Pharmacophore Mapping of Selective Binding Affinity of Estrogen Modulators through Classical and Space Modeling Approaches: Exploration of Bridged-Cyclic Compounds with Diarylethylene Linkage

Subhendu Mukherjee,† Shuchi Nagar,† Sanchita Mullick, Arup Mukherjee, and Achintya Saha*

Department of Chemical Technology, University of Calcutta, 92, A.P.C. Road, Kolkata − 700009, India

Received September 27, 2006

Research on Selective Estrogen Receptor Modulators (SERMs) has been driven by interest in discovering target selective molecules. In view of such significance, the present work explored the pharmacophores of estrogen receptor (ER) subtypes specific binding affinities of diverse compounds belonging to the category of bridged bicyclic-1,1-diarylethylene derivatives. Implementing classical QSAR and *CATALYST* based space-modeling approaches, it has been explored that attachment of aryl ring systems to unsaturated linkages, availability of phenolic hydroxyl group, global hydrophobicity, and stereochemistry of certain functional groups might be important for governing the subtype specific estrogenic behavior of this group of compounds. Supplementing this deduction, critical interfeature distances between hydrogen bond acceptor, hydrophobic, and ring aromatic features along with steric influence are found to primarily influence the ER-subtypes specific binding of this series of compounds.

INTRODUCTION

Estrogens have demonstrated remarkable effectiveness in the deterrence and management of pre- and postmenopausal diseases.[1,2] The effects of estrogen are mediated by an intranuclear transcription factor called the estrogen receptor (ER),[3] of which two structurally similar subtypes (α and β) are presently known. The existence of these two subtypes provide a possible explanation for the tissue selectivity.[4] The predominant ER in the female reproductive tract and the mammary glands is ER_α, whereas ER_β is primarily present in the vascular endothelial cells, bone, and male prostate tissue. Both the subtypes bind the endogenous estrogen, 17β-estradiol (E_2), but illustrate different tissue expression and marked response to different agonists and antagonists.[5,6] X-ray crystal structures of receptor-E_2 complexes have demonstrated that the binding areas of ER_α and ER_β are varied by only two amino acids. Leu384/Met421 in ER_α are found to be equivalent to Met336/Ile373 in ER_β,[7] a classical feature that probably explains the lack of selectivity of E_2 in binding to the ER.[8] A group of structurally diverse nonsteroidal compounds binding to the ER has been investigated to produce estrogen agonist effects in some tissues, while antagonist effects in others. These classes of compounds are known as selective estrogen receptor modulators (SERMs).[9,10] As such, research efforts are being devoted toward developing SERMs that can maintain the benefits of estrogens while avoiding the risks.[11] Drug molecules like raloxifene and tamoxifen that have tissue selective ER agonist or antagonist properties are some agents that ensure a safer alternative to estrogen.[12,13] There continues to be a serious effort to understand the molecular basis for the tissues-specific effects of this group of compounds.[14] The 1,1-diaryl component is a chemical moiety common to many potential SERMs, such as tamoxifen, toremifene, and idoxifene.[15] Our group has also explored this feature to be prime pharmacophore signals for estrogen mediated bioactivities of a different group of compounds, viz., 1-trifluoromethyl-1,2,2-triphenylethylenes, bromotriphenylethylenes, and triphenylacrylonitriles through Quantitative Structure−Activity Relationship (QSAR) studies.[16−18]

QSARs are mathematical methodology, statistically validated, and mostly used to correlate experimental or calculated properties derived from chemical structures with biological activities. They also may be applied to predict the activity values of nonsynthesized compounds structurally related to the training set. Structurally diverse 60 environmental estrogens[19] are classified through QSAR studies. Since the different classes of environmental estrogens examined contain no common structural elements, 3D rototranslational descriptors[20] are employed to predict activity. Groups of aromatic, phenolic, and aliphatic tetrasubstituted pyrazole,[21] raloxifene,[22] and xenoestrogens[23] are modeled for studying the SAR for binding affinity to the ER, while tetrahydroisoquinolines[24] and diphenylnapthyl propylene[25] are modeled for ER selective subtypes binding.

With the advent of 3D molecular space modeling, a pharmacophore hypothesis can visualize the potential interaction between the ligand and the receptor. A pharmacophore is a set of functional group/fragment types in a spatial arrangement that represents the interaction made in common by a set of small molecular ligands with a protein receptor.[26] The pharmacophore concept is based on the kinds of interaction observed in molecular recognition, i.e., hydrogen bonding, charge, and hydrophobic interaction and alternatively can be used as a query in a 3D database search to identify new structural classes of potential lead compounds; and it can serve as a template for generating alignment for 3D QSAR analysis.[27] Two types of pharmacophore hypoth-

* Corresponding author e-mail: achintya_saha@yahoo.com.
† Both authors contributed equally to this work.

10.1021/ci600419s CCC: $37.00
Published on Web 03/01/2007

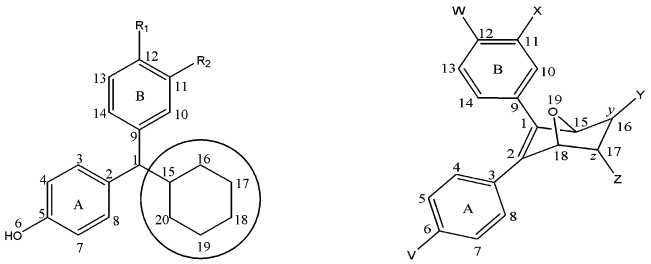


Figure 1. General structure of bridged-cyclic diaryl compounds.

eses are well established: receptor-based and receptor-independent pharmacophore. Receptor-based pharmacophore mapping of SERMs[25,28−30] are commonly employed, but nowadays receptor-independent pharmacophore mapping is growing interestingly for deriving bioactivities of diverse groups of compound, such as aromatase inhibitors,[31] serotonin inhibitors,[32] and antithormbin[33] agents and are now increasingly being handled by automated computational methods, such as CATALYST,[34] GASP, and DISCO which are commercially available programs.[35−37] Interestingly, there is no such receptor-independent pharmacophore hypothesis identified for SERMs. Consequently, the present work is taken up to study the 1,1-diaryl scaffold as a small ligand[38,39] with a view to deduce the active pharmacophore signals based on receptor-independent hypothesis, using the CATALYST program,[34] that can eventually aid in apprehending the tissue-specific effects of different compounds containing this unit. Due to the close structural similarity of these groups of compounds with that of many prospective SERMs,[15] it might be possible that some kind of pharmacodynamic similarity exists between these groups of compounds.

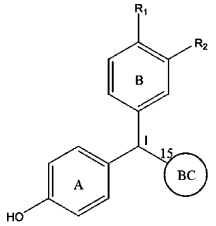
MATERIALS AND METHODS

In the present work, 14 compounds belonging to the categories of bridged bicyclic-1,1-diarylethylenes[39] (Figure 1a) and 20 compounds belonging to oxabicyclic-1,2-diarylethylenes[38] (Figure 1b) have been considered (Table 1) for modeling aspects through classical approach. The relative binding affinities (RBA) of these compounds to the ER subtypes (α and β) have been considered as the biological activity and implemented as logarithmic functions (pRBA) for modeling purpose. The primary objective of the work undertaken had been to generate relationships between the structure and corresponding activity through multiple linear regression (MLR) approach and deduce a pharmacophore map through receptor-independent space modeling technique. In the foremost section of the study, 3D structure of the molecules are energy minimized in *MOPAC* using the AM1 method to locate their global minima conformer, and subsequent calculation of different molecular properties, such as physiochemical, electronic, topological, and spatial and structural features are estimated for classical modeling. The partial charge is calculated using the Extended Huckel approach,[40] and *E*-state indices[41] of all the atoms are generated using a JAVA-based program,[42] while other descriptors are generated using Chem3D Pro 5.0[40] and Tsar 3.3.[43] MLR is performed using standard and forward stepwise regression methods.[44] QSAR model origination is accomplished by correlation analysis, and statistical parameters of the regression equation considered are as follows: r or R (correlation coefficient), EV (explained variance), F (variance ratio), df (degree of freedom), s (standard error of estimate), and AVRES (average of absolute values of calculated residuals). Leave-one-out (LOO) cross-validation[45] is performed that generated PRESS (predictive residual sum of squares), SDEP (standard deviation of error of predictions), Pres_{av} (average of absolute value of predicted residuals), and Q^2 (cross-validated variance).

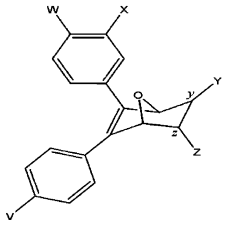
On the contrary, groups of compounds ($n = 34$) containing these bridged-cyclic diaryl scaffolds have been taken up for receptor-independent pharmacophore space modeling studies with regards to binding affinities at the ER_α and ER_β. The chemical features optimized for exploring the spatial pharmacophore maps of these groups of compound are hydrogen bond (HB) acceptor (a) and donor (d), hydrophobic (p), and ring aromatic (r). The pharmacophore models (hypothesis), generated by CATALYST,[34] consist of an array of features necessary for bioactivity of the ligands arranged in 3D space that can explain the variance in activity of the molecules w.r.t. geometric localization of the chemical features present in them. To be retrieved as a hit, a candidate ligand must possess appropriate functional groups which can simultaneously reside within the respective tolerance spheres of the pharmacophoric features. Each feature is associated with a weight (a measure of its proposed importance to the pharmacophore as a whole), and the better the overall superimposition of functional groups of the molecule to the appropriate features of the pharmacophore, the higher the score of the fit.[46]

The different control parameters employed for hypothesis generation (called a HypoGen process) are spacing, uncertainty, and weight variation. Spacing is a parameter representing the minimum interfeatures distance that may be allowed in the resulting hypotheses. In the generated *hypothesis*, each feature signifies some degree of magnitude of the compound's activity. The level to which this magnitude is explored by the *hypothesis* generator is controlled by the weight variation parameter. This is varied in some cases from 1 to 2. In other cases, the default value of 2 is generally considered. The uncertainty parameter reflects the error of prediction and denotes the standard deviation of a prediction error factor called the error cost. A default value of 3 is considered as the uncertainty parameter in the present work. While generating *hypothesis*, a total cost function is minimized comprising of three terms, viz., weight cost, error cost, and configuration cost. Weight cost is a value that increases as the weight variation in the model deviates from an ideal value of 2. The deviation between the estimated activities of the training set and their experimentally determined values is the error cost. A fixed cost (ideal *hypothesis* cost) depends on the complexity of the *hypothesis* space being optimized and is denoted as the configuration cost. The configuration cost is equal to the entropy of *hypothesis* space. The *CATALYST* program also calculates the cost of a null *hypothesis* that assumes no relationship in the data, and the experimental activities are normally distributed about their mean. Accordingly, the greater the difference between the total and the null costs, it is more likely that the *hypothesis* does not reflect a chance correlation. The minimum difference between the total and null

