Three-Dimensional Quantitative Structure−Activity Relationship Study of Nonsteroidal Estrogen Receptor Ligands Using the Comparative Molecular Field Analysis/Cross-Validated r^2-Guided Region Selection Approach

Beth R. Sadler,[†,‡] Sung Jin Cho,[§] Khalid S. Ishaq,[*,‡] Kun Chae,[†] and Kenneth S. Korach[†]

Laboratory of Reproductive and Developmental Biology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709, Division of Medicinal Chemistry and Natural Products, School of Pharmacy, and The Laboratory for Molecular Modeling, Division of Medicinal Chemistry and Natural Products, School of Pharmacy, University of North Carolina, Chapel Hill, North Carolina 27599-7360

Received August 20, 1997

A newly developed comparative molecular field analysis (CoMFA) technique, the cross-validated r^2-guided region selection (CoMFA/q^2-GRS) method, has been used to build a quantitative structure−activity relationship (3D-QSAR) for nonsteroidal estrogen receptor (ER) ligands. Ligands included in this study belong to a series of diethylstilbestrol (DES) and indenestrol analogues whose affinities for the mouse ER (mER) have been determined in our laboratory. The final model utilized 30 compounds and yielded a q^2_{GRS} (cross-validated r^2, guided region selection) of 0.796, as compared to a q^2 of 0.720 for conventional CoMFA, with a standard error of prediction of 0.594 at 3 principal components. This model was used to visualize steric and electrostatic features of the ligands that correspond with ER binding affinity. Results obtained from the CoMFA steric and electrostatic plots of this model have also been compared to information from the ER binding affinities of substituted estradiol analogues. This is in an effort to determine structural features of compounds in the CoMFA analysis that may correspond to those of the estradiol analogues and to further clarify the mode of binding of nonsteroidal ER ligands.

Introduction

The estrogen receptor (ER) belongs to the steroid/thyroid nuclear receptor superfamily whose members act as transcriptional activators via a direct interaction with DNA sequences termed response elements.[1] Ligand (e.g., estradiol) binding to the ER induces a conformational change in the receptor important for the association of the receptor−DNA complex with transcriptional coactivators and the transcriptional components of the cell.[2] This association then culminates in the synthesis of estrogen-responsive genes and an estrogen-induced cell and/or tissue response.

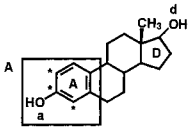
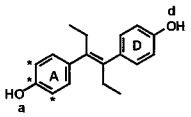
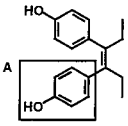
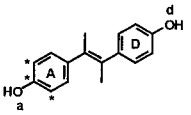
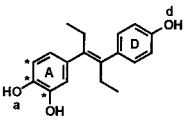
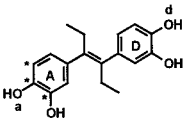
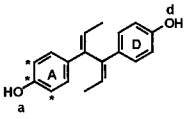
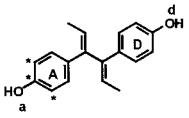
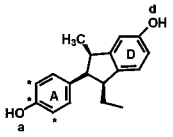
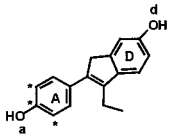
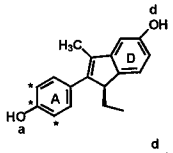
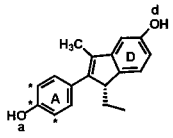
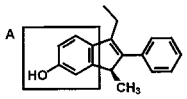
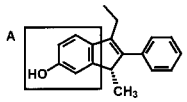
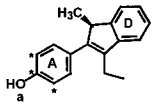
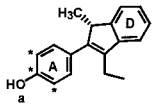
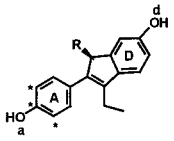
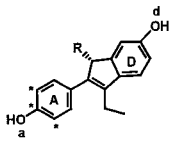
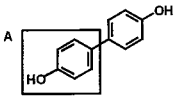
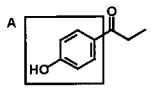
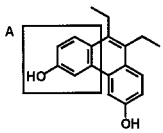
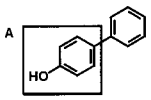
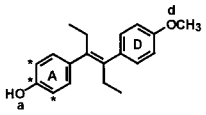
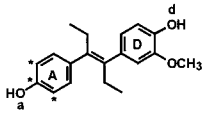
Because the natures of these responses depend on a specific ligand binding to the receptor,[3] it is important to examine ligand affinities for the ER and the structural characteristics of these ligands involved in receptor binding. This information is valuable in the design of new ER ligands from particular structural classes. The ER binding affinities of a number of diethylstilbestrol (DES, **1**; Table 1) metabolites and analogues have been reported.[4] Metabolism of DES, a potent nonsteroidal estrogen, results in products that retain a range of abilities to bind to the ER. These metabolites include indenestrols A and B, which have high ER affinities equivalent to that of the endogenous estrogen estradiol.[5] Various modes of binding can be proposed for many of these nonsteroidal compounds, and the lack of a resolved crystallographic structure of the ER ligand binding domain also hinders the identification of ligand orientation within the binding pocket. Structure−activity relationship studies can identify ligand substituents needed for high-affinity binding for the ER. Studies with indenestrol A (IA, **17**, the *S* enantiomer) analogues have indicated a requirement for both phenolic hydroxyl groups for maximum ER affinity.[6] Also, a hypothetical mode of binding has indicated that the 2-phenolic ring of IA mimics the A ring of estradiol in the ligand binding site.[7,8] An analysis of affinity data from these nonsteroidal ligands may help to confirm the proposed binding mode and model the ER binding site as it interacts with these ligands.

