Journal of Computer-Aided Molecular Design, 10 (1996) 186–200
ESCOM

J-CAMD 341

Three-dimensional quantitative structure–activity relationships of steroid aromatase inhibitors

Tudor I. Oprea* and Angel E. García

Theoretical Biology and Biophysics Group (T-10), Los Alamos National Laboratory, MS K710, Los Alamos, NM 87545, U.S.A.

Received 12 September 1995
Accepted 1 February 1996

Keywords: 3D QSAR methodology; Aromatase inhibition; CoMFA; GOLPE; Steroids; XED

Summary

Inhibition of aromatase, a cytochrome P450 that converts androgens to estrogens, is relevant in the therapeutic control of breast cancer. We investigate this inhibition using a three-dimensional quantitative structure–activity relationship (3D QSAR) method known as Comparative Molecular Field Analysis, CoMFA [Cramer III, R.D. et al., J. Am. Chem. Soc., 110 (1988) 5959]. We analyzed the data for 50 steroid inhibitors [Numazawa, M. et al., J. Med. Chem., 37 (1994) 2198, and references cited therein] assayed against androstenedione on human placental microsomes. An initial CoMFA resulted in a three-component model for $\log(1/K_i)$, with an explained variance r^2 of 0.885, and a cross-validated q^2 of 0.673. Chemometric studies were performed using GOLPE [Baroni, M. et al., Quant. Struct.–Act. Relatsh., 12 (1993) 9]. The CoMFA/GOLPE model is discussed in terms of robustness, predictivity, explanatory power and simplicity. After randomized exclusion of 25 or 10 compounds (repeated 25 times), the q^2 for one component was 0.62 and 0.61, respectively, while r^2 was 0.674. We demonstrate that the predictive r^2 based on the mean activity (Y_m) of the training set is misleading, while the test set Y_m-based predictive r^2 index gives a more accurate estimate of external predictivity. Using CoMFA, the observed differences in aromatase inhibition among C6-substituted steroids are rationalized at the atomic level. The CoMFA fields are consistent with known, potent inhibitors of aromatase, not included in the model. When positioned in the same alignment, these compounds have distinct features that overlap with the steric and electrostatic fields obtained in the CoMFA model. The presence of two hydrophobic binding pockets near the aromatase active site is discussed: a steric bulk tolerant one, common for C4, C6-alpha and C7-alpha substituents, and a smaller one at the C6-beta region.

Introduction

Naturally occurring estrogens (C_{18}-steroids) are formed in a multiple-step synthesis from androgens [1]. One of the key steps in their formation is the aromatization of the A-ring, with removal of the C19 methyl group attached at the 10-position [2]. Aromatase, the enzyme that catalyzes the conversion of androstenedione to estrone, of 16α-hydroxyandrostenedione to estriol and of testosterone to estradiol, is a cytochrome P450 complex [3,4]. Aromatase is present in the ovaries, the placenta, and, in postmenopausal women, in the adipose and muscular tissues [1]. Estrogen production via the aromatase pathway is important in the development and evolution of estrogen-dependent malign tumors (e.g., endometrial carcinoma and breast cancer in women) [5–8].

Aromatase inhibitors constitute a logical approach in the therapy of estrogen-dependent malignancies, since they block endogenous estrogen formation [8]. Such compounds are becoming an important asset in the arsenal against breast cancer [6,9]. Aminoglutethimide, a non-steroidal aromatase inhibitor, is an efficacious treatment for metastatic breast and prostate cancers [10]. Since it inhibits other pathways of steroid synthesis, ongoing research seeks more selective, potent inhibitors [11]. Current treatment of estrogen-dependent malignancies has focused on anti-estrogenic compounds such as tamoxifen [10], but recent clinical trials showed that 4-hydroxyandrostenedione is effective in tumors that relapse from tamoxifen [7]. For these reasons, the development of aromatase inhibitors is considered to be of major clinical importance [6].

*To whom correspondence should be addressed at: Department of Medicinal Chemistry, KJ3, Astra Hässle AB, S-43184 Mölndal, Sweden.