Table 1. Structural Features of Bridged-Cyclic Scaffolds

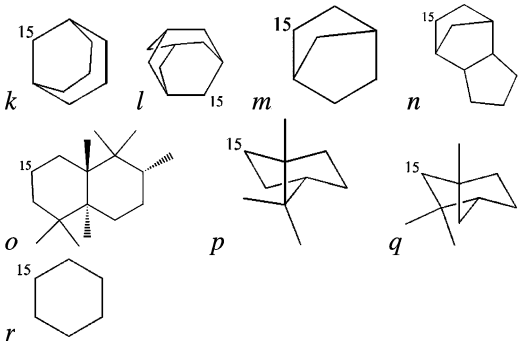


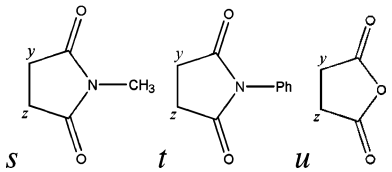
a



b

compd no.	substituents and ring systems				compd no.	substituents					
	R_1	R_2	BC[a]	C_1-C_{15}[b]		V	W	X	Y	Z	*y*−*z*
1	OH	H	*k*	U	**15**	OH	OH	H	SO_3Ph	H	S
2	F	H	*k*	U	**16**	OH	OH	H	SO_3Ph (α)	H	S
3	H	H	*k*	U	**17**	OH	OH	H	SO_2Ph	H	S
4	OH	OH	*k*	U	**18**	OH	OH	H	$SO_2CH_2CH_3$	H	S
5	H	OH	*k*	U	**19**	OH	OH	H	CO_2Et	CO_2Et	S
6	OH	H	*l*	U	**20**	OH	OH	H	−*s*−		S
7	OH	H	*m*	U	**21**	OH	OH	H	−*t*−		S
8	OH	H	*n*	U	**22**	OH	H	OH	SO_3Ph (mix.)	H	S
9	OH	H	*o*	U	**23**	OH	H	OH	SO_2Ph	H	S
10	OH	H	*p*	U	**24**	OH	OMe	H	SO_3Ph (mix.)	H	S
11	OH	H	*q*	U	**25**	OH	OMe	H	SO_2Ph	H	S
12	OH	H	*k*	S	**26**	H	H	H	SO_3Ph	H	S
13	OH	H	*r*	U	**27**	H	H	H	SO_2Ph	H	S
14	OH	H	*r*	S	**28**	H	H	H	−*u*−		S
					29	OH	OH	H	SO_3-Ph−OMe	H	S
bridged cycles[a]					**30**	H	H	H	SO_3-Ph−OH	H	S
					31	H	H	H	SO_3-Ph−OH (α)	H	S
					31	OH	OH	H	SO_3-Ph−OH	H	S
					32	H	H	H	CO_2Me	Ph	U
					33	H	H	H	CO_2Me	Ph−OH	U





[a] Bridged cycles. [b] See C_{15} in structures *k*−*r*. S: saturated; U: unsaturated; Y-substituent not mentioned denotes β configuration.

costs is taken as 60 bits for a *hypothesis* optimization. The generated *hypothesis* is further validated to nullify overprediction of bioactivities for inactive compounds through a process known as Hyporefine. In this process, the steric interactions of compounds are considered in the *hypothesis* generation, and if steric properties are crucial for bioactivities, then these are portrayed in the validated (refined) *hypothesis*. The quality of *hypothesis* generated is adjudged through a cross-validation technique using CatScramble.[34] The validation procedure is based on Fischer's randomization test,[47] where the biological activity data are randomized within a fixed chemical data set and the HypoGen process is initiated. By logic, the *hypothesis* generated prior to scrambling should be better to attest for a good pharmacophore model.

In the present work, the number of conformer generation is limited to a maximum of 250 using the 'best conformer generation' method with 20 kcal/mol of energy *cutoff*. The spacing is varied from 60 to 250 pm. The biological activities are expressed as RBA to the ER subtypes. The HypoGen algorithm is forced to find pharmacophore models that contain at least one and at most two of all the features.

RESULTS

Classical Modeling. A structure−activity relationship has been drawn, investigating physiochemical (partition coefficient, hydrophobicity, steric, and moments), electronic (atomic and partial charge functions, orbital energies), and electrotopological (*E*-state indices) features of molecular architecture for characterization of unique pharmacophore features. In all the regressional models, the 95% confidence intervals are shown in parentheses, and the *F*-values are significant at the 99% confidence level. The regression constants for all relations are significant at 95% (except superscripted with #).

(a) Bridged-Cyclic Diarylethylenes. While modeling RBA of bridged-cyclic diarylethylene derivatives (Figure 1a), the best single variate model for binding to the α-subtype could be developed with the charge function of atom C_1 (Ch_1) that explained 37.097% variance in activity, and the statistical quality of the relation is estimated to be

$$r = 0.648; \quad r^2 = 0.419; \quad s = 0.463; \quad n = 14$$

Table 2. Statistical Qualities of Best QSAR Relations for Binding Affinity of Bridged-Cyclic Scaffolds

		correlation statistics						prediction statistics				
eq no.	n	R	R^2	EV (%)	F	df	s	AVRES	PRESS	SDEP	$Pres_{av}$	Q^2
1	14	0.928	0.862	80.052	14.042	4.9	0.261	0.184	1.595	0.340	0.303	0.640
2	14	0.944	0.891	84.315	18.470	4.9	0.241	0.156	1.270	0.301	0.253	0.740
3	14	0.919	0.845	77.602	12.260	4.9	0.132	0.085	0.460	0.181	0.143	0.545
4	20	0.942	0.886	86.520	41.660	3.16	0.360	0.254	3.047	0.390	0.315	0.830
5	20	0.946	0.896	86.770	32.140	4.15	0.180	0.120	0.930	0.215	0.154	0.800
6	20	0.984	0.968	96.016	121.490	4.16	0.182	0.129	0.934	0.216	0.168	0.800
7	20	0.911	0.830	79.769	25.971	3.16	0.357	0.260	2.889	0.380	0.0067	0.759

Table 3. Intercorrelation Matrix of Independent Variables Used in the Models

compound	variables	S_1	S_6	S_{11}	S_{15}	S_{19}	PCh_8	PCh_{20}	I_{12-OH}
bridged-bicyclic diarylethylenes	S_1	1.00	1.00	0.50	0.53	0.10	0.29	0.02	0.05
	S_6		1.00	0.01	0.54	0.11	0.22	0.04	0.04
	S_{11}			1.00	0.17	0.14	0.42	0.12	0.23
	S_{15}				1.00	0.06	0.41	0.27	0.27
	S_{19}					1.00	0.12	0.20	0.23
	PCh_8						1.00	0.28	0.16
	PCh_{20}							1.00	0.16
	I_{12-OH}								1.00

compound	variables	Ch_5	Ch_{15}	Ch_{16}	Ch_{17}	E_{LUMO}	PMI_y	C_{SO_3Ph}
bridged oxabicyclic diarylethylenes	Ch_5	1.00	0.12	0.39	0.90	0.17	0.17	0.14
	Ch_{15}		1.00	0.89	0.31	0.27	0.29	0.03
	Ch_{16}			1.00	0.50	0.20	0.35	0.13
	Ch_{17}				1.00	0.02	0.26	0.17
	E_{LUMO}					1.00	0.44	0.56
	PMI_y						1.00	0.10
	C_{SO_3Ph}							1.00

and in the case of a bivariate relationship, the best significant relationship has been explored with Ch_1 and an indicator, N_p (no. of free terminal atoms, excluding hydrogens in the molecule), that explained 63.252% variance in activity. The quality of this relationship has been estimated to be

$$R = 0.830; \quad R^2 = 0.689; \quad s = 0.354; \quad n = 14$$

But, the best relationship for ER binding affinity to α-subtype has been deduced to be

$$[\text{pRBA}]_\alpha = 5.219(\pm 1.917)PCh_{20} - 0.784(\pm 0.119)S_{15} - 0.454(\pm 0.091)S_{19} + 0.739(\pm 0.196)I_{12-OH} + 3.900(\pm 0.381) \quad (1)$$

where S_{15}, S_{19}, and PCh_{20} are the *E*-state indices of atoms C_{15} and C_{19}, and the Extended Huckel atomic partial charge on C_{20}, respectively. I_{12-OH} indicates the presence of a hydroxyl group at the C_{12} atom.

For binding affinities to the β-subtype, the best univariate relation has again been deduced with Ch_1 that explained 48.103% variance in activity, and the statistical quality of the relationship is estimated to be

$$r = 0.722; \quad r^2 = 0.521; \quad s = 0.439; \quad n = 14$$

and similarly the best bivariate significant relationship has been explored with Ch_1 and N_p that explained 77.199% variance in activity. The quality of this relationship has been estimated as

$$R = 0.898; \quad R^2 = 0.807; \quad s = 0.291; \quad n = 14$$

The best relationship for binding affinity to ER_β has been found to be

$$[\text{pRBA}]_\beta = 7.088(\pm 1.735)PCh_{20} - 0.797(\pm 0.107)S_1 - 0.535(\pm 0.086)S_{19} + 0.822(\pm 0.177)I_{12-OH} + 3.190(\pm 0.272) \quad (2)$$

where S_1 is the *E*-state index of atom C_1.

