond, Estrogen receptor ligands. Part 4. The SAR of the syn-dihydrobenzoxathiin SERAMs, *Bioorg. Med. Chem. Lett.* 14 (2004) 2741–2745.
- [25] Q. Tan, E.T. Birzin, W. Chan, Y. Tien Yang, L.Y. Pai, E.C. Hayes, C.A. DaSilva, F. DiNinno, S.P. Rohrer, J.M. Schaeffer, M.L. Hammond, Estrogen receptor ligands. Part 5. The SAR of dihydrobenzoxathiins containing modified basic side chains, *Bioorg. Med. Chem. Lett.* 14 (2004) 3747–3751.
- [26] Q. Tan, E.T. Birzin, W. Chan, Y.T. Yang, L.Y. Pai, E.C. Hayes, C.A. DaSilva, F. DiNinno, S.P. Rohrer, J.M. Schaeffer, M.L. Hammond, Estrogen receptor ligands. Part 6. Synthesis and binding affinity of dihydrobenzodithiins, *Bioorg. Med. Chem. Lett.* 14 (2004) 3753–3755.
- [27] T.A. Blizzard, F. DiNinno, J.D. Morgan II, H.Y. Chen, J.Y. Wu, C. Gude, S. Kim, W. Chan, E.T. Birzin, Y. Tien Yang, L.Y. Pai, Z. Zhang, E.C. Hayes, C.A. DaSilva, W. Tang, S.P. Rohrer, J.M. Schaeffer, M.L. Hammond, Estrogen receptor ligands. Part 7. Dihydrobenzoxathiin SERAMs with bicyclic amine side chains, *Bioorg. Med. Chem. Lett.* 14 (2004) 3664–3861.
- [28] T.A. Blizzard, F. DiNinno, J.D. Morgan II, J.Y. Wu, H.Y. Chen, S. Kim, W. Chan, E.T. Birzin, Y.T. Yang, L.Y. Pai, Z. Zhang, E.C. Hayes, C.A. DaSilva, W. Tang, S.P. Rohrer, J.M. Schaeffer, M.L. Hammond, Estrogen receptor ligands. Part 8. Dihydrobenzoxathiin SERAMs with heteroatom-substituted side chains, *Bioorg. Med. Chem. Lett.* 14 (2004) 3865–3868.
- [29] T.A. Blizzard, F. Dininno, J.D. Morgan II, H.Y. Chen, J.Y. Wu, S. Kim, W. Chan, E.T. Birzin, Y.T. Yang, L.Y. Pai, P.M. Fitzgerald, N. Sharma, Y. Li, Z. Zhang, E.C. Hayes, C.A. DaSilva, W. Tang, S.P. Rohrer, J.M. Schaeffer, M.L. Hammond, Estrogen receptor ligands. Part 9. Dihydrobenzoxathiin SERAMs with alkyl substituted pyrrolidine side chains and linkers, *Bioorg. Med. Chem. Lett.* 15 (2005) 107–113.
- [30] Q. Tan, T.A. Blizzard, J.D. Morgan II, E.T. Birzin, W. Chan, Y.T. Yang, L.Y. Pai, E.C. Hayes, C.A. DaSilva, S. Warrier, J. Yudkovitz, H.A. Wilkinson, N. Sharma, P.M. Fitzgerald, S. Li, L. Colwell, J.E. Fisher, S. Adamski, A.A. Reszka, D. Kimmel, F. DiNinno, S.P. Rohrer, L.P. Freedman, J.M. Schaeffer, M.L. Hammond, Estrogen receptor ligands. Part 10. Chromanes: old scaffolds for new SERAMs, *Bioorg. Med. Chem. Lett.* 15 (2005) 1675–1681.
- [31] T.A. Blizzard, F. DiNinno, H.Y. Chen, S. Kim, J.Y. Wu, W. Chan, E.T. Birzin, Y.T. Yang, L.Y. Pai, E.C. Hayes, C.A. DaSilva, S.P. Rohrer, J.M. Schaeffer, M.L. Hammond, Estrogen receptor ligands. Part 13. Dihydrobenzoxathiin SERAMs with an optimized antagonist side chain, *Bioorg. Med. Chem. Lett.* 15 (2005) 3912–3916.
- [32] R.D. Cramer III, D.E. Patterson, J.D. Bunce, Recent advances in comparative molecular field analysis (CoMFA), *Prog. Clin. Biol. Res.* 291 (1989) 161–165.
- [33] W. Tong, D.R. Lowis, R. Perkins, Y. Chen, W.J. Welsh, D.W. Goddette, T.W. Heritage, D.M. Sheehan, Evaluation of quantitative structure–activity relationship methods for large-scale prediction of chemicals binding to the estrogen receptor, *J. Chem. Inf. Comput. Sci.* 38 (1998) 669–677.
- [34] M.S. Castilho, M.P. Postigo, C.B. de Paula, C.A. Montanari, G. Oliva, A.D. Andricopulo, Two- and three-dimensional quantitative structure–activity relationships for a series of purine nucleoside phosphorylase inhibitors, *Bioorg. Med. Chem.* 14 (2006) 516–527.