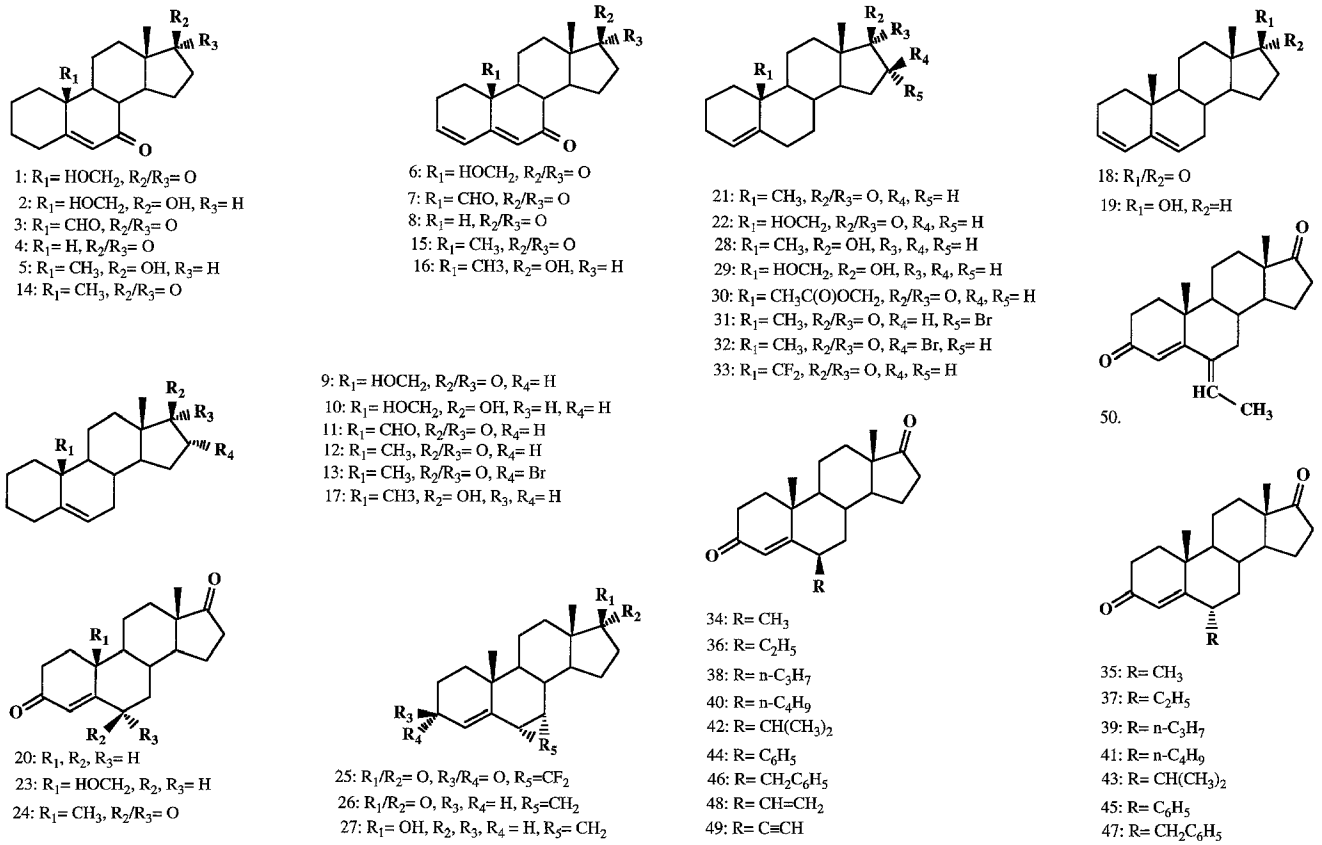

Fig. 1. The 50 steroid structures used to derive the CoMFA models.

Site-directed mutagenesis studies have identified amino acid residues that are important in steroid binding [12–18]. A structural model for aromatase was developed [19] based on the X-ray structure of $P450_{cam}$ – a class I cytochrome P450 from *Pseudomonas putida* [20]. This three-dimensional model for aromatase failed to correctly identify residues in the binding site [18]. While the exact structure of the aromatase binding site remains unknown, synthetic efforts in the design of aromatase inhibitors are based more on conventional methods and not on modern drug design strategies. Quantitative structure–activity relationship (QSAR) methods can extract information from existing experimental studies. This information can be used to improve existing receptor models for the binding site of aromatase, and can assist the evaluation of computer-designed compounds [21].

In this paper, we investigate the inhibition of human aromatase using a three-dimensional QSAR (3D QSAR) method [22], known as Comparative Molecular Field Analysis, CoMFA [23]. We analyzed the available data for 50 steroid inhibitors assayed against human placental aromatase and its natural substrate, androstenedione [24–28], and subsequently generated a CoMFA model. We discuss this model in terms of robustness, predictivity, explanatory power and simplicity. Using chemometric tools (GOLPE [29]), we show that change correlation [30] is unlikely in the resulting CoMFA model. We show that the often used training set Y_m-based predictive r^2 gives misleading results, while the test set Y_m-based predictive r^2 is more accurate in estimating external predictivity. We explain the variations in aromatase inhibition for several C6-substituted steroids, based on differences in molecular structure. The CoMFA fields are consistent with known, potent aromatase inhibitors that are not included in the model. Positioned in the same alignment, these external compounds overlap with the steric and electrostatic fields obtained in our CoMFA model. Our results support the existence of two binding pockets near the active site: a small one in the C6β region, favoring steric, but not electrostatic interactions, and a large one at the α-face of the C6-C7 region, which favors aromatic groups.

Computational Methods

The steroid aromatase inhibitor series

Reversible and irreversible inhibitors of human aromatase (analogs of androstenedione) have been characterized by Numazawa and co-workers [24–28]. They investigated the aromatase binding site by studying the effect of different substituents and/or unsaturated bonds on the binding affinity. These were substituents at C3 (keto or deoxy), C6 (α-, β- or unsaturated), C7 (α- or keto), C10 (hydro-

TABLE 1
SUMMARY OF AROMATASE INHIBITION BIOCHEMICAL ANALYSES

Screening assay procedure			Reference inhibitor		Reference
Substrate (μM)	Enzyme (μg)	K_m (nM)	IC_{50} (μM)	K_i (nM)	
1	40	57	0.53	36.5	24
1	40	60 ± 7	0.53	37	25
1	40	55 ± 5	0.53	37	26
1	20	18 ± 3	N.A.	N.A.	27
0.3	20	20	0.11	13	28

Substrate and enzyme amounts are listed as used in the screening assay procedure (0.5 ml final volume) and for determining K_m and IC_{50} for reference compound **21**, androsta-4-en-17-one (incubation time 20 min). K_i values reported for the reference inhibitor were obtained using a Dixon plot. N.A.: not available.

gen or methyl at C19), C16 (α- or β-bromine), C17 (keto or β-hydroxyl), and C19 (hydroxyl, aldo or acetyl), and unsaturated bonds at C3-C4, C5-C6 and C4-C5 (see Fig. 1 for structural details). Other important substituents, at C2, C4 and C7, used to derive potent aromatase inhibitors, are discussed for comparison in the Results.

Numazawa and co-workers determined the experimental activity according to the method of Thomson and Siiteri [4], by measuring the amount of tritiated water released by incubating [1β-^{3}H]androstenedione with placental microsomes. IC_{50} and K_i (from the initial screening assay) and K_I (time-dependent inactivation studies for suicidal inhibitors) values were obtained separately [27, 28]. Values for the apparent K_m of aromatase, summarized in Table 1, differ in the five Numazawa papers [24–28], because different amounts of the enzyme and of the tritiated substrate were used.

We selected the studied compounds based on two criteria: (i) screening studies of aromatase inhibition in the human placental microsomes assay that show competitive inhibition (IC_{50} and K_i); and (ii) exact chemical characterization. Several compounds were not included in this study, either because they were not shown to exhibit competitive inhibition (e.g., 19,19-difluoroandrost-4-en-17β-ol [25]), or because they were assayed as a racemic mixture (e.g., (*R/S*)-19-ethynyl-19-hydroxyandrost-4-en-17-one [25]). This precludes their use in CoMFA, because one of the underlying assumptions in this methodology is that the binding site is common for all compounds [23, 29], and because each compound included in the training set requires a well-defined chemical structure.

QSAR analyses are based on the formalism of linear free-energy relationships, where ΔG, the free energy of binding, is directly proportional with –logK, the binding affinity ($\Delta G \approx 1.4$ pK, in kcal/mol). Therefore, we transformed the biological activities of the 50 steroids as the $\log_{10}(1/K_i)$, expressed in μM (Table 2). We used the activity of androst-4-en-17-one, available for comparison in most studies [24–26,28], to determine the relative potency of the other compounds (separately for IC_{50} and K_i values). The logarithms of both absolute and relative IC_{50} and K_i values are also listed in Table 2.

For compounds **34–50**, the comparative activity of androst-4-en-17-one was not available [27]. We derived a Partial Least Squares (PLS) [30] regression model for compounds **1–33**, in which we used absolute pIC_{50} and pK_i to explain the relative pIC_{50} and pK_i values. We predicted relative potencies of compounds **34–50** based on this model.