In terms of selectivity (ER_β/ER_α) toward binding affinity to the receptor subtypes of the compounds investigated, the best significant single variate relation has been explored with a charge function at atom C_8 (Ch_8) that explained 29.471% variance in activity, and the statistical quality of the relationship is estimated to be

$$r = 0.591; \quad r^2 = 0.350; \quad s = 0.234; \quad n = 14$$

and in the case of a bivariate relationship, the best significant relationship has been explored with Ch_8 and the *E*-state index of atom C_4 that explained 50.436% variance in activity. The quality of this relationship has been estimated as

$$R = 0.762; \quad R^2 = 0.581; \quad s = 0.196; \quad n = 14$$

But, the overall best relationship for selective binding affinity to ER (β/α) has been found to be

$$[\text{pRBA}]_{\beta/\alpha} = 1.628(\pm 0.314)S_6 + 0.051(\pm 0.020)S_{11} - 28.670(\pm 5.528)PCh_8 + 3.232(\pm 0.899)PCh_{20} - 22.819(\pm 4.744) \quad (3)$$

Table 4. Calculated and Predicted Activities of Bridged-Cyclic Diaryl Compounds from the QSAR Models

	selective ER subtype binding affinity								
	ER_α			ER_β			selectivity, ER_β/ER_α		
compd no.	obs[g]	calc[g]	pred[g]	obs[g]	calc[g]	pred[g]	obs[g]	calc[g]	pred[g]
1	0.290[h]	0.679[a]	0.798[a]	0.521[h]	0.442[c]	0.421[c]	3.230[h]	3.025[e]	2.964[e]
2	1.398[h]	1.476[a]	1.515[a]	1.284[h]	1.344[c]	1.374[c]	2.886[h]	2.965[e]	2.983[e]
3	1.509[h]	1.305[a]	1.201[a]	1.456[h]	1.108[c]	0.929[c]	2.960[h]	3.052[e]	3.090[e]
4	0.967[h]	0.833[a]	0.805[a]	0.728[h]	0.673[c]	0.663[c]	2.770[h]	2.731[e]	2.682[e]
5	1.337[h]	1.463[a]	1.526[a]	1.051[h]	1.339[c]	1.484[c]	2.721[h]	2.769[e]	2.806[e]
6	0.614[h]	0.921[a]	1.025[a]	0.499[h]	0.815[c]	0.923[c]	2.886[h]	2.873[e]	2.868[e]
8	1.013[h]	0.824[a]	0.779[a]	0.903[h]	0.566[c]	0.482[c]	2.886[h]	2.897[e]	2.898[e]
9	2.046[h]	2.126[a]	2.397[a]	1.886[h]	1.850[c]	1.733[c]	2.850[h]	2.877[e]	2.964[e]
10	1.244[h]	1.201[a]	1.188[a]	0.724[h]	0.905[c]	0.969[c]	2.480[h]	2.549[e]	2.577[e]
11	2.000[h]	1.914[a]	1.651[a]	2.097[h]	1.974[c]	1.624[c]	3.097[h]	3.025[e]	2.933[e]
12	1.824[h]	2.134[a]	2.404[a]	1.886[h]	1.969[c]	2.050[c]	3.060[h]	2.969[e]	2.955[e]
13	1.168[h]	0.998[a]	0.961[a]	0.476[h]	0.494[c]	0.499[c]	2.310[h]	2.133[e]	1.864[e]
14	2.495[h]	2.279[a]	2.019[a]	2.046[h]	1.929[c]	1.808[c]	2.550[h]	2.617[e]	2.647[e]
15	3.968[i]	3.299[b]	3.164[b]	2.230[i]	2.010[d]	1.837[d]	1.732[i]	3.281[f]	3.391[f]
16	3.613[i]	3.358[b]	3.276[b]	1.176[i]	1.222[d]	1.258[d]	2.431[i]	3.474[f]	3.506[f]
17	2.806[i]	2.361[b]	2.293[b]	0.771[i]	0.772[d]	0.773[d]	2.041[i]	2.365[f]	2.265[f]
18	2.000[i]	1.822[b]	1.760[b]	0.886[i]	0.721[d]	0.695[d]	1.114[i]	1.840[f]	1.781[f]
19	1.322[i]	1.752[b]	1.826[b]	0.556[i]	0.714[d]	0.733[d]	0.763[i]	2.003[f]	2.037[f]
20	1.301[i]	1.307[b]	1.309[b]	0.398[i]	0.506[d]	0.530[d]	0.903[i]	1.650[f]	1.596[f]
21	1.114[i]	1.269[b]	1.315[b]	0.230[i]	0.609[d]	0.655[d]	0.886[i]	1.681[f]	1.675[f]
22	2.954[i]	3.070[b]	3.111[b]	1.322[i]	1.160[d]	1.139[d]	1.633[i]	3.201[f]	3.054[f]
23	1.838[i]	1.951[b]	1.958[b]	0.568[i]	0.569[d]	0.569[d]	1.278[i]	2.033[f]	2.069[f]
24	2.991[i]	2.906[b]	2.891[b]	1.041[i]	1.190[d]	1.209[d]	1.950[i]	3.003[f]	2.933[f]
25	1.832[i]	2.304[b]	2.432[b]	0.819[i]	0.822[d]	0.824[d]	1.000[i]	2.044[f]	2.117[f]
26	1.832[i]	1.918[b]	1.938[b]	1.079[i]	1.306[d]	1.501[d]	0.756[i]	2.111[f]	2.139[f]
27	0.903[i]	0.879[b]	0.872[b]	0.114[i]	0.215[d]	0.225[d]	0.792[i]	1.020[f]	1.102[f]
28	1.146[i]	1.210[b]	1.236[b]	0.301[i]	0.024[d]	0.071[d]	0.845[i]	1.641[f]	1.555[f]
29	2.505[i]	2.632[b]	2.653[b]	1.000[i]	0.880[d]	0.857[d]	1.518[i]	2.806[f]	2.841[f]
30	1.342[i]	1.798[b]	1.897[b]	0.431[i]	0.410[d]	0.406[d]	0.914[i]	2.020[f]	2.041[f]
31	1.398[i]	1.791[b]	1.883[b]	0.462[i]	0.481[d]	0.491[d]	0.934[i]	1.917[f]	1.914[f]
32	3.342[i]	3.474[b]	3.512[b]	1.204[i]	0.992[d]	0.957[d]	2.146[i]	3.320[f]	3.361[f]
33	0.845[i]	0.460[b]	0.349[b]	0.380[i]	0.489[d]	0.505[d]	0.462[i]	1.359[f]	1.311[f]
34	1.079[i]	0.572[b]	0.444[b]	0.491[i]	0.598[d]	0.626[d]	0.591[i]	1.494[f]	1.444[f]

[a] As per eq 1. [b] As per eq 4. [c] As per eq 2. [d] As per eq 6. [e] As per eq 3. [f] As per eq 7. [g] obs = observed values; calc = calculated values; pred = predicted values. [h] Reference 19. [i] Reference 20.

where S_6, S_{11}, and PCh_8 indicate *E*-state indices of atoms C_6 and C_{11} and the Extended Huckel atomic partial charge on C_8, respectively.

(b) Bridged-Oxabicyclic Diarylethylenes. While modeling RBA of bridged-oxabicyclic diarylethylene derivatives, the best univariate model for binding affinity to the α-subtype of ER could be developed with a charge function of atom C_{16} (Ch_{16}) that explained 53.441% variance in activity, and the statistical quality of the relation is estimated to be

$$r = 0.748; \quad r^2 = 0.559; \quad s = 0.660; \quad n = 20$$

and in the case of a bivariate relationship, the best significant relationship has been explored with the charge functions of atoms C_6 (Ch_6) and C_{16} (Ch_{16}) that explained 71.830% variance in activity. The quality of this relationship has been estimated to be

$$R = 0.865; \quad R^2 = 0.748; \quad s = 0.514; \quad n = 20$$

The best relationship deduced for ER_α binding affinity could explain 86.520% activity variance with good predictive property, and the relation is

$$[\text{pRBA}]_\alpha = -6.253(\pm 0.950)\text{Ch}_5 - 1.053(\pm 0.392)\text{Ch}_{16} - 3.240(\pm 0.711)E_{\text{LUMO}} - 3.045(\pm 0.604) \quad (4)$$

where Ch_5 and E_{LUMO} indicate the atomic charge function of C_5 and the lowest unoccupied molecular orbital energy, respectively.

In the case of β-receptor subtype binding, the best single variate model could be developed with an indicator, C_{SO_3Ph} (configuration of $-SO_3Ph$ substitution at Y) that explained 46.855% variance in activity, and statistical quality of the relation is estimated to be

$$r = 0.705; \quad r^2 = 0.496; \quad s = 0.360; \quad n = 20$$

and in the case of a bivariate relationship, the best significant relationship has been explored with C_{SO_3Ph} and charge function of atom C_3 (Ch_3) that explained 71.442% variance in activity, and the quality of this relationship has been estimated to be

$$R = 0.863; \quad R^2 = 0.744; \quad s = 0.264; \quad n = 20$$

But, for binding affinities to the ER_β, the best relation developed is

$$[\text{pRBA}]_\beta = 0.835(\pm 0.226)\text{Ch}_{17} - 2.309(\pm 0.457)\text{Ch}_5 + 1 \times 10^{-4}(\pm 4 \times 10^{-5})\text{PMI}_y + 1.083(\pm 0.130)C_{\text{SO}_3\text{Ph}} - 0.193(\pm 0.164)^{\#} \quad (5)$$

where Ch_{17} and PMI_y indicate the atomic charge function of

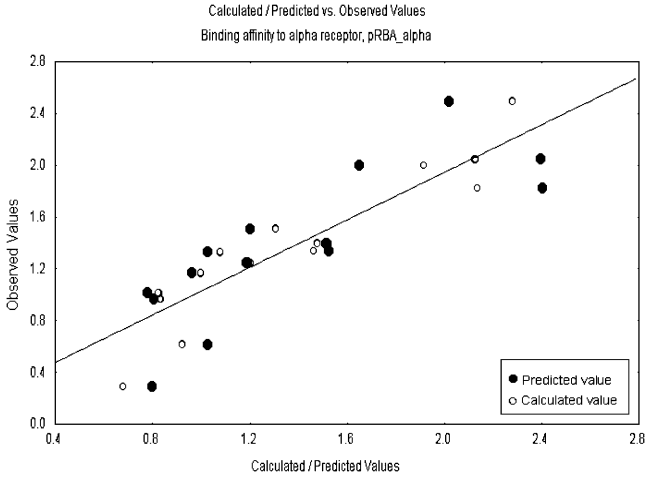


Figure 2. Observed, calculated, and predicted values as per eq 1.

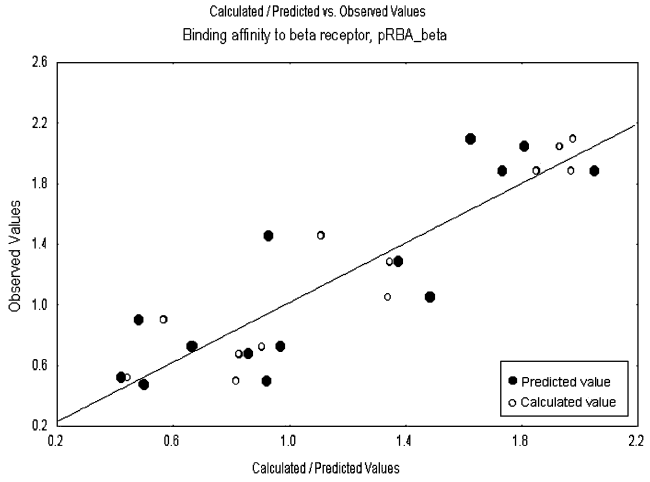


Figure 3. Observed, calculated, and predicted values as per eq 2.