A three-dimensional quantitative structure−activity relationship (3D-QSAR) model of the ER, built with the validated technique of comparative molecular field analysis (CoMFA)[9] and using binding affinity data from known ligands, should be useful to interpret results from affinity studies and relate this information to the design of new ligands. The CoMFA method has been used previously to develop 3D-QSAR estrogen receptor models with estradiol derivatives,[10] polychlorinated biphenyls,[11] and recently a diverse set of environmental estrogens.[12] In this paper, we report the development of a 3D-QSAR model with 30 ER ligands, tested in our laboratory for receptor binding affinity, using the recently described CoMFA/q^2-GRS routine which was developed to reduce variability within the CoMFA routine.[13] This analysis includes DES metabolites, indenestrol analogues, and several other ER ligands in order to obtain information on ligand−receptor interac-

* To whom correspondence should be addressed.
† National Institutes of Health.
‡ Division of Medicinal Chemistry and Natural Products.
§ The Laboratory for Molecular Modeling.

S0022-2623(97)00552-9 CCC: $15.00
Published on Web 05/23/1998

Table 1. Structures of Compounds Included in CoMFA Study

Compound	Structure[a]	RBA[b]	Ref.
1		100	24
2		286	24
3		0.8	4
4		33	K.S. Korach, personal communication
5		100	K.S. Korach, personal communication
6		25	K.S. Korach, personal communication
7		0.3	4
8		20	4
9		2	4
10		14	24
11		100	5
12		142	5
13		1.8	6
14		0.2	6
15		5.6	6
16		0.9	6
17-20		17 R = CH_3 287 18 R = CH_2CH_3 292 19 R = $CH_2CH_2CH_3$ 229 20 R = $CH_2CH_2CH_2CH_3$ 178	24 24 c c
21-24		21 R = CH_3 13 22 R = CH_2CH_3 11 23 R = $CH_2CH_2CH_3$ 18 24 R = $CH_2CH_2CH_2CH_3$ 8	24 24 c c
25		0.02	33
26		0.1	4
27		0.16	33
28		0.02	32
29		20	K.S. Korach, personal communication
30		10	K.S. Korach, personal communication

[a] Alignments are explained in the text. [b] RBA as defined in Biological Activity Data section. [c] Unpublished data.

tions that lead to either an increase or decrease in ER affinity for this group of compounds. The final model resulted in a q^2_{GRS} (cross-validated r^2, guided region selection) of 0.796 and a standard error of prediction of 0.594 at 3 principal components.

Computational Details

SYBYL Molecular Modeling Software[14] (version 6.0) running on an IBM RISC6000 model 340 workstation was used for structure generation and CoMFA procedures. Optimization of structures and field-fit minimizations were performed using

the standard Tripos force field[15] with the maximum iteration cutoff of 1000 steps. Charges were calculated using the Gasteiger−Huckel method as implemented in SYBYL. The SYBYL systematic search method (10° increment and energy option on with electrostatic and maximum energy difference of 0.1 kcal/mol) was used to search for lowest-energy conformations. Default settings were used for other options except where noted.

Biological Activity Data. The selected compounds (Table 1) were tested for their affinity for the mouse estrogen receptor using an assay that has been previously described.[16] The activity data was expressed as relative binding affinities (RBA) relative to estradiol which is set to 100. These numbers were then transformed to the natural log values of the RBA (log RBA) which were used in the CoMFA/q^2-GRS studies.

Structure Generation and Alignment Rules. Compounds in the set were generated from the X-ray crystal structures or by modification of the crystal structure of a similar compound using the SYBYL BUILD option. X-ray crystal data were available for compounds **1**,[17] **2**,[18] **8**,[19] and **9**, **11**, and **17**.[7] Compounds **4**−**6**, **29**, and **30** were generated from the crystal structure of compound **2**. Compound **12** was generated from the crystal structure of compound **11** which is the enantiomer of compound **12**. Compounds **10**, **13**−**16**, and **18**−**24** were generated from the crystal structure of compound **17**. Remaining compounds were generated from the BUILD option in SYBYL. All compounds were geometrically optimized, and the lowest-energy conformation of all compounds was searched for by SYBYL systematic search applied to all rotatable bonds. After the conformational search, all compounds were then reoptimized.