We observed a strong numerical correlation between absolute and relative IC_{50} and K_i values for the initial 33-compound set, as well as between absolute IC_{50} and K_i values in the 50-compound set. Principal Component Analysis (PCA) [31] on these four activities showed that one component explains 97.6% of the variance (33 compounds). This confirms that these descriptors are not independent. There is a high correlation between experimentally determined pIC_{50} and pK_i values ($r^2 = 0.963$, $F = 722.98$, $s = 0.175$, $n = 33$; where r^2 is the explained variance, F is the fraction of explained versus unexplained variance, s is the standard error of the estimate, and n is the number of compounds), and between absolute and relative K_i values ($r^2 = 0.988$, $F = 3867.41$, $s = 0.133$, $n = 50$).

This interdependence prompted us to present only QSARs based on absolute K_i values, because: (i) these originate from experimental results; and (ii) both K_i descriptors have similar informative content. We prefer pK_i over pIC_{50} because the former is directly related to binding affinity, hence to the free energy of binding. QSAR results for absolute pIC_{50} and relative pIC_{50} and pK_i values are similar.

Results for the first 33 compounds (with relative potencies available [24–26,28]) are reported as a separate model, to which the other 17 compounds [27] were subsequently added. We discuss the similarities and differences between the 33- and 50-compound model systems in the section on the 50-compound model.

The CoMFA method

The CoMFA approach [23,29] is based on the following assumptions: (i) structural details of the common receptor binding site are not known for all (or most) of the compounds in the studied series; (ii) theoretical physicochemical aspects of the modeling of ligand–receptor

interactions, e.g. force-field parameterization and solvent effects, maintain a degree of uncertainty; and (iii) observed variations in experimental activity among different compounds may be explained with calculated physicochemical properties within a given series.

In CoMFA, steric (Lennard-Jones) and electrostatic (Coulombic) interactions of a ligand evaluated for probe atoms regularly spaced on a grid replace the noncovalent ligand–receptor interactions. The steric and electrostatic interactions of probe atoms with the ligand are calculated at grid points, then tabulated for each molecule (row) in the series. The resulting matrix is analyzed with multivariate statistics (PLS), yielding an equation that relates the CoMFA field values (X-block) to the activity (Y-block). This process highlights those features of the receptor that are investigated by the given structure–activity data set. The resulting QSAR coefficients (fields) are then graphically examined as contours in 3D space, and suggest areas for improvement in ligand design.

We used the following specifications for CoMFA runs: 2-Å spaced grid ($36 \times 26 \times 22$ Å); probe atom: Fe^{2+} (its larger van der Waals radius of 1.844 Å causes a tight overlap in the 2-Å lattice); a dielectric constant of 1.0; a ±20 kcal/mol cutoff value with no electrostatic interactions at steric overlap points (i.e., electrostatic field values are not considered at grid points overlapping with ligand atoms); field smoothing (average of field values surrounding the lattice point) using a sphere with a radius of 1.8 Å.

We used the SYBYL implementation [32] of the PLS algorithm [30] for regression analyses, initially with cross-validation (leave-one-out model; LOO), and five principal components. From the analysis with the highest cross-validated r^2 value (termed q^2, after Clementi [33]) we chose the optimal number of components. For models with very similar or identical q^2 values, we chose the model with the smallest standard deviation error of prediction (SDEP). We used a 0.5 kcal/mol energy column filter (minimum sigma) for all CoMFA (not GOLPE) calculations, to improve the signal-to-noise ratio.

The cross-validated r^2 (q^2) measures the robustness of the model [23], and is defined as:

$$q^2 = 1 - \text{PRESS/SD} \quad (1)$$

$$\text{PRESS} = \sum_{i}^{N} (Y_p - Y_a)^2 \quad (2)$$

$$\text{SD} = \sum_{i}^{N} (Y_a - Y_m)^2 \quad (3)$$

where PRESS is the sum of squared deviations between predicted (Y_p) and measured (actual, Y_a) biological activity values for all N molecules in the set, and SD is the sum of the squared deviations between measured biological activity values of the molecules in the test set and the mean (Y_m) activity value for all N molecules in the training set.

TABLE 2
ACTIVITIES OF THE COMPOUNDS INCLUDED IN THE TRAINING SET

Compound	pIC_{50}	pK_i	Relative pIC_{50}	Relative pK_i	Ref.
1	−1.7482	−1.0414	−2.7068	−2.9274	28
2	−2.1461	−1.6532	−3.1051	−3.5406	28
3	−1.7782	−1.1139	−2.7368	−3.0000	28
4	−1.4472	−0.3802	−2.4058	−2.2663	28
5	−1.6232	−0.7404	−2.5819	−2.6265	28
6	−1.7782	−1.1761	−2.7368	−3.0625	28
7	−1.0414	−0.2553	−2.0000	−2.1413	28
8	−1.5185	−0.4771	−2.4772	−2.3632	28
9	−0.8388	0.0000	−1.7974	−1.8861	28
10	−1.7924	−0.9912	−2.7510	−2.8775	28
11	−0.9542	−0.1461	−1.9129	−2.0322	28
12	0.1805	0.9208	−0.7781	−0.9653	28
13	−1.7404	−1.0414	−2.6990	−2.9274	28
14	−0.2553	0.6021	−1.2139	−1.2840	28
15	−0.2553	0.6576	−1.2139	−1.2284	28
16	−1.6232	−0.7243	−2.5818	−2.6103	28
17	−1.4771	−0.4771	−2.4358	−2.3632	28
18	0.4685	1.2366	−0.4901	−0.6495	28
19	−1.2553	−0.3010	−2.2139	−2.1871	28
20	−0.0792	0.8539	−1.0378	−1.0322	28
21	0.9586	1.8861	0.0000	0.0000	28
22	0.5686	1.9031	0.2929	0.4564	24
23	−0.6721	0.5961	−0.9478	−0.8416	24
24	0.3188	1.5870	0.0430	0.1493	24
25	−0.0607	1.3010	−0.3364	−0.1308	26
26	0.8239	2.3010	0.5482	0.8692	26
27	−0.4314	0.9208	−0.7071	−0.5110	26
28	−0.8633	0.0809	−1.1391	−1.3509	25
29	−0.3617	0.7696	−0.6375	−0.6627	25
30	−0.6021	0.5391	−0.8778	−0.8927	25
31	−0.6128	0.6421	−0.8885	−0.7897	25
32	−0.7709	0.3420	−1.0466	−1.0898	25
33	−0.0792	0.3536	−0.3549	−1.0782	25
34	0.1427	1.9586	−0.0562	0.5645	27
35	0.4559	2.2518	0.2768	0.8697	27
36	0.8539	2.8539	0.8162	1.5627	27
37	0.5850	2.3279	0.3913	0.9360	27
38	0.6198	2.3372	0.4165	0.9391	27
39	0.4089	2.1739	0.2096	0.7789	27
40	0.2924	2.0555	0.0810	0.6529	27
41	0.1024	1.9208	−0.0991	0.5251	27
42	−0.0414	1.6576	−0.3172	0.2139	27
43	−0.1461	1.5086	−0.4544	0.0442	27
44	−0.2304	1.4318	−0.5429	−0.0352	27
45	−0.0414	1.6778	−0.3070	0.2408	27
46	−0.6021	1.2007	−0.8788	−0.2422	27
47	0.1805	2.0000	−0.0130	0.6093	27
48	0.4949	2.2924	0.3203	0.9132	27
49	−0.2788	1.2076	−0.6850	−0.3202	27
50	0.4089	2.3098	0.2785	0.9595	27