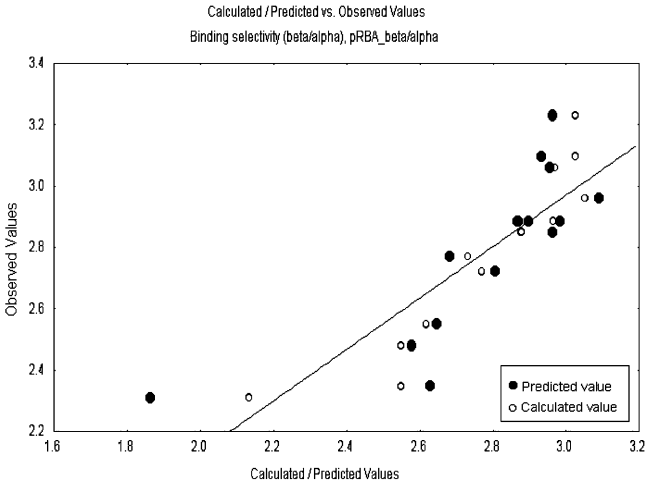


Figure 4. Observed, calculated, and predicted values as per eq 3.

C_{17} and the principal moment of inertia along the *y*-Cartesian coordinate axis, respectively. In this model (eq 5), the intercept is found statistically insignificant, and deletion of the same did not affect the relation but could explain as much as 96% activity variance, and the relationship is

$$[\mathrm{pRBA}]_{\beta} = 0.249(\pm 0.062)\mathrm{Ch}_{17} - 0.551(\pm 0.113)\mathrm{Ch}_{5} + 0.363(\pm 0.108)\mathrm{PMI}_{y} + 0.419(\pm 0.050)C_{\mathrm{SO_3Ph}} \quad (6)$$

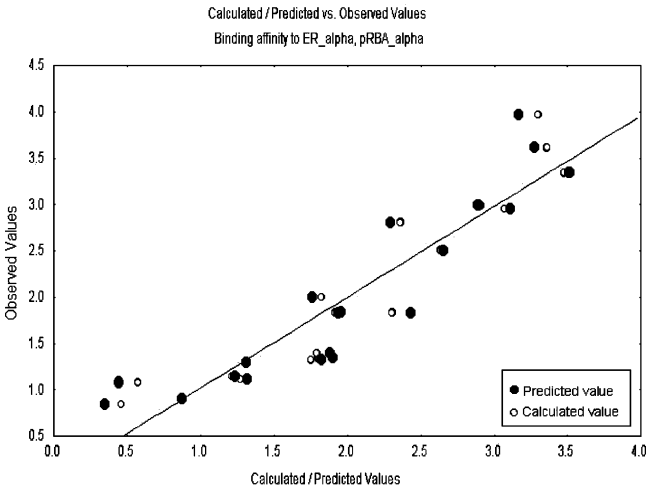


Figure 5. Observed, calculated, and predicted values as per eq 4.

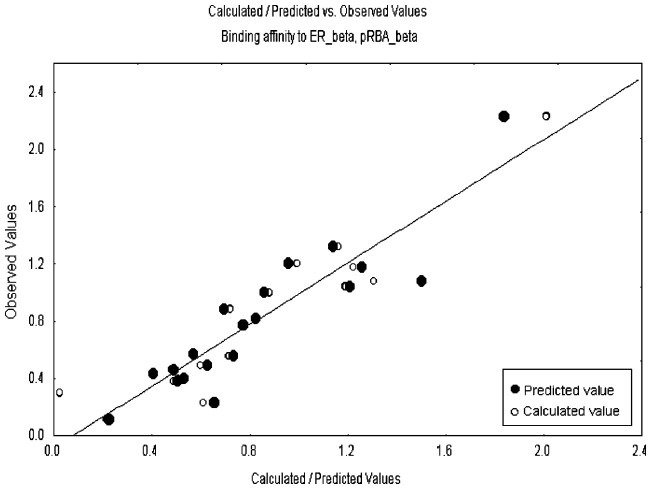


Figure 6. Observed, calculated, and predicted values as per eq 6.

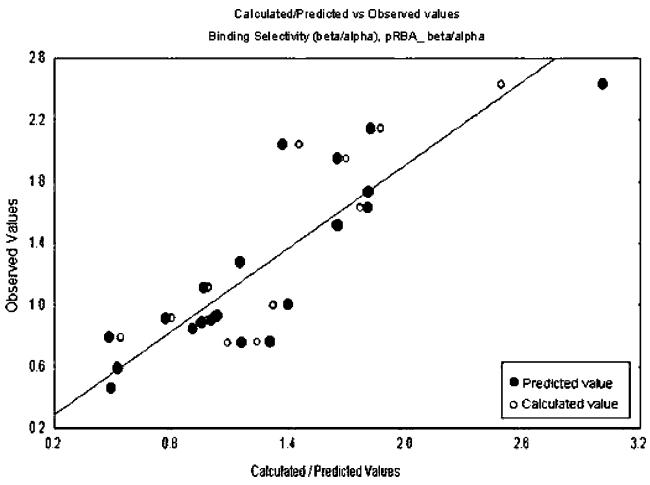


Figure 7. Observed, calculated, and predicted values as per eq 7.

For selectivity toward the binding affinity for ER (β/α), the best univariate model could be developed with E_{LUMO} that explained 33.247% variance in activity, and the statistical quality of the relationship is estimated to be

$$r = 0.606; \quad r^2 = 0.368; \quad s = 0.648; \quad n = 20$$

and in the case of a bivariate relationship, the best significant relation has been explored with the atomic charge function

Table 5. Hypotheses Parameters Observed in Successive Runs for ER_α Binding Affinity of Bridged-Cyclic Diaryl Analogs

run parameters					cost[d]				
run no.	spacing (pm)	input features[a]	hypothesis no.	pharmacophore features in generated hypothesis[a]	configuration	total	Δ	*R*	rmsd[e]
1	250	a, d, p, r	1	a, p, 2 × r	10.532	112.963	71.617	0.916	1.137
			3	d, p, 2 × r	10.532	115.778	68.802	0.901	1.232
2	180	a, d, p, r	1	a, p, 2 × r	11.381	110.847	73.733	0.932	0.994
			2	d, p, 2 × r	11.381	115.62	68.96	0.906	1.199
3	100	a, d, p, r	1	a, p, 2 × r	11.577	110.407	74.173	0.936	0.994
			3	d, p, 2 × r	11.577	116.564	68.016	0.902	1.224
4	10	a, d, p, r	1	a, p, 2 × r	12.104	112.879	71.701	0.927	1.059
			2	d, p, 2 × r	12.104	117.03	67.550	0.903	1.221
			3	d, p, 2 × r	12.104	117.201	67.379	0.901	1.228
5	60	a, d, p, r	1	a, p, 2 × r	11.577	110.407	74.173	0.936	0.994
			2	a, p, 2 × r	11.577	112.963	71.617	0.923	1.089
			3	d, p, 2 × r	11.577	116.564	68.016	0.902	1.224
6	60	a, p, r	1	a, p, 2 × r	9.947	111.314	73.266	0.923	1.093
			2	a, p, 2 × r	9.947	114.769	69.811	0.903	1.216
7	60	d, p, r	1	d, p, 2 × r	10.140	113.643	70.937	0.909	1.175
			2	d, p, 2 × r	10.140	114.413	70.167	0.906	1.201
8[b]	60	a, d, p, r	1	a, p, 2 × r	23.187	122.661	61.919	0.932	1.028
			3	d, p, 2 × r	23.187	127.862	56.718	0.904	1.213
9[c]	60	a, d, p, r	1	a, p, 2 × r	23.187	124.71	59.87	0.934	1.014
			2	d, p, 2 × r	23.187	130.368	54.212	0.903	1.218
10[b,c]	60	a, d, p, r	1	a, p, 2 × r	34.796	135.197	49.383	0.938	0.980
			3	d, p, 2 × r	34.796	140.647	43.933	0.909	1.179

[a] a: HB acceptor; d: HB donor; p: hydrophobic factor; r: ring aromatic. [b] Weight variation changed to 1 from default value. [c] Variable tolerance changed to 1 from default value. [d] Null cost = 184.580; Δ = null cost − total cost. [e] rmsd: rms deviation; $n = 25$.

Table 6. Hyporefine Parameters Observed for Hypothesis 1 of Run 5 in ER_α Binding Affinity of Bridged-Cyclic Diaryl Analogs

run parameters					cost[b]				
run no.	spacing (pm)	input features[a]	hypothesis no.	pharmacophore features in generated hypothesis[a]	configuration	total	Δ	*R*	rmsd[c]
11	60	a, d, p, r	1	a, 2 × p, r, e	12.173	110.532	74.048	0.938	0.976
			2	d, 2 × p, r, e	12.173	112.568	72.012	0.928	1.052
			3	a, p, 2 × r, e	12.173	115.743	68.837	0.909	1.179

[a] a: HB acceptor; d: HB donor; p: hydrophobic factor; r: ring aromatic; e: excluded volume. [b] Null cost = 184.580; Δ = null cost − total cost. [c] rmsd: rms deviation; $n = 25$.

at C_5 (Ch_5) along with E_{LUMO} that explained 75.299% variance in activity. The quality of this relationship has been estimated to be

$$R = 0.883; \quad R^2 = 0.719; \quad s = 0.393; \quad n = 20$$

But, the best relation for binding selectivity toward ER (β/α) has been explored to be

$$[\mathrm{pRBA}]_{\beta/\alpha} = 0.961(\pm 0.441)\mathrm{Ch}_{17} - 5.734(\pm 0.875)\mathrm{Ch}_5 - 3.711(\pm 0.534)E_{\mathrm{LUMO}} - 2.202(\pm 0.512) \quad (7)$$

where Ch_{17} indicates atomic charge function of C_{17}.

The statistical parameters of the best relations (eqs 1−7) are described in Table 2, and independent variables used in various equations are not intercorrelated ($r < 0.5$) (Table 3). The calculated and predicted activities from the equations have been delineated in Table 4 and Figures 2−7.