Estradiol analogues where the 3-OH is deleted have a lower affinity for the receptor than analogues where the 17-OH has been deleted.[20] However, both analogues have a lower affinity for the receptor than estradiol. It has been proposed that the 3-OH of estradiol is important for initiating high-affinity ER binding while the 17-OH is important for estrogenic activity.[21] Because of the importance of the hydroxyls for ER affinity, they were used for alignment of the compounds in this exercise with estradiol. Two major alignments were generated for comparison via CoMFA analysis. In both alignments compound **1**, estradiol, the endogenous ligand for the ER, was chosen as the template molecule because of its high affinity for the receptor and its structural rigidity. Centroids were defined in the A and D rings of **1** and in the two aryl rings (A and D) of compounds **2**, **4**−**12**, **15**−**24**, **29**, and **30**. For alignment 1 (Table 1), these compounds were root mean square (rms)-fitted to compound **1** with the O-3 (a), the centroid of the A ring, the centroid of the D ring, and the O-17 (d) of **1** corresponding to the phenolic oxygen a, the centroid of aryl ring A, the centroid of aryl ring D, and the hydroxy or methoxy oxygen d of the compounds if present, respectively. The indicated (in box A) hydroxyaryl rings of compounds **3**, **13**, **14**, and **25**−**28** were rms-fitted to the A ring of compound **1** to allow for the importance of a phenol ring for receptor binding.

To obtain maximum steric and electronic overlap between these compounds and compound **1**, the field-fit option was applied to compounds generated in alignment 1. The structures of all compounds were then reminimized with the field-fit option turned off. This is referred to as alignment 2.

In alignment 3, compounds **2**, **4**−**12**, **15**−**24**, **29**, and **30** were rms-fitted to compound **1** as indicated in Table 1 where the C-2, C-3, C-4, O-3 (a), and O-17 (d) atoms of compound **1** were aligned with the respective starred atoms in other compounds. This fit has been previously reported[7] in a comparison of the crystal structures of compounds **1**, **2** and **17**, and this mode of binding orientation has also been proposed elsewhere. With this fit of compound **1** aligned with the respective highlighted atoms in compounds **2** and **17**, a good overall fit was obtained. As in alignment 1, the indicated hydroxyaryl rings of compounds **3**, **13**, **14**, and **25**−**28** were rms-fitted to the A ring of compound **1**. In alignment 4, the compounds in alignment 3 were field-fitted and reoptimized with the field-fit option turned off.

Conventional CoMFA. Conventional CoMFA was performed using the QSAR option in SYBYL. Default settings were used except where otherwise noted. Minimum σ was set to 2.0 for each cross-validated CoMFA analysis and set to 0.0 for each non-cross-validated analysis. The steric and electrostatic energies were calculated using sp^3 carbon probes with a +1 charge. The CoMFA grid spacing in all dimensions was 2.0 Å within the defined region which extended beyond the van der Waals envelopes of all molecules by at least 4.0 Å. The CoMFA QSAR equations were calculated within the partial least-squares (PLS) algorithm,[9] and the optimal number of components (ONC) in the final PLS model was determined from the standard error of prediction value, obtained from the leave-one-out cross-validation procedure; the number of components with the lowest standard error of prediction value was selected as the ONC.

CoMFA/q^2-GRS Routines. The q^2-GRS routine has been described in detail elsewhere[13] and applied to other data sets.[22,23] Acting on observations that values of q^2 in a CoMFA model are variable depending on the orientation on the computer screen of the aggregated molecules, Cho and Tropsha[13] developed a supplemental approach to CoMFA which attempts to reduce this variability. The CoMFA/q^2-GRS approach eliminates from the routine areas of space where steric and electronic fields do not correlate with changes in biological activity, i.e., individual regions that have low q^2 values. This allows for an optimal region selection for PLS analyses and gives as a result an orientation-independent q^2 which usually exceeds the value obtained with conventional CoMFA.

The steps in the CoMFA/q^2-GRS routine are as follows: (1) perform conventional CoMFA with automatically generated region file; (2) divide this region file encompassing the aligned molecules into 125 small region files using the Cartesian coordinates of the original region file to compute the coordinates of the 125 small boxes; (3) perform CoMFA with step size of 1.0 Å on each of these 125 small region files; (4) select the regions with a q^2 greater than or equal to a defined cutoff; (5) create a master region file with these cutoff regions included; (6) perform final CoMFA using master region file.

Results

CoMFA of ER Ligands. The CoMFA/q^2-GRS results on the estrogen receptor ligand set are presented in Table 2. Conventional CoMFA was initially run for all alignments; then the CoMFA/q^2-GRS was run to optimize the CoMFA model. For alignment 1, initial CoMFA produced a q^2 of 0.640 with 3 PCs (principal components). The highest q^2_{GRS} value (0.768; 2 PCs) was obtained using the q^2 cutoff value of 0.5.

Alignment 2, generated using the field-fit algorithm to alignment 1, gave a similar q^2 (0.647) for the conventional CoMFA to that of alignment 1 (0.640). The highest q^2_{GRS} value (0.744; 2 PCs) was obtained for the q^2 cutoff value of 0.3. This value is lower than the highest q^2_{GRS} obtained in alignment 1, the rms fit.

In alignment 3 the compounds were rms-fitted according to a published alignment[7,24] which fit selected compounds contained in this set. The q^2 obtained from the conventional CoMFA (0.720; 2 PCs) was significantly higher than that obtained from alignment 1 (0.640). The highest q^2_{GRS} value (0.796; 3 PCs), obtained with a cutoff value of 0.6, was an increase over the conventional CoMFA q^2 and also was higher than the comparable q^2_{GRS} calculated for alignment 1.