Experimental IC_{50} (pIC_{50}) and K_i (pK_i) values are expressed as the $-\log_{10}$ of the micromolar concentration. Relative IC_{50} and K_i values are $\log_{10}$ values of the ratio between the activity of androst-4-en-17-one (**21**) and the compound's corresponding activity (calculated for compounds **34**–**50**; see text for details). References are given for each compound. For comparison, pK_I (μM) values derived from a kinetic assay [28] are given for the seven suicidal inhibitors: **1** = −1.041; **3** = −1.301; **4** = −0.415; **6** = −0.845; **7** = −0.462; **14** = 0.824; **15** = 0.744.

The same formula can be applied to calculate predictive r^2 (r^2_{pred} or q^2) values. In this case, Y_p and Y_m are derived from the test set. Because the nature of the Y_m value used in the calculation of r^2_{pred} is subject to discussion [34], we used both the external (test) set Y_m and the training set Y_m when calculating the predictive r^2. From its definition, the best predictive model yields a q^2 of 1, but values can go to zero (or below) for models that predict values equal to or worse than Y_m. Negative values imply that the width of the model distribution is larger than the standard deviation of the actual activity distribution.

A parameter that does not depend on the mean activity, SDEP, has been proposed [33]:

$$SDEP = \sqrt{\frac{\sum_{i}^{N}(Y_p - Y_a)^2}{N}} \quad \text{(for 1 model)} \qquad (4)$$

$$SDEP_{Tot} = \sum_{j}^{K}\frac{SDEP}{K} \quad \text{(for K models)} \qquad (5)$$

The use of SDEP has been extended from 1 model (Eq. 4) to all models derived during cross-validation ($SDEP_{Tot}$, Eq. 5) when the results of Eq. 4 are averaged over the K models computed during the process. In both cases, a lower SDEP value reflects a good predictive ability. The $SDEP_{Tot}$ index, referred to as SDEP in this paper, is used to compare various cross-validated CoMFA models. The standard deviation error of calculations, SDEC, is defined in a similar fashion [33]:

$$SDEC = \sqrt{\frac{\sum_{i}^{N}(Y_c - Y_a)^2}{N}} \qquad (6)$$

where Y_c is the calculated activity for the compounds used as data points in the regression model. Like r^2, SDEC is a measure of the goodness of fit of the regression model, and does not evaluate its predictive ability.

We use the mean absolute error of prediction (MAEP) to estimate the binding affinity prediction error [35,36]:

$$MAEP = \frac{\sum_{i}^{N}|Y_p - Y_a|}{N} \qquad (7)$$

where $|Y_p - Y_a|$ is the absolute value of the difference between actual and predicted activities. SDEP (Eq. 4) gives higher weights to large variations, whereas in MAEP (Eq. 7) all variations are weighted equally.

Chemometric methods

To analyze the CoMFA fields with multiple cross-validation and variable selection procedures, we used GOLPE [37] (Generating Optimal Linear PLS Estimations). This program performs chemometric analyses on GRID [38] and CoMFA fields, and can be used to further refine the PLS model using variable selection procedures according to D-optimal and fractional factorial design strategies [39]. These procedures select the variables with the highest informative content from the X-block, highlighting CoMFA field regions of interest. We used a D-optimal design preselection of variables in the partial-weights space prior to fractional factorial design, as suggested by Clementi and co-workers [39].

We determined the optimal number of components using the two-random-groups cross-validation model. In this procedure, the data set is randomly divided into two equal groups, one set being used to derive the model, the other one for prediction. The process is repeated 25 times,

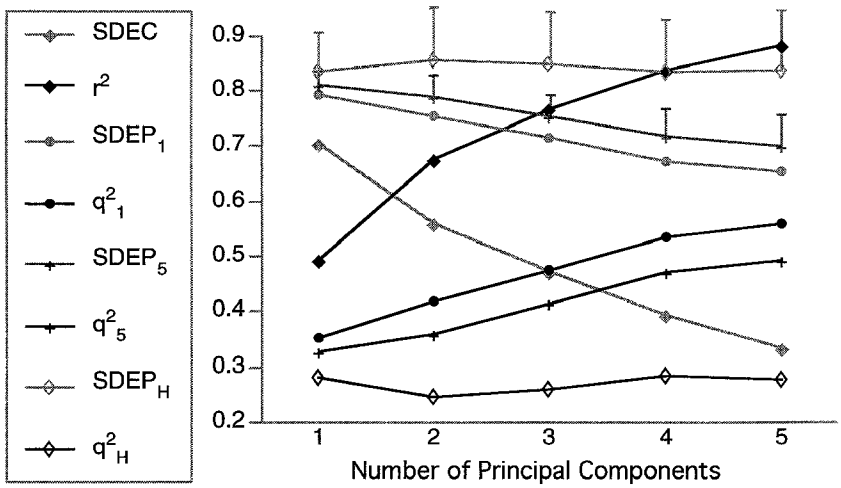

Fig. 2. Correlation results for the 33-compound model. SDEP and q^2 subscribted indices for cross-validation are: 1 = leave-one-out model; 5 = five-groups random model; H = two-groups random model. The SDEP standard deviation among the 25 CV2 and CV5 cross-validation runs is presented as a positive error bar. For the LOO procedure, the SDEP standard deviation is on the order of 10^{-7} (not shown).