Pharmacophore Space Modeling. The results of this study have been depicted in Tables 5−13. The *hypothesis* 1 of *run* number 5 (Table 5) has been adjudged to be the best pharmacophore hypothesis for binding affinity to ER_α, while that of *run* number 13 for *hypothesis* 1 (Table 8) has been adjudged to be the best hypothesis for binding affinity to

Table 7. Results of CatScramble Cross-Validation[a] of Run 5 (Hypothesis 1) for ER_α Binding Affinity of Bridged-Cyclic Diaryl Analogs

trial/ spreadsheet no.	cost[b]					
	configuration	fixed	total	Δ	*R*	rmsd[c]
1	10.288	96.389	166.88	17.700	0.547	2.373
2	17.711	103.872	161.925	22.655	0.650	2.154
3	11.808	97.968	173.574	11.006	0.498	2.457
4	17.362	103.522	179.31	5.270	0.496	2.462
5	10.739	96.900	166.719	17.861	0.553	2.363
6	17.774	103.935	154.60	29.980	0.704	2.012
7	16.230	102.39	181.634	2.946	0.461	2.515
8	14.520	100.68	157.302	27.278	0.661	2.127
9	18.885	108.046	171.799	12.781	0.589	2.291
10	19.253	105.414	168.913	15.667	0.608	2.251
11	11.777	97.937	136.728	47.852	0.785	1.755
12	17.306	103.467	175.142	9.438	0.566	2.340
13	17.318	103.479	162.557	22.023	0.642	2.173
14	10.357	96.518	184.504	0.076	0.353	2.653
15	14.154	100.314	156.374	28.206	0.665	2.116
16	16.943	103.104	175.285	9.295	0.531	2.402
17	18.939	105.10	164.103	20.477	0.643	2.173
18	17.471	103.631	150.684	33.896	0.730	1.937
19	11.851	98.012	169.781	14.799	0.538	2.395

[a] 95% confidence level. [b] Null cost = 184.580; Δ = null cost − total cost. [c] rmsd: rms deviation.

Table 8. Hypotheses Parameters Observed in Successive Runs for ER_β Binding Affinity of Bridged-Cyclic Diaryl Analogs

run parameters					cost[d]				
run no.	spacing (pm)	input features[a]	hypothesis no.	pharmacophore features in generated hypothesis[a]	configuration	total	Δ	*R*	rmsd[e]
12	250	a, d, p, r	1	a, p, 2 × r	10.445	104.781	110.659	0.969	0.791
			3	d, p, 2 × r	10.445	106.049	109.391	0.964	0.857
13	180	a, d, p, r	1	a, p, 2 × r	11.208	103.911	111.529	0.976	0.707
			2	d, p, 2 × r	11.208	105.691	109.749	0.969	0.796
14	100	a, d, p, r	1	a, p, 2 × r	11.407	104.243	111.197	0.975	0.707
			3	d, p, 2 × r	11.407	107.002	108.438	0.964	0.862
15	10	a, d, p, r	1	a, p, 2 × r	11.981	106.411	109.029	0.969	0.798
			3	d, p, 2 × r	11.981	106.919	108.521	0.967	0.817
16	180	a, p, r	1	a, p, 2 × r	9.599	103.311	112.129	0.971	0.767
			2	a, p, 2 × r	9.599	104.003	111.437	0.969	0.800
17	180	d, p, r	1	d, p, 2 × r	9.603	104.355	111.085	0.967	0.817
			2	d, p, 2 × r	9.603	105.636	109.804	0.962	0.875
18[b]	180	a, d, p, r	1	a, p, 2 × r	22.818	115.396	100.044	0.977	0.682
			2	d, p, 2 × r	22.818	118.578	96.862	0.963	0.862
19[c]	180	a, d, p, r	1	a, p, 2 × r	22.818	116.526	98.914	0.977	0.680
			3	d, p, 2 × r	22.818	120.88	94.560	0.961	0.901
20[b,c]	180	a, d, p, r	1	a, p, 2 × r	34.428	131.953	83.487	0.963	0.863
			3	d, p, 2 × r	34.428	132.863	82.577	0.960	0.901

[a] a: HB acceptor; d: HB donor; p: hydrophobic factor; r: ring aromatic. [b] Weight variation changed to 1 from default value. [c] Variable tolerance changed to 1 from default value. [d] Null cost = 215.439; Δ = null cost − total cost. [e] rmsd: rms deviation; $n = 25$.

Table 9. Hyporefine Parameters Recorded for Hypothesis 1 of Run 13 in ER_β Binding Affinity of Bridged-Cyclic Diaryl Analogs

run	run parameters		hypothesis	pharmacophore features	cost[b]				
no.	spacing (pm)	input features[a]	no.	in generated hypothesis[a]	configuration	total	Δ	*R*	rmsd[c]
21	180	a, d, p, r	1	a, p, 2 × r, e	11.944	103.713	111.727	0.979	0.652
			2	d, 2 × p, r, e	11.944	104.821	110.619	0.975	0.713
			3	a, 2 × p, r, e	11.944	105.062	110.378	0.974	0.727

[a] a: HB acceptor; d: HB donor; p: hydrophobic; r: ring aromatic; e: excluded volume. [b] Null cost = 215.439; Δ = null cost − total cost. [c] rmsd: rms deviation; $n = 25$.

Table 10. Results of CatScramble Cross-Validation[a] of Run 13 (Hypothesis 1) for ER_β Binding Affinity of Bridged-Cyclic Diaryl Analogs

trial/ spreadsheet	cost[b]					
no.	configuration	fixed	total	Δ	*R*	rmsd[c]
1	11.365	97.525	168.231	47.208	0.687	2.356
2	15.203	101.363	180.876	34.563	0.628	2.521
3	9.614	95.775	224.344	−8.905	0.150	3.204
4	11.244	97.404	181.43	34.009	0.605	2.579
5	17.535	103.894	153.201	62.238	0.80	1.945
6	13.655	99.815	181.062	34.377	0.617	2.549
7	13.703	99.863	167.615	47.824	0.701	2.311
8	10.021	96.181	169.606	45.833	0.673	2.397
9	14.879	101.04	199.368	16.071	0.526	2.758
10	13.726	99.886	171.929	43.510	0.682	2.372
11	9.791	95.951	214.495	0.944	0.312	3.079
12	12.222	98.382	193.937	21.502	0.523	2.762
13	15.472	101.633	181.432	34.007	0.628	2.520
14	15.460	101.62	166.674	48.765	0.711	2.277
15	11.521	97.682	202.696	12.743	0.447	2.898
16	17.312	103.472	182.156	33.283	0.634	2.507
17	13.909	100.069	161.377	54.062	0.733	2.205
18	10.581	96.742	192.597	22.842	0.520	2.768
19	14.556	100.716	199.741	15.698	0.497	2.812

[a] 95% confidence level. [b] Null cost = 215.439; Δ = null cost − total cost. [c] rmsd: rms deviation.

ER_β. This has been characterized in terms of the highest cost difference, the lowest root-mean-square divergence (rmsd), and the best correlation coefficient. The generated best hypotheses are validated to nullify overprediction of bioactivities for inactive compounds through Hyporefine. In this process, the steric interactions of compounds are considered in the hypothesis generation and are portrayed in the validated (refined) hypothesis (*run* nos. 11 and 21). The mapped pharmacophore features for binding affinity to the receptor subtypes (α and β) are described in Table 11 and Figures 8 and 9, and estimated fit scores are delineated in Figures 10 and 11, respectively. The quality of hypothesis generated is adjudged through a cross-validation technique using Fischer's randomization test,[47] where the biological activity data are randomized within a fixed chemical data set and the HypoGen process is initiated to explore possibilities of other hypotheses of predictive values (Tables 7 and 10). The predicted hypothesis for α and β ER binding are further tested with some estrogenic and nonestrogenic compounds as reference compounds, and the fit score of the compounds are listed in Table 13.

The binding affinity of bridged-cyclic diaryl scaffolds toward the ER subtypes (α and β) demonstrated the importance of HB acceptor (a), hydrophobic (p), and ring aromatic (r) features along with the steric influence (e) of the compounds. In the hyporefined models, the distance between a and p constraints are observed to be 8.496 Å and 7.830 Å, while the separation of a and r constraints are found to be 7.080 Å and 7.699 Å for α- and β-subtypes binding, respectively (Table 11, Figures 8 and 9).

DISCUSSION

Reasonably well predictable models for binding affinity to ER-subtypes of diaryl compounds have been obtained with

Table 11. Features Obtained from Refined Pharmacophore Hypotheses[a]

Hyporefine for ERα	interfeature distance (Å)[b]					Hyporefine for ERβ	interfeature distance (Å)[b]				
	p1	p2	a	r1	e		p1	a	r1	r2	e
p1	0.000	5.587	8.496	6.349	10.299	p1	0.000	7.830	5.832	5.978	6.195
p2		0.000	2.964	5.149	6.617	a		0.000	2.820	7.699	8.016
a			0.000	7.080	5.707	r1			0.000	4.975	5.476
r1				0.000	8.614	r2				0.000	4.060
e					0.000	e					0.000

[a] Run nos. 11 and 21. [b] p1 = hydrophobic1; p2 = hydrophobic2; a = hydrogen bond acceptor; r1 = ring aromatic 1; r2 = ring aromatic 2; e = excluded volume.

Table 12. Observed and Estimated Activities of Training Set for Bridged-Cyclic Diaryl Analogs

	ER binding affinity (RBA)					
	α-subtype[a]			β-subtype[a]		
compd no.	obs	est run 5 (hypo 1)	est Hyporefine (hypo 1)	obs	est run 13 (hypo 1)	est Hyporefine (hypo 1)
1	1.95	14.0	4.2	3.32	8.3	1.6
2	25.00	12.0	4.2	19.23	9.3	1.5
3	32.26	15.0	4.9	28.57	9.6	1.4
4	9.26	9.3	2.7	5.35	9.4	1.1
5	21.74	14.0	3.8	11.2	11.0	1.6
6	4.12	17.0	2.4	3.16	9.3	0.9
7	21.28	26.0	5.8	4.74	10.0	1.0
8	10.31	11.0	4.3	8.00	4.3	1.0
9	111.11	50.0	1.6	76.92	38.0	1.0
10	17.54	18.0	4.7	5.29	17.0	1.2
11	100.00	270.0	9.8	125.00	110.0	2.0
12	66.67	9.1	6.1	76.92	19.0	3.3
15	107.53	1.1×10^3	15.0	588.24	2.5×10^3	1.6×10^4
16	243.90	230.0	3.5	6.67×10^3	4.5×10^3	210.0
17	1.56×10^3	2.2×10^3	6.9	1.70×10^4	7.2×10^3	250.0
18	1.00×10^4	7.9×10^3	1.9×10^3	1.30×10^4	2.1×10^4	2.5×10^4
19	4.76×10^4	1.4×10^4	4.6	2.78×10^4	1.8×10^4	280.0
20	5.00×10^4	1.2×10^5	2.1×10^3	4.00×10^4	4.1×10^4	2.8×10^4
21	7.69×10^4	1.4×10^4	150.0	5.88×10^4	4.2×10^4	1.9×10^4
22	1.11×10^3	990.0	1.7	4.76×10^3	7.8×10^3	4.0
23	1.45×10^4	3.2×10^3	2.6	2.70×10^4	7.7×10^3	40.0
24	1.02×10^3	4.1×10^3	3.5	9.09×10^3	1.1×10^4	830.0
25	1.47×10^4	4.7×10^3	14	1.51×10^4	1.4×10^4	1.2×10^4
29	3.13×10^3	4.9×10^3	5.8	1.00×10^4	1.0×10^4	490.0
32	454.545	960.0	4.7	6.25×10^3	1.6×10^4	340.0