Alignment 4 was generated using the field-fit algorithm to alignment 3. The conventional q^2 obtained for alignment 4 (0.707; 1 PC) was lower than the conventional q^2 obtained for alignment 3. However, the maximum q^2_{GRS} obtained for alignment 4 (0.794; 6 PCs)

Table 2. q^2 Values Obtained Using the CoMFA/q^2-GRS Approach with Different q^2 Cutoff Values[a]

		no. of components					
	q^2 cutoff	1	2	3	4	5	6
alignment 1	none[b]	0.537	0.622	**0.640**	0.622	0.615	0.608
	0.1	0.678	**0.725**	0.722	0.686	0.679	0.676
	0.2	0.704	**0.745**	0.738	0.701	0.694	0.686
	0.3	0.713	**0.748**	0.741	0.708	0.703	0.697
	0.4	0.752	**0.764**	0.743	0.722	0.718	0.718
	0.5	0.748	**0.768**	0.746	0.729	0.745	0.763
alignment 2	none	0.628	**0.647**	0.601	0.548	0.542	0.537
	0.1	0.651	**0.708**	0.697	0.700	0.702	0.713
	0.2	0.652	**0.706**	0.695	0.703	0.700	0.719
	0.3	0.668	**0.724**	0.720	0.743	0.741	0.761
	0.4	0.682	**0.744**	0.746	0.735	0.758	0.765
	0.5	0.717	**0.719**	0.713	0.736	0.733	0.752
alignment 3	none	0.666	**0.720**	0.692	0.666	0.676	0.691
	0.1	0.655	0.744	**0.760**	0.765	0.760	0.765
	0.2	0.658	0.748	**0.765**	0.771	0.765	0.770
	0.3	0.655	0.750	**0.769**	0.775	0.766	0.768
	0.4	0.671	0.753	0.773	**0.785**	0.776	0.776
	0.5	0.699	0.765	0.771	**0.782**	0.774	0.777
	0.6	0.755	0.775	**0.796**	0.799	0.797	0.790
alignment 4	none	**0.707**	0.643	0.648	0.615	0.590	0.559
	0.1	**0.710**	0.705	0.725	0.719	0.732	0.738
	0.2	**0.713**	0.707	0.727	0.723	0.734	0.741
	0.3	0.722	0.730	**0.744**	0.736	0.750	0.752
	0.4	**0.722**	0.733	0.739	0.739	0.743	0.751
	0.5	0.726	**0.737**	0.738	0.730	0.729	0.736
	0.6	0.743	0.753	0.764	0.732	0.777	**0.794**

[a] The numbers in bold represent the q^2 values for the optimal number of components. [b] The results of conventional CoMFA.

Table 3. Summary of CoMFA/q^2-GRS Results: Alignment 3

optimal number of components	3
probe atom	C (sp^3, +1)
q^2	0.796
q^2-GRS cutoff	0.6
standard error of estimate	0.290
standard error of prediction	0.594
R^2	0.951
F values	169.830
probability of $R^2 = 0$ ($n_1 = 3$, $n_2 = 26$)	0.000
contributions	
steric	0.451
electrostatic	0.549

was very similar to the maximum of alignment 3 (0.796) but with a higher number of components (6 vs 3). The q^2 cutoff value which gave the maximum q^2_{GRS} for both alignments 3 and 4 was 0.6.

Among all alignments, alignment 3 produced the highest q^2_{GRS} (0.796, 3 PCs, 0.6 cutoff). The PLS run without cross-validation on this alignment yielded the conventional r^2 of 0.951 and a standard error of estimate of 0.290 (Table 3). The relative contribution for sterics is 0.451 versus 0.549 for electrostatics. This PLS model was used to calculate the activity of each compound, and this was compared with the actual value (Figure 1). The actual, calculated, and residual activities of all compounds are shown in Table 4.

CoMFA Fields. The CoMFA steric and electrostatic fields were obtained using an sp^3 carbon with a +1 charge under alignment 3. The field values were calculated by multiplying the β-coefficient and the standard deviation of columns in the QSAR table (stdev*coeff). Figure 2 shows the overlay of all 30 compounds in the CoMFA model rms-fitted to compound **1** (RBA = 100) as used in alignment 3. Clearly seen are the phenolic rings of all compounds superimposed to compound **1**. Figure 3 shows the CoMFA steric

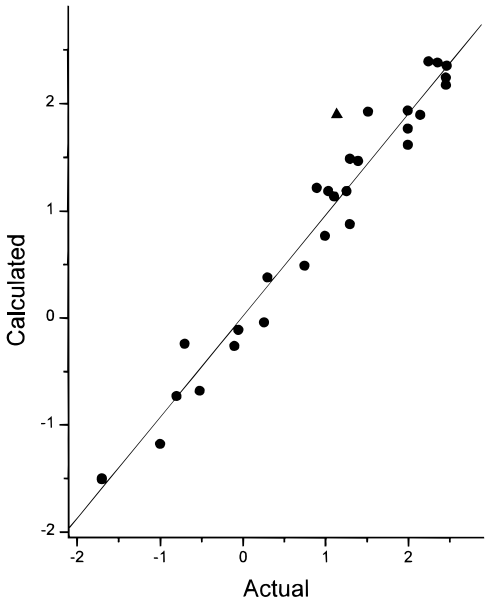


Figure 1. Actual vs calculated log RBA using alignment 3 (see Table 4). Compound **10** is indicated as ▲.