TABLE 3
PREDICTIVE ABILITY EVALUATION: COMPARISON BETWEEN INTERNAL AND EXTERNAL PARAMETERS

External prediction parameters						Internal parameters			
PCs	PRESS	q^2_{TS}	q^2_{TR}	MAEP (kcal)	$SDEP_E$	SDEC	$SDEP_1$	$SDEP_5$	$SDEP_H$
Initial model (no variable selection)									
1	5.852	–0.803	0.893	0.493 (0.660)	0.587	0.696	0.792	0.807	0.833
2	4.049	–0.247	0.926	0.418 (0.559)	0.488	0.557	0.758	0.792	0.861
3	9.906	–2.052	0.819	0.644 (0.862)	0.763	0.438	0.710	0.751	0.847
DOD model									
1	5.779	–0.780	0.894	0.489 (0.654)	0.583	0.692	0.786	0.801	0.827
2	4.201	–0.294	0.923	0.420 (0.562)	0.497	0.552	0.752	0.784	0.854
3	10.248	–2.157	0.812	0.654 (0.876)	0.776	0.436	0.707	0.749	0.839
FFD model									
1	8.690	–1.677	0.841	0.614 (0.822)	0.715	0.601	0.699	0.720	0.758
2	5.945	–0.832	0.891	0.489 (0.654)	0.591	0.459	0.585	0.618	0.683
3	4.630	–0.426	0.915	0.428 (0.573)	0.522	0.367	0.589	0.612	0.685

For q^2 calculations, the following parameters were used: for the external set-based index, q^2_{TS}, $Y_m = 1.951$ (with SD = 3.246), and, for the training set-based index, q^2_{TR}, $Y_m = 0.211$ (with SD = 1.003). The SDEP subscribted indices are as follows: E = external set; 1 = leave-one-out; 5 = five-groups random model; H = two-groups random model. Note that internal PRESS values are not given, since $SDEP = (PRESS/N)^{1/2}$.

using recalculation of weights. We compared these results with the five-random-groups cross-validation model. In this procedure, 20% of the compounds are randomly excluded from the data set and used for prediction. This process is also repeated 25 times, with recalculation of weights. In what follows, the labels CV2 and CV5 are used to identify results from the two- and five-groups cross-validation models, respectively.

A recent 3D QSAR application using GOLPE and GRID [40] has shown that, even if the alignment rule is determined by X-ray crystallography, variable selection procedure criteria have to be established within GOLPE before selecting the final PLS model. In this work, the X-distribution for both steric and electrostatic fields was studied in GOLPE, and a standard deviation (SDEV) filter was applied to each field: 0.1 kcal for steric field values, and 0.8 kcal for electrostatic field values. To gain insight in the individual contribution of each field, different weights were used for the steric/electrostatic weighting ratio, going from 0.0 to 1 in both fields using a 0.5 increment. The field-weighting coefficients study is summarized in Fig. 7 and discussed in the section on the 50-compound model.

CoMFA/PLS results were recalculated using the GOLPE implementation of the PLS algorithm. These latter results are reported for all models. Because the number of PLS models obtained precludes their listing in tabular form, a graphical representation was used in this study. Important numerical results are given in the text.

The XED alignment procedure

One of the central issues in a 3D QSAR study is the choice of the geometry and superimposition rules within the given series (i.e., alignment rules). Pharmacophoric search methods, e.g. DISCO [41], generate alignments by finding various pharmacophoric map superimpositions between classes of conformers – alignments that can be directly used in CoMFA models. Other methods, e.g. COMPASS [42], use neural network algorithms that select the conformational and superimposition rules by maximizing the predicted activity for all given compounds. When crystallographic data are available, the alignment rule is experimentally determined [35,40]. Alignment rules can be influenced by physicochemical hypotheses (e.g., energy minimizations for neutral versus ionized compounds) which affect the quality of the resulting models [35,36].

To obtain an alignment rule and to understand steroid recognition features for aromatase we have used the XED (eXtended Electron Definition) program [43]. XED explicitly treats the π-electrons and lone pairs of polar groups in a molecular mechanics framework. The program locates molecular property extrema (MPE) for the interaction with three probes (H^+, OH^- and H_2O) after adding XED points to a given molecule. The approach is conceptually similar to the GRID [38] method. These MPE can be used to superimpose two different ligands, regardless of their exact atomic structure [44]. The MPE pattern for one ligand (e.g., the most active or rigid compound) serves as a template for systematic rigid docking of the other ligand's MPE structure, which is minimized with a Simplex algorithm onto the inverse-polarity MPE template of the first ligand. The resulting superimposition is conformationally dependent, but reflects the pharmacophoric pattern of the two ligands without an exact geometric superimposition. This method successfully reproduced [44] the binding site orientation for three different HIV-1 protease inhibitors and for three steroids binding to the same monoclonal antibody, as observed with X-ray crystallography [45,46].

TABLE 4
DIFFERENCES BETWEEN ACTUAL AND PREDICTED pK_i VALUES FOR THE TEST SET COMPOUNDS

No.	Actual	Initial	DOD	FFD
46	1.201	−0.305	−0.265	−0.286
49	1.208	−0.413	−0.372	−0.457
44	1.432	0.260	0.241	−0.285
43	1.509	−0.380	−0.382	−0.643
42	1.658	0.130	0.144	0.182
45	1.678	−0.379	−0.383	−0.849
41	1.921	0.035	0.013	−0.182
34	1.959	0.454	0.481	0.455
47	2.0	0.181	0.156	0.078
40	2.056	0.320	0.365	0.379
39	2.174	0.303	0.280	0.063
35	2.252	0.464	0.463	0.358
48	2.292	0.877	0.881	0.794
50	2.310	0.409	0.449	0.174
37	2.328	0.480	0.456	0.242
38	2.337	0.602	0.646	0.674
36	2.854	1.111	1.156	1.181

The order of the compounds coincides with the plot in Fig. 3. Mean absolute errors of prediction (MAEP) are given in Table 3.

In the present study, 33 steroids and the natural substrate androstenedione were superimposed using the XED procedure, with compound **21** as template. The resulting XED superimposition pattern comprises the same orientation of the A-, B-, C- and D-rings throughout the series. The distance between superimposed compounds varied between 0.07 and 0.20 Å at the fused ring system. The XED pattern suggests that these steroid inhibitors and the substrate have a common binding site.

This result is supported by the following observations: (i) the C19 methyl of the substrate has to be within 5 Å of the heme iron atom [20]; steroidal inhibitors that compete with the substrate occupy that region in space; (ii) seven (**1**, **3**, **4**, **6**, **7**, **14**, **15**) of the 33 compounds, which inactivate aromatase in a time-dependent manner, were shown to initially form a dissociable enzyme–inhibitor complex, followed by unimolecular inactivation [28]; these results suggest that the seven suicidal inhibitors occupy, at least in part, the substrate binding site. On the basis of these observations, we have included compounds **1**, **3**, **4**, **6**, **7**, **14** and **15** in this CoMFA study. In the Results section concerning the 50-compound model, we show that removing these compounds from the training set does not exert a significant influence on the CoMFA/GOLPE results.

We used the XED superimposition pattern to generate the CoMFA alignment. Because PLS attributes changes in biological activity to regions of high variance in the CoMFA fields, we used the rms fit procedure to reduce the influence of regions with high similarity between compounds, e.g., the common part of the steroidal framework.