[a] obs: observed activity; est: estimated activity.

cross-validated variance (CVV) exceeding 50%. The models generated for bridged-cyclic diaryl derivatives account for more than 75% variance in observed activity with low estimation errors. However, the best relations (eqs 1 and 2) revealed the importance of atoms C_{19}, C_{20} and the presence of a hydroxyl group[23] at C_{12} of bridged-cyclic diaryl-5-hydroxy analogs for binding affinities to both the ER subtypes. Furthermore, the ethylenic fragment (atom C_{15} in eq 1 and atom C_1 in eq 2) signifies the impact for binding to both the ER_α and ER_β. The negative coefficient of S_{19} indicates that increase in the *E*-state value of atom C_{19} or an increment in the e^- density in the vicinity of this atom will result in enhanced binding to both subtypes. A negative coefficient of the binary indicator also suggests that the presence of a hydroxyl group at atom C_{12} (in ring B) is further favorable for binding affinity to both the ER subtypes. Again, the positive coefficient of the partial charge at atom C_{20} signifies that amplification of the partial charge or more electronegativity at C_{20} will result in decreased binding to the receptor surface, which is selectively more in ER_β than in ER_α subtype. Negative coefficients of S_{15} (eq 1) and S_1 (eq 2) explain that an increase in e^- densities around atoms C_{15} and C_1 (ethylenic fragment) will result in enhanced binding to both ER subtypes. Thus, judging from eqs 1 and 2, it can be inferred that the presence of the hydroxyl group at C_{12}, electronic distributions in the vicinity of atoms C_{19} and C_{20} of bridged-cyclic diaryl-C_5-hydroxy analogues, and atoms C_{15} and C_1 (ethylenic fragment) are essential requirements for binding at the ER-ligand binding regions, irrespective of the subtypes. The significant contributions of atoms C_{15}, C_{19}, and C_{20} in either case, as obtained during the course of the present investigation, showed the importance of the cyclic ring system for ER binding of these groups of compounds. The importance of atoms C_1 and C_{15} can be conceptualized as the influence of unsaturation in the C_1-C_{15} linkage. The significance of an ethylenic fragment in nonsteroidal estrogens has been explored in a series of diverse compounds containing the diarylhydroxy unit.[16−18,39] In terms of binding selectivity to ER (β/α), the best model (eq 3) generated with 78% binding selectivity indicated that increase in *E*-state values of atoms O_6 and C_{11} will decrease selectivity of ER_β over ER_α. Consequently, substituents that

Table 13. Fit Scores of Training, Test, and Reference Compounds

		fit score			
		ER_α binding		ER_β binding	
set[a]	compd no./name	run 5 (hypo 1)	Hyporefine (hypo 1, run 11)	run 13 (hypo 1)	Hyporefine (hypo 1, run 21)
Tr	**1**	13.410	12.956	13.436	13.427
Tr	**2**	13.390	12.955	13.421	13.467
Tr	**3**	13.410	12.892	13.430	13.499
Tr	**4**	13.420	13.154	13.444	13.603
Tr	**5**	13.403	13.000	13.420	13.432
Tr	**6**	13.186	13.204	13.192	13.689
Tr	**7**	13.187	12.813	13.241	13.645
Tr	**8**	13.487	12.946	13.302	13.652
Tr	**9**	13.656	13.378	13.407	13.632
Tr	**10**	13.650	12.906	12.992	13.564
Tr	**11**	12.414	12.588	12.516	13.334
Tr	**12**	13.426	12.793	13.360	13.121
Tr	**15**	13.564	12.396	12.689	10.443
Tr	**16**	13.210	13.031	12.461	11.314
Tr	**17**	13.426	12.740	12.555	11.237
Tr	**18**	10.440	10.293	10.253	10.247
Tr	**19**	13.570	12.917	13.454	11.193
Tr	**20**	10.050	10.236	10.097	10.194
Tr	**21**	12.025	11.409	12.273	10.366
Tr	**22**	13.272	13.360	12.595	13.039
Tr	**23**	12.770	13.170	12.112	12.034
Tr	**24**	13.275	13.033	12.196	10.273
Tr	**25**	13.103	12.438	12.107	10.566
Tr	**29**	13.302	12.819	12.945	10.953
Tr	**32**	13.158	12.911	13.082	11.107
Ts	**13**	13.297	12.839	13.347	13.509
Ts	**14**	13.374	13.093	13.435	13.397
Ts	**26**	10.559	10.418	10.449	10.187
Ts	**27**	10.391	10.032	10.195	10.185
Ts	**28**	10.023	10.093	9.235	9.674
Ts	**30**	10.666	10.334	10.507	10.357
Ts	**31**	10.685	11.017	12.787	10.186
Ts	**33**	10.556	10.255	9.850	10.274
Ts	**34**	10.736	11.374	13.048	10.503
ref	raloxifene	13.071	13.059	12.922	12.710
ref	tamoxifene	13.308	13.359	13.256	13.048
ref	estradiol	9.684	10.056	9.707	8.333
ref	cyclofenil	13.371	13.059	13.400	13.605
ref	tetrahydroisoquinoline	12.768	13.326	12.402	12.691
ref	clidinium	7.136	7.015	7.004	7.014
ref	phenytoin	8.421	7.779	7.814	7.962
ref	norgestrel	7.139	8.619	7.019	7.019
ref	acetaminophine	6.694	6.917	6.685	6.928
ref	ephedrine	6.522	6.521	6.797	6.773
ref	morphine	6.655	8.035	6.396	7.988
ref	propranolol	7.138	7.897	7.003	7.016
ref	testosterone	6.859	6.992	6.623	6.478
ref	stanalone	7.137	8.523	6.981	7.018
ref	letrozole	10.117	9.586	10.166	9.960

[a] Tr = training set; Ts = test set; ref = reference compound.

tend to decrease the e^- density around these atoms should result in increased selectivity. The importance of the 5-phenolic hydroxyl group for selective estrogenic activity is demonstrated earlier.[48] The negative coefficient of PCh_8 and the positive coefficient of PCh_{20} correspondingly signify that increase and decrease in electronegativity around these atoms should result in an enhanced selectivity profile. From the observation, atom C_{20} appeared to be a fragment pivotal for both ER binding as well as ER_β selectivity.

The best relations explored for subtypes binding affinity of bridged-oxabicyclic diarylethylenes in eqs 4 and 6 have brought into the picture the importance of charge functions of atoms C_5, C_{16} and C_{17}, overall electron affinity along with

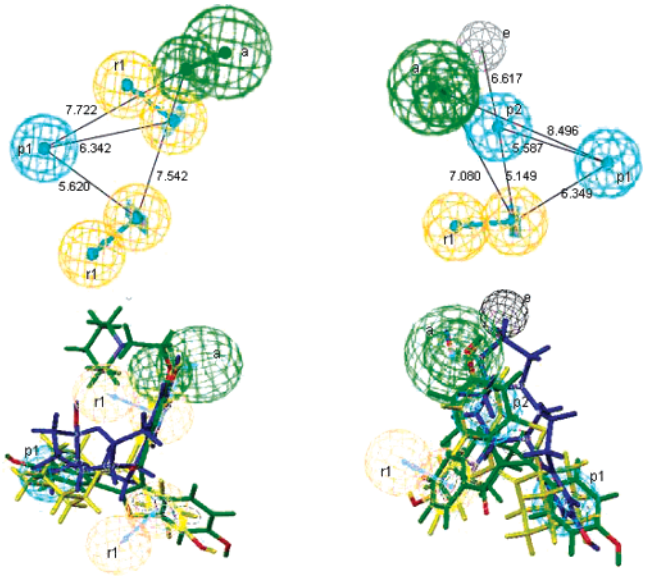


Figure 8. Pharmacophore features (*hypothesis* 1 of *run* nos. 5 and 11) for RBA to ER_α. Mapped features are HB acceptor (a), hydrophobic (p), ring aromatic (r), and excluded volume (e); green: raloxifene, yellow: compn no. 1; blue: E_2.

the orientation, and conformational rigidity of the oxabicyclic core for binding affinities to the ER subtypes. These relations revealed that increased negative charge on atom C_5 (influenced by the V-substituent) shall increase the binding affinity to both the receptor subtypes, since Ch_5 has negative contribution. Further the negative contribution of Ch_{16} in eq 4 and the positive contribution of Ch_{17} in eq 6 depicted that a decrease and an increase of negative charges at atoms C_{16} (influenced by the Y-substituent) and C_{17} (influenced by the Z-substituent) will cause lowering of the receptor binding affinity to ER_α and ER_β, respectively. A negative contribution of E_{LUMO} in eq 4 revealed that enhanced electron affinity of the molecules is detrimental for binding to ER_α, whereas conformational rigidity enhanced the binding affinity of the ligands to ER_β as is evident from the coefficient of PMI_y in eq 6. The positive contribution of PMI_y indicated that increase in the value of this parameter can increase the binding affinity to ER_β. PMI_y values of compounds **15** and **16** are 5563.300 and 4881.240 g/mol Å^2, and as such the former is more potent in binding to ER_β. Eq 6 also indicated the significance of configuration of the $-SO_3Ph$ (phenyl sulfonate) group attached to atom C_{16} in affecting the receptor binding affinity and demonstrated that β-configuration is favorable for binding affinity to ER_β. In terms of binding selectivity toward ER (β/α), the best model (eq 7) can portray that increased negative charge on atom C_5 (influenced by V-substituent), decreased negative charge on atom C_{17} (influenced by Z-substituent), and overall decrease in electron affinity of the molecule will favor the selectivity. Thus electron affinity and the orientation and conformational rigidity of oxabicyclic diarylethylenes along with Y- and Z-substituents differentiate the binding affinity to ER-subtypes.