Table 4. CoMFA Actual, Calculated, and Residual Activities[a] for Alignment 3[b]

compd	actual[a]	calculated	residual
1	2.00	1.77	0.23
2	2.46	2.25	0.21
3	−0.10	−0.26	0.16
4	1.52	1.93	−0.41
5	2.00	1.62	0.38
6	1.40	1.47	−0.07
7	−0.52	−0.68	0.16
8	1.30	0.88	0.42
9	0.30	0.38	−0.08
10	1.14	1.90	−0.76
11	2.00	1.94	0.06
12	2.15	1.90	0.25
13	0.26	−0.04	0.30
14	−0.70	−0.24	−0.46
15	0.75	0.49	0.26
16	−0.05	−0.11	0.06
17	2.46	2.18	0.28
18	2.47	2.36	0.11
19	2.36	2.39	−0.03
20	2.25	2.40	−0.15
21	1.11	1.14	−0.29
22	1.04	1.19	−0.15
23	1.26	1.19	0.07
24	0.90	1.22	−0.32
25	−1.70	−1.50	−0.20
26	−1.00	−1.18	0.18
27	−0.80	−0.73	−0.07
28	−1.70	−1.51	−0.19
29	1.30	1.49	−0.19
30	1.00	0.77	0.23

[a] Activities are expressed as log RBA as RBA is explained in Table 1. [b] See text for alignment rules.

contour plots of alignment 3 with compounds **7** (RBA = 0.3; red) and **8** (RBA = 20; blue) included. The green (sterically favorable) and yellow (sterically unfavorable) contours represent actual values of 0.006 and −0.003, respectively. Compounds **7** and **8** are isomers which have very different affinities for the ER. The active compound **8** extends into the green contour regions, while the less active compound **7** extends into the unfavorable yellow contour regions. The contours from the electrostatic field are shown in Figure 4.

The blue (positive charge favored) and red (negative

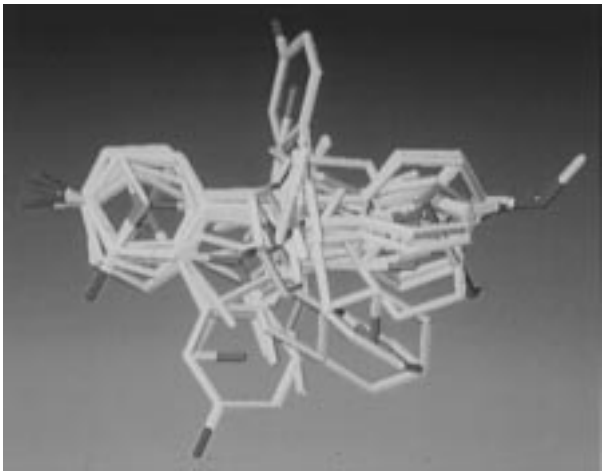


Figure 2. rms fit of all 30 compounds used for CoMFA model (alignment 3).

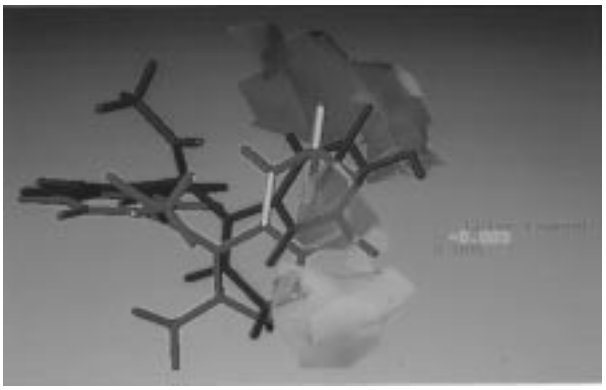


Figure 3. CoMFA steric contour plots (alignment 3) with compounds **7** (red) and **8** (blue). Green-shaded areas represent sterically favored regions, while yellow-shaded areas represent sterically unfavorable areas.

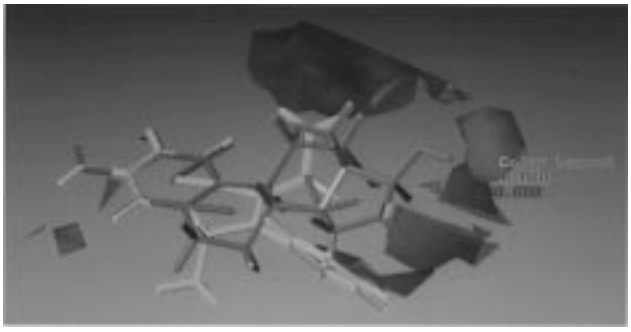


Figure 4. CoMFA electrostatic contour plots (alignment 3) with compounds **1** (green) and **14** (yellow). Blue-shaded areas represent positive-charge-favored regions, while red-shaded areas represent negative-charge-unfavored areas.

charge favored) contours represent 80% and 20% level contribution, respectively. The negative-charge-favored regions are clearly seen in the areas of the 3- and 17-OH's of compound **1** (green). Compound **14** (RBA = 0.2; yellow) is also included to represent analogues with no group comparable to the 17-OH of compound **1**.