Two rms fit alignments were tested for compounds **1**–**33**: superimposition of all 17 carbons in the steroid skeleton, and superimposition of the B-, C- and D-rings, to enhance the contribution of the A-ring field variance to the biological activity (13 carbons). The superimposition of all four rings (17 atoms) yields higher regression results ($r^2 = 0.886$ and $q^2 = 0.562$ for the five-component model) when compared to the rms fit of the B-, C- and D-rings ($r^2 = 0.824$ and $q^2 = 0.494$ for the five-component model, respectively). We used the rms fit of all 17 carbons

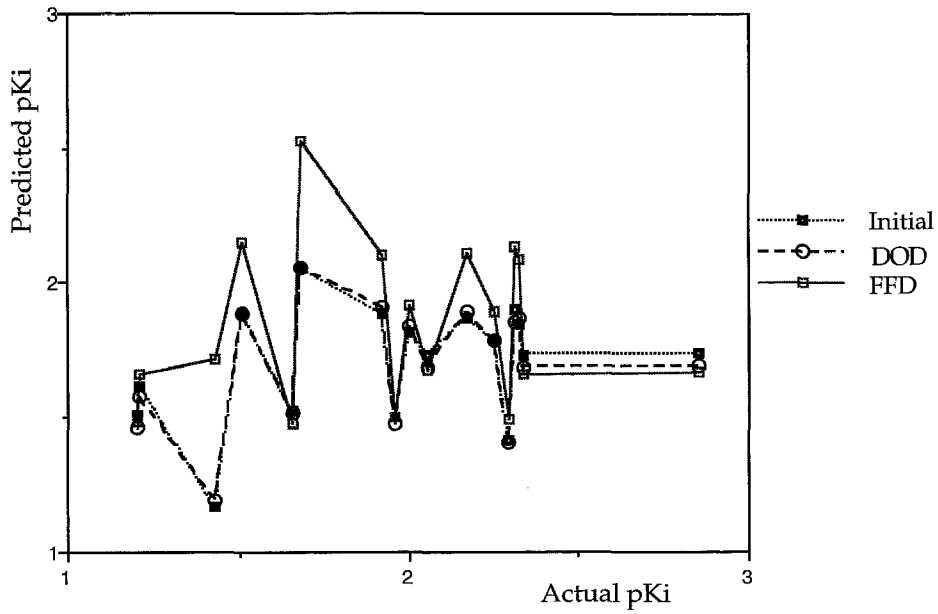

Fig. 3. Predictive power comparison for variable selection procedures: actual versus predicted pK_i (μM) activities for the 17 compounds in the external set, based on the 33-compound models with the lowest PRESS and SDEP values. The two-component model was used for both the initial and the D-optimal design (DOD) preselected models, while the three-component model was used for the fractional factorial design model (FFD).

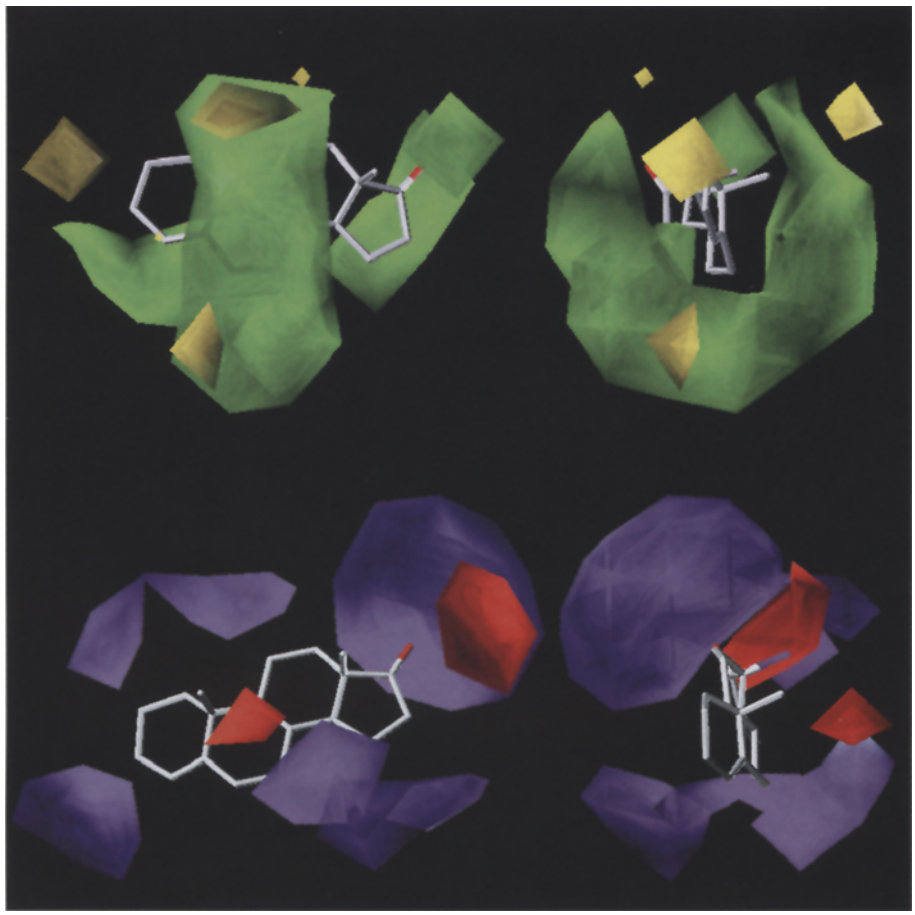

Fig. 4. Orthogonal views of the major steric (a) and electrostatic (b) features of the CoMFA for aromatase inhibition for the 33-compound model. Field contours surround regions of high (green) and low (yellow) steric bulk tolerance. Blue and red contours surround regions where negative polar interactions are favored (blue) or detrimental (red).

in the steroid framework as alignment rule throughout this study.

Molecular modeling

CoMFA model systems are based on defined molecular geometries and on atom-centered partial charges. Molecular geometries were determined starting from crystallized steroid structures. The following crystallographic structures from the Cambridge Crystallographic Data Center were used for the starting geometry of modeled compounds: NOTEST01 (17β-hydroxy-19-nor-androst-4-en-3-one) for compound **29**, HTESBZ (19-hydroxy-17β-benzoyloxy-androst-4-en-3-one) for **32** and other 19-hydroxyl and 19-keto compounds, and HANDDO (17β-hydroxy-androsta-4,14-dien-3-one) for all other compounds.

Each steroid structure was optimized with the MOPAC [47] 6.0 program (MNDO Hamiltonian, using the Eigenvector Following minimizer). Alkyl and alkyl–aryl chains were modelled in the all-trans (extended) conformation. The same extended conformation was considered by Numazawa and Oshibe in their MOPAC (PM3) analysis of these 6-alkyl- and 6-aryl-substituted steroids [27]. Partial atomic charges from the semiempirical MNDO method were used throughout this study, except for the XED superimposition, when empirical XED point charges were added onto the MNDO-derived ones, as described by Vinter [43].