The quality of best *hypothesis* generated in either case for pharmacophore space modeling studies on RBA are significant with regards to cost differences, correlation coefficients, and rmsd recorded. The best hypothesis (*hypo* 1, *run* no. 5) and hyporefine (*hypo* 1 of *run* no. 11) on the same for ER_α binding demonstrated around 94% correlation with the

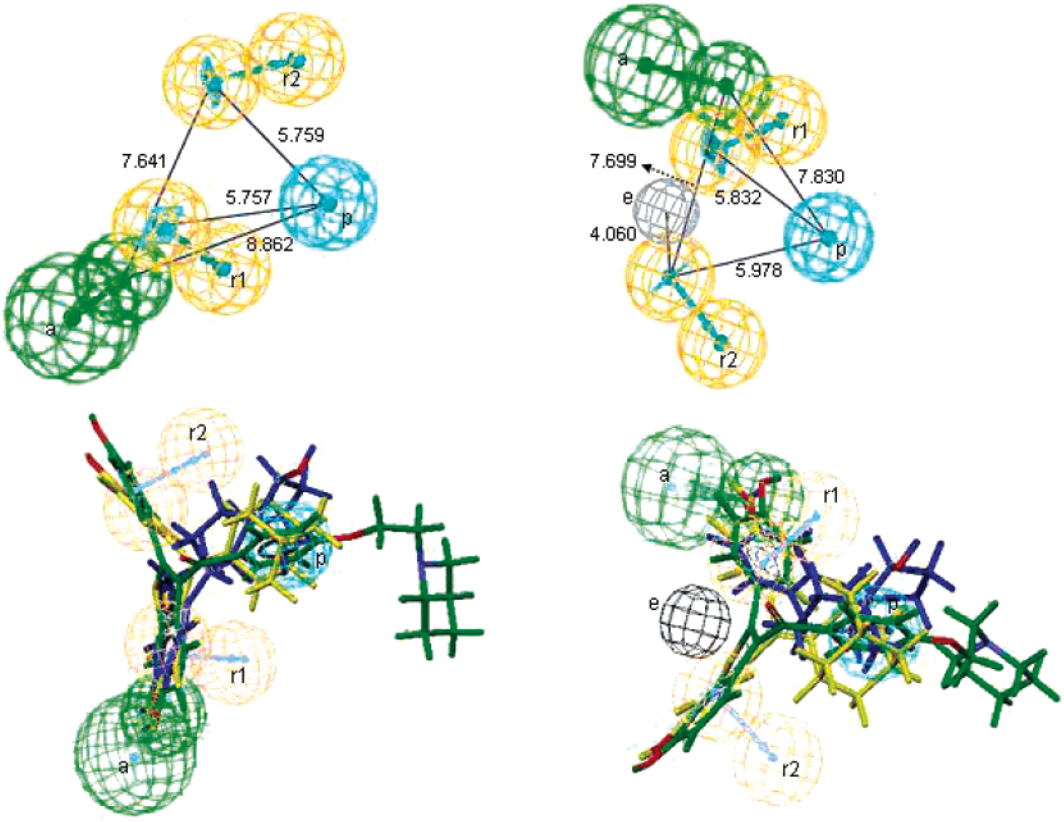


Figure 9. Pharmacophore features (*hypothesis* 1 of *run* nos. 13 and 21) for RBA to ER$_\beta$. Mapped features are HB acceptor (a), hydrophobic (p), ring aromatic (r), and excluded volume (e); green: raloxifene, yellow: compd no. 6; blue: E_2.

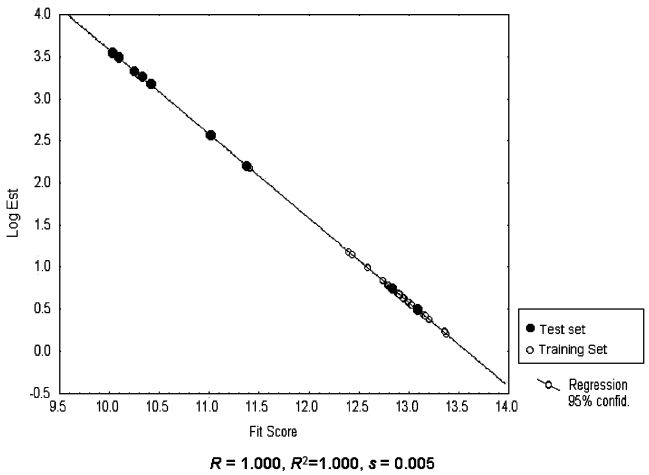


Figure 10. Estimated versus fit score from hyporefine of ER$_\alpha$ binding.

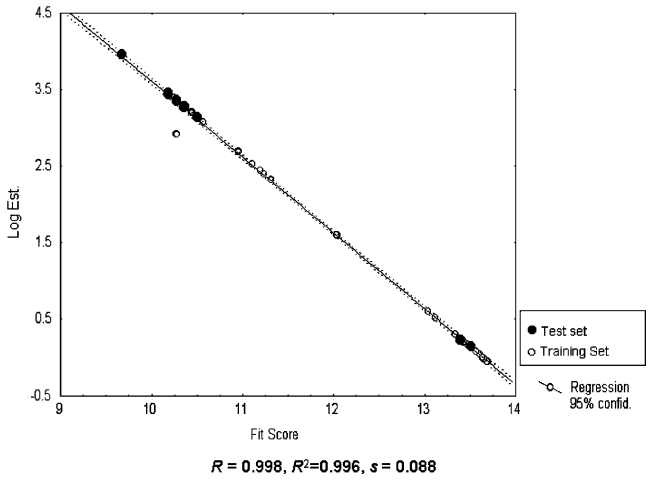


Figure 11. Estimated versus fit score from hyporefine of ER$_\beta$ binding.

binding affinity of compounds, while that for ER$_\beta$ (*hypo* 1 of *run* nos. 13 and 21) is to the extent of 98%. It has been deduced that HB acceptor, hydrophobic, ring aromatic, and steric features might function as prime biophores for ER binding affinity to both subtypes. However, the receptor subtype specific binding of compounds might arise due to geometrical distances between these features within the compounds. From Table 11 and Figures 8 and 9, the interbiophore distances between similar features can be perceived to be widely varying, which clearly differentiate the binding affinity to both ER subtypes. HB donor is not figured out as a significant contributor toward the ER-binding affinity of bridged-cyclic diaryl analogs. From Tables 7 and 10 for CatScrambling based on the cross-validation of both ER$_\alpha$ and ER$_\beta$ binding affinity respectively, it is observed that none of the spreadsheets generated better parameters in comparison to the original hypotheses. The cross-validation results clearly indicated the superiority of the hypotheses selected and also provide strong confidence on the initial pharmacophore *hypo* 1 of both the cases. The results also demonstrated superior predictive ability of the models for binding affinity to ER subtypes, while tested with reference compounds (Table 13). The standard SERMs are highly fitted (fit score >12), with both the models for the binding affinity to the ER subtypes, while non-SERMs, including E_2,[9] showing low fit score (<10).

The space modeling study can also be correlated with the classical approach of QSAR studies done on the two sets of

compounds. The bridged-cyclic core showed the importance of the hydrophobic feature of the molecule, and the presence of the HB-acceptor in aromatic ring A along with steric factors (influenced by the orientation and conformational rigidity) are found to be detrimental for receptor binding.

CONCLUSIONS

This work supports the fact that *p*-hydroxyl substitutions in rings A and B and bridged-cyclic systems connected through unsaturated linkages within the bridged-cyclic diaryl ligands could result in enhanced binding affinity to both ER subtypes. For ER_α and ER_β preferential bindings, increased electron density distribution near atoms C_{15} and/or C_1 could be crucial. Substitutions and/or bridged-cyclic systems capable of lowering electron densities around C_{11} could increase the binding selectivity (β/α) that can be further aided by increased electronegativity at atom C_8 and decreased electronegativity at atom C_{20}.

For bridged-oxabicyclic diarylethylenes, the analyses substantiate that enhanced electron density distribution near atom C_5 will increase the binding affinity to both the receptor subtypes. Additionally, decrease and increase in negative charges on atoms C_{16} and C_{17} are also liable to cause the lowering of receptor binding affinity to ER_α and ER_β. Enhanced electron affinity of the molecules is detrimental for binding to ER_α, but conformational rigidity of the compound might also increase the binding affinity to ER_β. The β-configuration of the phenyl sulfonate group also increases receptor binding affinity to ER_β. In terms of ER binding selectivity (β/α), it is likely that increased and decreased electron density distribution on atoms C_5 (due to the presence of a hydroxy substituent at C_6) and C_{17} (the presence of Z-substituents), respectively, and an overall decrease in electron affinity of the compounds favor for selectivity.

From the pharmacophore space modeling studies, it can be concluded that critical interfeature distances between HB acceptor, hydrophobic, and ring aromatic features along with steric influence primarily govern the ER-subtypes specific binding of scaffolds containing the 1,1-diaryl compounds of SERMs.

ACKNOWLEDGMENT

One of the authors, S. Mukherjee, wishes to thank CSIR, New Delhi (India) for the award of a senior research fellowship. Financing assistance from AICTE, New Delhi (India) under RPS program and University TEQIP scheme are also thankfully acknowledged.

REFERENCES AND NOTES

(1) Yaffe, K.; Sawaya, G.; Lieberburg, I.; Grady, D. Estrogen Therapy in Postmenopausal Women: Effects on Cognitive Function and Dementia. *JAMA* **1998**, *279*, 688−695.

(2) Writing Group for the Women's Health Initiative Investigators. Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women. *JAMA* **2002**, *288*, 321−333.

(3) Kuiper, G. G.; Enmark, E.; Pelto-Huikko, M.; Nilsson, S.; Gustafsson, J. A. Cloning of a Novel Receptor Expressed in Rat Prostate and Ovary. *Proc. Natl. Acad. Sci.* **1996**, *93*, 5925−5930.

(4) Levenson, A. S.; Jordan, V. C. Selective Oestrogen Receptor Modulation: Molecular Pharmacology for the Millennium. *Eur. J. Cancer* **1999**, *35*, 1628−1639.

(5) Kuiper, G. G.; Carlsson, B.; Grandien, K.; Enmark, E.; Haggblad, J.; Nilsson, S.; Gustafsson, J. A. Comparison of the Ligand Binding Specificity and Transcript Tissue Distribution of Estrogen Receptors α and β. *Endocrinology* **1997**, *138*, 863−870.

(6) Barkhem, T.; Carlsson, B.; Nilsson, Y.; Enmark, E.; Gustafsson, J.; Nilsson. S. Differential Response of Estrogen Receptor α and Estrogen Receptor β to Partial Estrogen Agonists/Antagonists. *Mol. Pharmacol.* **1998**, *54*, 105−112.

(7) Pike, A. C. W.; Brzozowski, A. M.; Hubbard, R. E.; Bonn, T.; Thorsell, A.; Engstrom, O.; Ljunggren, J.; Gustafsson, J.; Carlquist, M. Structure of the Ligand-binding Domain of Oestrogen Receptor beta in the Presence of a Partial Agonist and a Full Antagonist. *EMBO J.* **1999**, *18*, 4608−4618.