Discussion

CoMFA has been used to develop 3D-QSAR equations for various receptors and enzymes. To reduce some of the variability of the CoMFA procedure which is due to overall orientation and lattice placement of the molecules, the CoMFA/q^2-GRS approach was developed[13] and has been utilized for ER ligands.

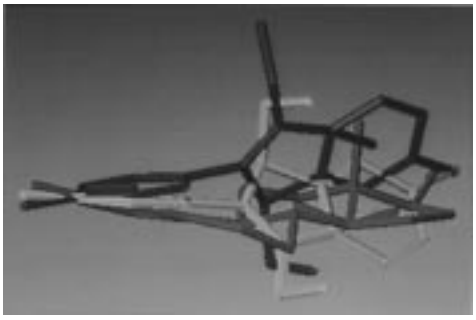


Figure 5. Alignment 3 rms fit of compounds **1** (red), **18** (yellow), and **22** (blue).

A number of different alignments were used in this CoMFA/q^2-GRS model. Among four alignments, alignments 3 and 4 yielded comparable values of q^2_{GRS} (Table 2). Alignment 3 was selected for examination because of its lower number of components and better overall statistics. For this alignment, GRS routine produced a q^2_{GRS} of 0.796 exceeding the q^2 of 0.720 obtained with conventional CoMFA. For all alignments, the q^2 obtained by q^2-GRS at the optimal number of components was higher than the q^2 obtained with conventional CoMFA. A lower q^2 for conventional CoMFA has been attributed to a possible poor orientation of the molecular aggregate on the screen since the q^2-GRS routine is orientation-independent.[13]

Figure 5 illustrates the alignment 3 rms-fit of compounds **1** (RBA = 100; red), **18** (RBA = 292; yellow), and **22** (RBA = 11; blue). Compounds **18** and **22** are enantiomers with the *S* enantiomer (**18**) having a higher binding affinity for the ER than the *R* enantiomer (**22**). Comparison of the rms fit of **18** and **22** to estradiol shows that **18** has a better overall fit to estradiol than **22**, which may explain its higher affinity for the receptor. The C-1 ethyl side chain of **18** extends into a similar region of space that would be occupied by an 11β-estradiol substituent. Previous examination of estradiol analogues with 11β-ethyl and -allyl substituents has shown that they possess a higher affinity for lamb ER than estradiol.[25] Others have also described the existence of a hydrophobic pocket in the 11β region of the receptor similar to that seen in this model.[26,27] Further, it has been proposed that the phenyl ring containing the dimethylaminoethoxy side chain of the antiestrogen tamoxifen also extends into this 11β region, supporting the presence of a binding pocket in this region of the receptor.[28] These observations are in agreement with the results from binding affinities of the indenesterol analogues **17**−**20** where the *S* orientation of the side chain places it in the 11β region, and they support the proposed mode of binding of these analogues. The C-1-unsubstituted IA analogue (**10**) with no side chain has a 20-fold lower binding affinity than IA-*S* (**17**), suggesting again that the C-1 substituent confers higher affinity binding for this series of compounds.[24]

The steric contour plot of alignment 3 is shown in Figure 3 with (*Z*,*Z*)-dienestrol (**7**, RBA = 0.3; red) and (*E*,*E*)-dienestrol (**8**, RBA = 20; blue). The higher affinity of **8** may be explained by its extension into the sterically favored (green) space of the receptor as determined by this CoMFA study. Compound **7** does not occupy this

space, but instead it extends into the inactive (yellow) sterically unfavorable region, correlating with its lower binding affinity. The orientation of the double bonds is important in the 3D orientation of each molecule, and the *E*,*E* isomer more closely fits into the active site as described by this model. From this model, it appears that substitutions that lead to hydrophobic bulk below the D ring of estradiol decrease ER affinity. Steric bulk intolerance in the D-ring region has also been noted in a CoMFA analysis of environmental estrogens.[12] Substitutions in the D ring of estradiol with alkylamide substituents (C-15 position) also lead to dramatic decreases in the affinities of these compounds for the ER.[29]