Results and Discussion

The 33-compound model: Robustness and predictive power analysis

The initial CoMFA model was based on 33 compounds (**1**–**33**). For this set, relative and absolute potencies were available. This model shows a good degree of robustness, judging from the LOO and CV5 cross-validation models (see Fig. 2 for regression model results, and Fig. 4 for CoMFA fields). The optimal model, with the lowest SDEP and the highest q^2, has five components: $r^2 = 0.886$ (5); $q_l^2 = 0.562$ (5); $q_5^2 = 0.494$ (5); and $q_H^2 = 0.276$ (5). The subscripts used are explained in Fig. 2. However, according to CV2 cross-validation, the optimal model has only one component, $q_H^2 = 0.283$ (1), because q^2 drops at the two-component model, $q_H^2 = 0.247$ (2). The minor improvement at four components, $q_H^2 = 0.284$ (4), is due to noise accumulation. Test set predictions (compounds

34–50) confirm the apparent discrepancy between *robustness* (predictive power evaluated by internal cross-validation) and *external predictivity* (as tested in CV2, which simulates external prediction by drastically reducing the internal data set): the two-component model has the lowest SDEP and highest q^2 (see Table 3).

To improve the signal-to-noise ratio, we used a 0.1 cutoff for the steric field and a 0.8 cutoff for the electrostatic field, with a weighting ratio of 1.359 (steric) to 0.828 (electrostatic) for this PLS model. Variable preselection using D-optimal design (partial weights, 50% reduction level) reduces the number of X variables from 2460 to 577, without improving model robustness and/or predictivity (see Table 4). Variable selection using fractional factorial design (cross-validation at two components, using the model with five random groups, 25 randomizations) further reduces the number of variables to 130, yielding a more predictive model, as observed by comparing SDEC and SDEP results in the LOO, CV5 and CV2 procedures (Table 3).

We have tested external predictivity using compounds **34–50**. The initial PLS model (two components) has slightly higher predictivity, as observed by comparing the predictive power evaluation indices: 4.049 (PRESS), 0.488 (SDEPE), and 0.418 (MAEP). Errors of prediction of individual compounds show minor variation among the three GOLPE models, except for compounds **44** and **45** in the fractional factorial model (see Table 4). With few exceptions (**36**, **38** and **48**), all compounds in the external set are predicted within less than 0.6 log units (0.76 kcal/mol) of the actual pK_i. These 6β-alkyl-substituted compounds (**36**, **38** and **48**) are some of the most active in the test set. Seven out of nine 6β-substituted steroids are underpredicted (see Table 4 for errors of prediction). Except for one 6-keto- (**24**) and three 6α,7α-substituted (**25–27**) steroids, no C6-substituents are present in the training set. Therefore, predictions for the C6 region are based on extrapolation. This, and the use of absolute pK_i values, diminishes model predictivity for this test set.

During predictive power analysis, we have observed a large difference between two predictive r^2-based indices (Table 3): q^2_{TS}, obtained using the test set Y_m, and q^2_{TR}, obtained using the training set Y_m [32]. Results using the training-set-based predictive r^2, q^2_{TR}, proved to be misleading, as their value (higher than 0.8 in all cases) suggests a very good predictive power. The test-set-based predictive r^2 (q^2_{TS}) was found to be a better estimate of the predictive power, as there is no correlation between predictive and actual activities in this case ($r^2 = 0.097$ for the actual versus predicted values in the initial PLS model; see Fig. 3). The difference between the Y_m values explains the discrepancy between the two predictive r^2 indices: the test set K_i has a different distribution compared to the training set K_i. This shows the inability of q^2_{TR} to accurately evaluate predictive power.

Although it is often used in the CoMFA literature, we recommend that the training-set-based predictive r^2 (q^2_{TR}) should be avoided, or used only in conjunction with Y_m-independent indices such as PRESS, SDEP and MAEP. We suggest the use of q^2_{TS} instead of q^2_{TR} to compare the reliability of prediction between different QSARs. The reliability of prediction cannot be compared using PRESS, SDEP and MAEP, because these are context-dependent parameters.

For this aromatase inhibitors series, external SDEP values are consistently smaller than cross-validation-based (internal) SDEPs for all optimal models, except for fractional factorial design. This apparently suggests a good predictive power, which is not confirmed in the actual versus predicted plot (Fig. 3). This is misleading, because SDEP values are context-dependent: internal SDEPs refer to a different data set than external SDEPs. Results from the CV2 cross-validation suggest a poor predictive ability in all five regression models, e.g., $q^2_H = 0.247$ (2), $SDEP_H = 0.857$ (2). Therefore, among cross-validation procedures, CV2 is the better estimate of external predictivity. While cross-validation proves useful in selecting among many potential predictors [48], SDEP (PRESS, q^2) parameters remain an estimate of the predictive power, inherently limited by data set composition. The distinction between robustness and predictive power analysis further supports the idea that QSAR models should be analyzed both in terms of robustness and external predictivity.

The predictive power analyses show that accurate prediction of activity is not correlated to the local CoMFA fields in a direct manner, hence it should be evaluated separately. Alterations in training set composition to include some of the 6-alkyl or 6-aryl steroids in the system improve external predictivity, as shown by the CV2 cross-validation results (discussed in the 50-compound model section).

The explanatory power of the 33-compound model is not discussed, because similar results are detailed for the 50-compound model in the next section. For comparative examination, we present the CoMFA steric and electrostatic field * coefficients contours in Figs. 4a and 4b. The contours represent the scalar products of the average steric and electrostatic fields for active compounds (50 nM > K_i > 5 nM; i.e., compounds **18**, **21**, **22**, **24**, **25**, **26**) with the PLS (β) coefficients. The conventional graphic representations for CoMFA [34] were used. We have discussed elsewhere [49] the choice of this field over the traditional stdev*β field.

In summary, the 33-compound model exhibits good explanatory power compared to chemical structures from the test set and to other external compounds. Its lack of predictive power for the test set is due to extrapolation procedures, although the error in predicting the binding affinity is less than 0.6 kcal/mol (Table 3). Variable selection methods (D-optimal design, fractional factorial de-

sign) do not improve external predictivity. These methods emphasize beneficial and detrimental CoMFA fields and can be used for regional chemical optimization.

The 50-compound model

We obtained a new CoMFA model by enlarging the data set – compounds **1–50** (see Fig. 5 for statistic results and Fig. 6 for CoMFA fields). Field features of this model do not differ significantly from the previous one. We applied the same SDEV cutoff values for the steric and electrostatic CoMFA fields (0.1 and 0.8, respectively), and the same weighting coefficients (1.359 and 0.828, respectively). We obtained the following results (using the above settings): $r^2 = 0.674$ (1), $q_1^2 = 0.619$ (1), $q_5^2 = 0.613$ (1), $q_H^2 = 0.618$ (1). Although we obtained higher values for higher-component models, all q^2 values dropped at two components, suggesting that further q^2 improvements reflect noise addition to the X-block. Changes in the weight coefficients for the two fields explain this situation, as summarized in Fig. 7 and discussed below.