(8) Renaud, J.; Bischoff, S. F.; Buhl, T.; Floersheim, P.; Fournier, B.; Halleux, C.; Kallen, J.; Keller, H.; Schlaeppi, J. M.; Stark, W. Estrogen Receptor Modulators: Identification and Structure-Activity Relationships of Potent ERα-Selective Tetrahydroisoquinoline Ligands. *J. Med. Chem.* **2003**, *46*, 2945−2957.

(9) Goldstein, S. R.; Siddhanti, S.; Ciaccia, A. V.; Plouffe, L., Jr. A Pharmacological Review of Selective Oestrogen Receptor Modulators. *Hum. Reprod. Update* **2000**, *6*, 212−224.

(10) Burger, H. G. Selective Oestrogen Receptor Modulators. *Horm. Res.* **2000**, *53*, 25−29.

(11) Henke, B. R.; Consler, T. G.; Go, N.; Hale, R. L.; Hohman, D. R.; Jones, S. A.; Lu, A. T.; Moore, L. B.; Moore, J. T.; Orband-Miller, L. A.; Robinett, R. G.; Shearin, J.; Spearing, P. K.; Stewart, E. L.; Turnbull, P. S.; Weaver, S. L.; Williams, S. P.; Wisely, G. B.; Lambert, M. H. A New Series of Estrogen Receptor Modulators that Display Selectivity for Estrogen Receptor β. *J. Med. Chem.* **2002**, *45*, 5492−5505.

(12) Grese, T. A.; Dodge, J. A. Selective Estrogen Receptor Modulators (SERMs). *Curr. Pharm. Des.* **1998**, *4*, 71−92.

(13) Lin, X.; Huebner, V. Non-steroidal Ligands for Steroid Hormone Receptors. *Curr. Opin. Drug Discovery Dev.* **2000**, *3*, 383−398.

(14) McDonnell, D. P. The Molecular Pharmacology of SERMs. *Trends Endocrinol. Metab.* **1999**, *10*, 301−311.

(15) Jordan, V. C. Antiestrogens and Selective Estrogen Receptor Modulators as Multifunctional Medicines. 2. Clinical Considerations and New Agents. *J. Med. Chem.* **2003**, *46*, 1081−1111.

(16) Mukherjee, S.; Mukherjee, A.; Saha, A. Predicting Pharmacophore Signals for Post-Coital Antifertility Activity of 1-Trifluoromethyl-1, 2, 2-Triphenylethylene Derivatives: A Statistical Approximation Using E-state Index. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 897−900.

(17) Mukherjee, S.; Mukherjee, A.; Saha, A. Exploring Pharmacophore Signals for Uterotrophic Activity of Bromotriphenylethylene Derivatives. *Internet Electron. J. Mol. Des.* **2005**, *4*, 270−278.

(18) Mukherjee, S.; Mukherjee, A.; Saha A. QSAR Studies with E-state Index: Predicting Pharmacophore Signals for Estrogen Receptor Binding Affinity of Triphenylacrylonitriles. *Biol. Pharm. Bull.* **2005**, *28*, 154−157.

(19) Suzuki, T.; Ide, K.; Ishida, M.; Shapiro, S. Classification of Environmental Estrogens by Phyisicochemical Principal Component Analysis and Hierarchical Cluster Analysis. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 718−726.

(20) Famini, G. R.; Wilson, L. Y. Using Theoritical Descriptors in Linear Solvation Energy Relationships. *Theor. Comput. Chem.* **1994**, *1*, 213−241.

(21) Fink, B. E.; Mortensen, D. S.; Stauffer, S. R.; Aron, Z. D.; Katzenellenbogen, J. A. Novel Structural Templates for Estrogen Receptor Ligands and Prospects for Combinatorial Synthesis of Estrogens. *Chem. Biol.* **1999**, *6*, 205−219.

(22) da Cunha, E. F. F.; Martins, R. C. A.; Albuquerque, M. G.; de Alencastro, R. B. LIV-3D-QSAR Model for Estrogen Receptor Ligands. *J. Mol. Model.* **2004**, *10*, 297−304.

(23) Fang, H.; Tong, W.; Shi, L. M.; Blair, R.; Perkins, R.; Branham, W.; Hass, B. S.; Xie, Q.; Dial, S. L.; Moland, C. L.; Sheehan, D. M. Structure-Activity Relationship for a Large Diverse Set of Natural, Synthetic and Environmental Estrogens. *Chem. Res. Toxicol.* **2001**, *14*, 280−294.

(24) Mukherjee, S.; Saha, A.; Roy, K. QSAR of Estrogen Receptor Modulators Exploring Selectivity Requirements for ER (alpha) versus ER (beta) Binding of Tetrahydroquinoline Derivatives Using E-state and Physicochemical Parameters. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 957−961.

(25) Schmidt, J. M.; Mercure, J.; Tremblay, G. B.; Page, M.; Feher, M.; Dunn-Dufault, R.; Peter, M. G.; Redden, P. R. De Novo Design, Synthesis and Evaluation of a Non-Steroidal Diphenylnapthyl Propylene Ligand for the Estrogen Receptor. *Bioorg. Med. Chem.* **2003**, *11*, 1389−1396.

(26) Kristam, R.; Gillet, V. J.; Lewis, R. A.; Thorner, D. Comparison of Conformational Analysis Techniques to Generate Pharmacophore Hypotheses Using Catalyst. *J. Chem. Inf. Model.* **2005**, *45*, 461−476.

(27) Guner, O. F. In *Pharmacophore Perception, Development, and Use in Drug Design;* International University Line: California, 2000.

(28) Mortensen, D. S.; Rodriguez, A. L.; Carlson, K. E.; Sun, J.; Katzenellenbogen, B. S.; Katzenellenbogen, J. A. Synthesis and Biological Evaluation of a Novel Series of Furans: Ligands Selective for Estrogen Receptor α. *J. Med. Chem.* **2001**, *44*, 3838−3848.
(29) Henke, B. R.; Consler, T. G.; Go, N.; Hale, R. L.; Hohman, D. R.; Jones, S. A.; Lu, A. T.; Moore, L. B.; Moore, J. T.; Orband-Miller, L. A.; Robinett, R. G.; Shearin, J.; Spearing, P. K.; Stewart, E. L.; Turnbull, P. S.; Weaver, S. L.; Williams, S. P.; Wisely, G. B.; Lambert, M. H. A New Series of Estrogen Receptor Modulators that Display Selectivity for Estrogen Receptor *β*. *J. Med. Chem.* **2002**, *45*, 5492−5505.
(30) Zong-Ru, G.; Xiang, Y.; Zhi-Bin, X. Molecular Simulation of Interactions between Estrogen Receptor and Selective Estrogen Receptor Modulators. *Acta Pharmacol. Sin.* **2002**, *23*, 208−212.
(31) Schuster, D.; Laggner, C.; Steindl, T. M.; Palusczak, A.; Hartmann, R. W.; Langer, T. Pharmacophore Modeling and in Silico Screening for New P450 19 (Aromatase) Inhibitors. *J. Chem. Inf. Model.* **2006**, *46*, 1301−1311.
(32) Lopez-Rodrýguez, M. L.; Benhamu, B.; de la Fuente, T.; Sanz, A.; Pardo, L.; Campillo, M. A Three-Dimensional Pharmacophore Model for 5-Hydroxytryptamine6 (5-HT6) Receptor Antagonists. *J. Med. Chem.* **2005**, *48*, 4216−4219.
(33) Taha, M. O.; Qandil, A. M.; Zaki, D. D.; AIDamen, M. A. Ligand-based Assessment of Factor Xa Binding Site Flexibility via Elaborate Pharmacophore Exploration and Genetic Algorithm based QSAR Modeling. *Eur. J. Med. Chem.* **2005**, *40*, 701−727.
(34) *Catalyst, version 4.11*; Accelrys Inc.: San Diego, CA, 2005.
(35) Barnum, D.; Greene, J.; Smellie, A.; Sprague, P. Identification of Common Functional Configurations Among Molecules. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 563−571.
(36) Jones, G.; Willett, P.; Glen, R. C. A Genetic Algorithm for Flexible Molecular Overlay and Pharmacophore Elucidation. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 532−549.
(37) Martin, Y. C.; Bures, M. G.; Danaher, E. A.; DeLazzer, J.; Lico, I.; Pavlik, P. A. A Fast New Approach to Pharmacophore Mapping and its Application to Dopaminergic and Benzodiazepine Agonists. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 83−102.
(38) Zhou, H. B.; Comninos, J. S.; Stossi, F.; Katzenellenbogen, B. S.; Katzenellenbogen, J. A. Synthesis and Evaluation of Estrogen Receptor Ligands with Bridged Oxabicyclic Cores Containing a Diarylethylene Motif: Estrogen Antagonists of Unusual Structure. *J. Med. Chem.* **2005**, *48*, 7261−7274.
(39) Muthyala, R. S.; Sheng, S.; Carlson, K. E.; Katzenellenbogen, B. S.; Katzenellenbogen, J. A. Bridged Bicyclic Cores Containing a 1,1-Diarylethylene Motif are High-Affinity Subtype-Selective Ligands for the Estrogen Receptor. *J. Med. Chem.* **2003**, *46*, 589−1602.
(40) *CS Chem3D Pro and CS MOPAC Pro, version 5.0*; Cambridge Soft, Corporation: Cambridge, U.S.A.
(41) Kier, L. B.; Hall, L. H. An Electrotopological-state Index for Atoms in Molecules. *Pharm. Res.* **1990**, *7*, 801−807.
(42) *ETSA-CS, Cyber-Mate*; Srirampur, Hooghly, W. Bengal, India.
(43) *TSAR 3D, version 3.3*; Oxford Molecular Limited: Accelrys.
(44) *Statistica, version 5.0*; StatSoft Inc.: Tulsa, U.S.A.
(45) Wold, S.; Eriksson, L. Chemometric Methods in Molecular Design. In *Chemometric Methods*; de Waterbeemd, H. V., Ed.; VCH Verlagsgesellschaft mbh: Weinheim, 1995; Vol. 2, Chapter 5, p 312.
(46) Barnum, D.; Greene, J.; Smellie, A.; Sprague, P. Identification of Common Functional Configurations Among Molecules. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 563−571.
(47) Snedecor, G. W.; Cochran, W. G. In *Statistical Methods*, 6th ed.; Iowa State University Press: Iowa, 1967.
(48) Jordan, V. C.; Mittal, S.; Gosdon, B.; Koch, R.; Lieberman, M. E. Structure-activity Relationships of Estrogens. *Environ. Health Perspect.* **1985**, *61*, 97−110.

CI600419S