The electrostatic contour plot of alignment 3 is included with estradiol (**1**; green) and 4′-deoxy-IA (**13**, RBA = 0.2; yellow) in Figure 4. Clearly evident is the negative charges (red) favored in the regions of the 3-OH and 17-OH of estradiol. The decreased binding of **14** is probably due to its lack of a hydroxy group comparable to the 17-OH of estradiol. Previously it has been shown that both OH's of estradiol are necessary for high-affinity binding,[20] and this seems to be also true for this series of indenestrol compounds as exemplified by the affinity of compound **17** (RBA = 287) compared to that of its monohydroxylated analogues **13** (RBA = 1.8) and **15** (RBA = 5.6). Corresponding hydroxy substituents in the tissue-specific estrogen raloxifene have also been found to be important in receptor binding.[30] Comparison of the steric versus electrostatic contributions to binding affinity reveals a greater contribution of the latter to such activity (Table 3), indicating that electrostatic interactions between ligand and the receptor are important in affinity determination. The greater contribution from the electrostatic component in this model supports the observed requirement for two appropriately spaced hydroxy groups to mimic the A- and D-ring hydroxy groups of estradiol for highest receptor affinity for these compounds. This is also consistent with the observation that compounds from diverse structural classes bind to the estrogen receptor, but a common feature important for highest binding affinity is a phenolic ring mimicking the A ring of estradiol in the binding site.[31]

In summary, we have utilized the CoMFA/q^2-GRS approach to investigate the binding of a number of nonsteroidal ER ligands. The results of this analysis confirm the importance of hydroxy substituents which mimic the 3-OH and 17-OH of estradiol for high-affinity binding. The proposed mode of binding of the indenestrol analogues is also supported by the CoMFA model and affinity data from substituted estradiol analogues indicating the presence of a hydrophobic pocket in the 11β region of estradiol. Although the prediction of ER binding affinity for a ligand may not accurately indicate the biological activity of the compound in vivo, ER binding activity is required for activities of ER-directed therapies such as raloxifene and its analogues.[30] The determination of ligand characteristics leading to ER affinity is important for the development of therapeutic agents that exert their actions via the ER in a general and tissue-specific manner because receptor binding is a necessary requisite for stimulation of biological activity.

References

(1) Manglesdorf, D. J.; Thummel, C.; Beato, M.; Herrlich, P.; Shutz, G.; Umesono, K.; Blumberg, B.; Kastner, P.; Mark, M.; Charmbon, P.; Evans, R. M. The nuclear receptor superfamily: the second decade. *Cell* **1995**, *83*, 835−839.

(2) Beekman, J. M.; Allan, G. F.; Tsai, S. Y.; Tsai, M.-J.; O'Malley, B. W. Transcriptional activation by the estrogen receptor requires a conformational change in the ligand binding domain. *Mol. Endocrinol.* **1993**, *267*, 19513−19520.

(3) Katzenellenbogen, J. A.; O'Malley, B. W.; Katzenellenbogen, B. S. 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.* **1996**, *10*, 119−131.

(4) Korach, K. S. Biochemical and estrogenic activity of some diethylstilbestrol metabolites and analogues in the mouse uterus. In *Hormones and Cancer*; Leavitt, W., Ed.; Plenum: New York, 1982; pp 39−62.

(5) Korach, K. S.; Chae, K.; Levy, L. A.; Duax, W. L.; Sarvar, P. J. Diethylstilbestrol metabolites and analogues: stereochemical probes for the estrogen receptor binding site. *J. Biol. Chem.* **1989**, *264*, 5642−5647.

(6) Chae, K.; Gibson, M.; Korach, K. S. Estrogen receptor stereochemistry: ligand binding orientation and influence on biological activity. *Mol. Pharmacol.* **1991**, *40*, 806−811.

(7) Duax, W. L.; Swenson, D. C.; Strong, P. D.; Korach, K. S.; McLachlan, J.; Metzler, M. Molecular structures of metabolites and analogues of diethylstilbestrol and their relationship to receptor binding and biological activity. *Mol. Pharmacol.* **1984**, *26*, 520−525.

(8) Anstead, G. M.; Wilson, S. R.; Katzenellenbogen, J. A. 2-Arylindenes and 2-arylindenones: molecular structures and considerations in the binding orientation of unsymmetrical nonsteroidal ligands to the estrogen receptor. *J. Med. Chem.* **1989**, *32*, 2163−2171.

(9) Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959−5967.

(10) Gantchev, T. G.; Ali, H.; van Lier, J. E. Quantitative structure−activity relationships/comparative molecular field analysis (QSAR/CoMFA) for receptor-binding properties of halogenated estradiol derivatives. *J. Med. Chem.* **1994**, *37*, 4164−4176.

(11) Waller, C.; Minor, D. L.; McKinney, J. D. Using three-dimensional quantitative structure−activity relationships to examine estrogen receptor binding affinities of polychlorinated biphenyls. *Environ. Health Perspect.* **1995**, *103*, 702−707.

(12) Waller, C.; Oprea, T. I.; Chae, K.; Park, H.-K.; Korach, K. S.; Laws, S. C.; Wiese, T. E.; Kelce, W. R.; Gray, L. E., Jr. Ligand-based identification of environmental estrogens. *Chem. Res. Toxicol.* **1996**, *9*, 1240−1248.

(13) Cho, S. J.; Tropsha, A. Cross-validated R^2 guided region selection for comparative molecular field analysis (CoMFA): a simple method to achieve consistent results. *J. Med. Chem.* **1995**, *38*, 1060−1066.

(14) Tripos Inc., 1699 S. Hanley Rd., Suite 303, St. Louis, MO 63144.