- [35] G. Jones, P. Willett, R.C. Glen, A.R. Leach, R.D. Taylor, Development and validation of a genetic algorithm for flexible docking, *J. Mol. Biol.* 267 (1997) 727–748.
- [36] M.L. Verdonk, J.C. Cole, M.J. Hartshorn, C.W. Murray, R.D. Taylor, Improved protein–ligand docking using GOLD, *Proteins* 52 (2003) 609–623.
- [37] W. Zhu, G. Chen, L. Hu, X. Luo, C. Gui, C. Luo, C.M. Puah, K. Chen, H. Jiang, QSAR analyses on ginkgolides and their analogues using CoMFA, CoMSIA, and HQSAR, *Bioorg. Med. Chem.* 13 (2005) 313–322.
- [38] M.A. Avery, M. Alvim-Gaston, C.R. Rodrigues, E.J. Barreiro, F.E. Cohen, Y.A. Sabnis, J.R. Woolfrey, Structure–activity relationships of the anti-malarial agent artemisinin. 6. The development of predictive in vitro potency models using CoMFA and HQSAR methodologies, *J. Med. Chem.* 45 (2002) 292–303.
- [39] K.M. Honorio, R.C. Garratt, A.D. Andricopulo, Hologram quantitative structure–activity relationships for a series of farnesoid \times receptor activators, *Bioorg. Med. Chem. Lett.* 15 (2005) 3119–3125.
- [40] M.R. Doddareddy, Y.J. Lee, Y.S. Cho, K.I. Choi, H.Y. Koh, A.N. Pae, Hologram quantitative structure activity relationship studies on 5-HT₆ antagonists, *Bioorg. Med. Chem.* 12 (2004) 3815–3824.
- [41] H. Gao, J.A. Katzenellenbogen, R. Garg, C. Hansch, Comparative QSAR analysis of estrogen receptor ligands, *Chem. Rev.* 99 (1999) 723–744.
- [42] J.H. Poupaert, D.M. Lambert, J. Vamecq, Y.J. Abul-Hajj, Molecular modeling studies on 11 β -aminoethoxyphenyl and 7 α -aminoethoxyphenyl estradiols. Evidence suggesting a common hydrophobic pocket in estrogen receptor, *Bioorg. Med. Chem. Lett.* 5 (1995) 839–842.
- [43] H. Fang, W. Tong, L.M. Shi, R. Blair, R. Perkins, W. Branham, B.S. Hass, Q. Xie, S.L. Dial, C.L. Moland, D.M. Sheehan, Structure–activity relationships for a large diverse set of natural, synthetic, and environmental estrogens, *Chem. Res. Toxicol.* 14 (2001) 280–294.
- [44] C.L. Waller, A comparative QSAR study using CoMFA, HQSAR, and FRED/SKEYS paradigms for estrogen receptor binding affinities of structurally diverse compounds, *J. Chem. Inf. Comput. Sci.* 44 (2004) 758–765.
- [45] C.L. Waller, T.I. Oprea, K. Chae, H.K. Park, K.S. Korach, S.C. Laws, T.E. Wiese, W.R. Kelce, L.E. Gray Jr., Ligand-based identification of environmental estrogens, *Chem. Res. Toxicol.* 9 (1996) 1240–1248.
- [46] I.R. Menezes, A. Leitao, C.A. Montanari, Three-dimensional models of non-steroidal ligands: a comparative molecular field analysis, *Steroids* 71 (2006) 417–428.
- [47] F. Minutolo, M. Antonello, S. Bertini, S. Rapposelli, A. Rossello, S. Sheng, K.E. Carlson, J.A. Katzenellenbogen, M. Macchia, Synthesis, binding affinity, and transcriptional activity of hydroxy- and methoxy-substituted 3,4-diarylsalicylaldoximes on estrogen receptors α and β , *Bioorg. Med. Chem.* 11 (2003) 1247–1257.
- [48] A.S. Levenson, V.C. Jordan, The key to the antiestrogenic mechanism of raloxifene is amino acid 351 (aspartate) in the estrogen receptor, *Cancer Res.* 58 (1998) 1872–1875.
- [49] J.H. Kim, M.H. Lee, B.J. Kim, J.H. Kim, S.J. Han, H.Y. Kim, M.R. Stallcup, Role of aspartate 351 in transactivation and active conformation of estrogen receptor α , *J. Mol. Endocrinol.* 35 (2005) 449–464.
- [50] C. Zhao, J. Abrams, D.F. Skafar, Targeted mutation of key residues at the start of helix 12 in the hER α ligand-binding domain identifies the role of hydrogen-bonding and hydrophobic interactions in the activity of the protein, *J. Steroid Biochem. Mol. Biol.* 98 (2006) 1–11.
- [51] G. Dayan, M. Lupien, A. Auger, S.I. Anghel, W. Rocha, S. Croisietiere, J.A. Katzenellenbogen, S. Mader, Tamoxifen and raloxifene differ in their functional interactions with aspartate 351 of estrogen receptor α , *Mol. Pharmacol.* 70 (2006) 579–588.