(15) Clark, M.; Cramer, R. D., III; Van Opdenbosch, N. Validation of the general purpose Tripos 5.2 force field. *J. Comput. Chem.* **1989**, *10*, 982−1012.

(16) Korach, K. S. Estrogen action in the mouse uterus: characterization of cytosol and nuclear receptor systems. *Endocrinology* **1979**, *104*, 1324−1332.

(17) Duax, W. L. The structure of the crystalline complex estradiol urea (1:1). *Acta Crystallogr.* **1972**, *B28*, 1864−1871.

(18) Busetta, P. B.; Hospital, C. C. Etude cristallographique de differentes formes solvatees du diethylstilboestrol. (Crystallographic study of different solvated forms of diethylstilbestrol.) *Acta Crystallogr.* **1973**, *B29*, 2456−2462.

(19) Doyle, T. D.; Stewart, J. M.; Filipescu, N.; Benson, W. R. Configuration of dienestrol. *J. Pharm. Sci.* **1975**, *64*, 1525−1528.

(20) Chernyaev, G. A.; Barkova, T. I.; Egorova, V.; Sorokina, I. B.; Anachenko, S. N.; Matarodze, G. D.; Sokolova, N. A.; Rozen, V. B. A series of optical structural and isomeric analogues of estradiol: a comparative study of the biological activity and affinity to cytosol receptor of rabbit uterus. *J. Steroid Biochem.* **1975**, *6*, 1483−1488.

(21) Duax, W. L.; Weeks, C. M. Molecular basis of estrogenicity: X-ray crystallographic studies. In *Estrogens in the Environment*; McLachlan, J., Ed.; Elsevier/North-Holland: New York, 1980; pp 111−131.

(22) Cho, S. J.; Tropsha, A.; Suffness, M.; Cheng, Y.-C.; Lee, K.-H. Antitumor Agents. 163. Three-dimensional quantitative structure−activity relationship study of 4′-O-demethylepipodophyllotoxin analogues using the modified CoMA/q^2-GRS approach. *J. Med. Chem.* **1996**, *39*, 1383−1395.

(23) Cho, S. J.; Garsia, M. L. S.; Bier, J.; Tropsha, A. Structure-based alignment and comparative molecular field analysis of acetylcholinesterase inhibitors. *J. Med. Chem.* **1996**, *39*, 5064−5071.

(24) Korach, K. S.; Chae, K.; Gibson, M.; Curtis, S. Estrogen receptor stereochemistry: ligand binding and hormonal responsiveness. *Steroids* **1991**, *56*, 263−270.
(25) Napolitano, E.; Fiaschi, R.; Carlson, K. E.; Katzenellenbogen, J. A. 11-Beta substituted estradiol derivatives − potential high-affinity carbon-11-labeled probes for the estrogen receptor: a structure−activity-relationship. *J. Med. Chem.* **1995**, *38*, 429−434.
(26) Leclercq, G.; Heuson, J. C. Drug interaction with estrogen receptors for the control of breast neoplasia (review). *Anticancer Res.* **1981**, *1*, 217−228.
(27) Poupaert, J. H.; Lambert, D. M.; Vamecq, J.; Abul-Hajj, Y. J. Molecular modeling studies on 11β-aminoethoxyphenyl and 7α-aminoethoxyphenyl estradiols. evidence suggesting a common hydrophobic pocket in estrogen receptor. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 839−842.
(28) Duax, W. L.; Griffin, J. F.; Rohrer, D. C.; Swenson, D. C.; Weeks, C. M. Molecular details of receptor binding and hormonal action of steroids derived from X-ray crystallographic investigations. *J. Steroid Biochem.* **1981**, *15*, 41−47.
(29) Poirier, D.; Merand, Y.; Labrie, C.; Labrie, F. D-ring alkylamide derivatives of estradiol: effect on ER-binding affinity and antiestrogenic activity. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2537−2542.
(30) Grese, T. A.; Cho, S.; Finely, D. R.; Godfrey, A. G.; Jones, C. D.; Lugar, C. W., III; Martin, M. J.; Matsumoto, K.; Pennington, L. D.; Winter, M. A.; Adrian, M. D.; Cole, H. W.; Magee, D. E.; Phillips, D. L.; Rowley, E. R.; Short, L. L.; Glasebrook, A. L.; Bryant, H. U. Structure−activity relationships of selective estrogen receptor modulators: modifications to the 2-arylbenzothiophene core of raloxifene. *J. Med. Chem.* **1997**, *40*, 146−167.
(31) Duax, W. L.; Griffen, J. F.; Weeks, C. M.; Wawrzak, Z. The mechanism of action of steroid antagonists: insights from crystallographic studies. *J. Steroid Biochem.* **1988**, *31*, 481−492.
(32) Korach, K. S.; Sarver, P.; Chae, K.; McLachlan, J. A.; McKinney, J. Estrogen receptor binding activity of polychlorinated hydroxybiphenyls: conformationally restricted structural probes. *Mol. Pharmacol.* **1987**, *33*, 120−126.
(33) Korach, K. S.; Metzler, M.; McLachlan, J. A. Estrogenic activity in vivo and in vitro of some diethylstilbestrol metabolites and analogues. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 468−471.

JM9705521