In this study, the filtering SDEV values were maintained constant (0.1 and 0.8, respectively). Using the steric field only, the LOO q^2 peaks at three components, while in the model based on the electrostatic field it peaks at one component. We compared three intermediate steric to electrostatic field weights: 2:1, 1:2 and 1.3:0.8 (optimal as suggested in GOLPE). In both r^2 and q^2 (LOO), the 2:1 weighting scheme performs better than the optimal value, and they both outperform the 1:2 weighting model. For all models, except for the steric field only, q^2 values dropped at two components. Higher components add noise to the electrostatic field model. This observation is consistent with all cross-validation results presented in Fig. 5. A one-component model is better for the electrostatic field, but does not optimally model the steric interaction. In a three-component model (better for the steric field), noise is added to the electrostatic field. To select the final model, we applied the parsimony principle (*less is better* [48]), hence we preferred the one-component model. This parallels results from the 33-compound model, where CV2 statistics also suggested the one-component model (Fig. 2). Further analyses used the 1.3:0.8 weighting scheme, for direct comparison with the 33-compound model.

The CoMFA fields resulting from the 50-compound model are not essentially different from those obtained in the 33-compound model (see Figs. 4 and 6 for comparison). We used the average steric and electrostatic field for the active compounds (13 nM > K_i > 1.4 nM, i.e., compounds **21**, **22**, **26**, **34–41**, **47**, **48**, **50**) for the field * coefficients graphical analysis. We note areas of high steric-bulk tolerance at the B-ring, at positions 6 and 7 (both on the α- and β-face), These areas extend towards the A-ring (positions 3 and 4, equatorial), cover the A-ring on the α- and β-face (including C19), and less of the B-ring (β-face) and the D-ring (position 16α). Areas of restricted steric-bulk tolerance occupy the same areas as in the 33-compounds model: mostly at the A-ring, between positions 2 and 3 (α-face), at C19 (outside the steroid rings), and at positions 6α and 12α. The high-tolerance area at position 6α is oriented towards positions 4, 7 and the C-ring. We examine the possibility of a common binding pocket for 4-, 6α- and 7α-substituents in the next section.

Polar (negative) groups are favored at the A-ring (positions 2–4, mostly on the α-face and equatorial), between the B- and D-rings (positions 7 and 15–17); these are different from the 33-compound model (e.g., no beneficial electrostatic fields present at positions 1 and 2, on the α-

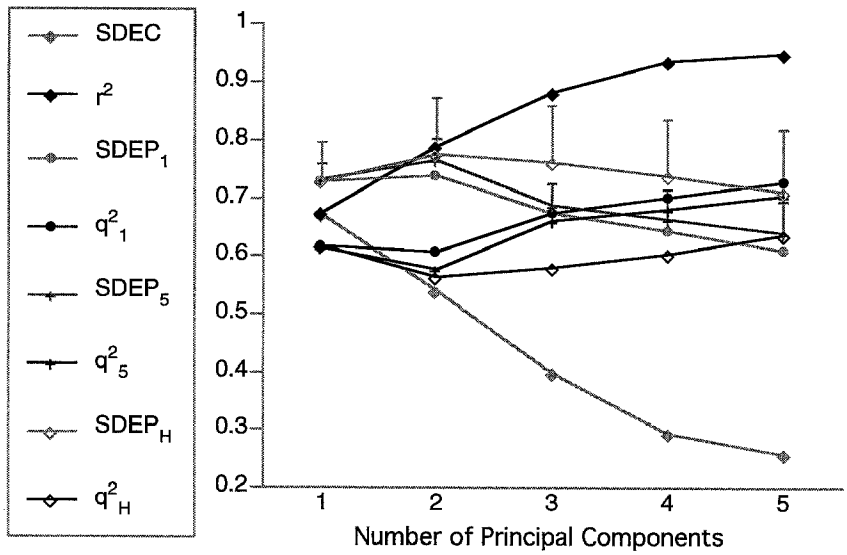

Fig. 5. Correlation results for the 50-compound data set. SDEP and q^2 are subscribted as in Fig. 2. The SDEP standard deviation among the 25 CV2 and CV5 cross-validation runs is presented as a positive error bar. For the LOO procedure, the SDEP standard deviation is on the order of 10^{-7} (not shown).

and β-face). Polar groups are detrimental both on the α- and β-face, above the B-ring (positions 6 and 10; both faces), the D-ring (position 17β), and the A-ring (position 2α). The beneficial electrostatic fields are consistent with the presence of the keto groups at C3 and C17 in the natural substrate and some of the active inhibitors (e.g., compounds **24**, **25**, **34**–**41**, **47**, **48** and **50**). Compounds with a keto group at C7 show weak activity (e.g., compounds **1**–**10**). The detrimental electrostatic fields at C17β are consistent with the loss of activity observed when reducing the 17-keto to a 17β-hydroxyl (e.g., compounds **26** and **27**).

Compounds with 7α-substituents have been synthesized [51–53], some of which are potent aromatase inhibitors [54]. These compounds have steric bulk in a beneficial region observed in our CoMFA fields (Figs. 4a and 6a). Aryl groups are, to some extent, polar functions, e.g., 7α-thioaryl-ethers, provide this steric bulk. Other potent aromatase inhibitors bridge C2 and C10 on the β-face [55–57]. These compounds have a propyl, ether or thioether bridge between C2 and C10, and may be irreversible inhibitors.

A dimethyl-ether bridging C10 and C6, based on the androstenedione structure, has been synthesized [57]. Our CoMFA model predicted this compound to be highly active (estimated K_i of 2 nM). In vitro, it produces 0% inhibition of aromatase at 1 μM, after 20 minutes preincubation in the human placental aromatase assay (N.P. Peet, 1995, personal communication). According to electrostatic field analysis, the ether-bridge oxygen of the above compound is situated in a detrimental area and should be inactive (red above the B-ring in Fig. 6b). However, because of the regions of steric-bulk tolerance above the B-ring (Fig. 6a), this compound is predicted to be active. This result documents the lack of predictivity above the C6–C10 region.

Synthesis of 4-amino-androstenedione and some derivatives [58,59] confirms, indirectly, the observed beneficial steric and electrostatic fields at the A-ring (positions 3 and 4). Some compounds in this series are potent irreversible inhibitors of aromatase. One of them, FCE 24304, with a methylene substituted at C6, is currently in

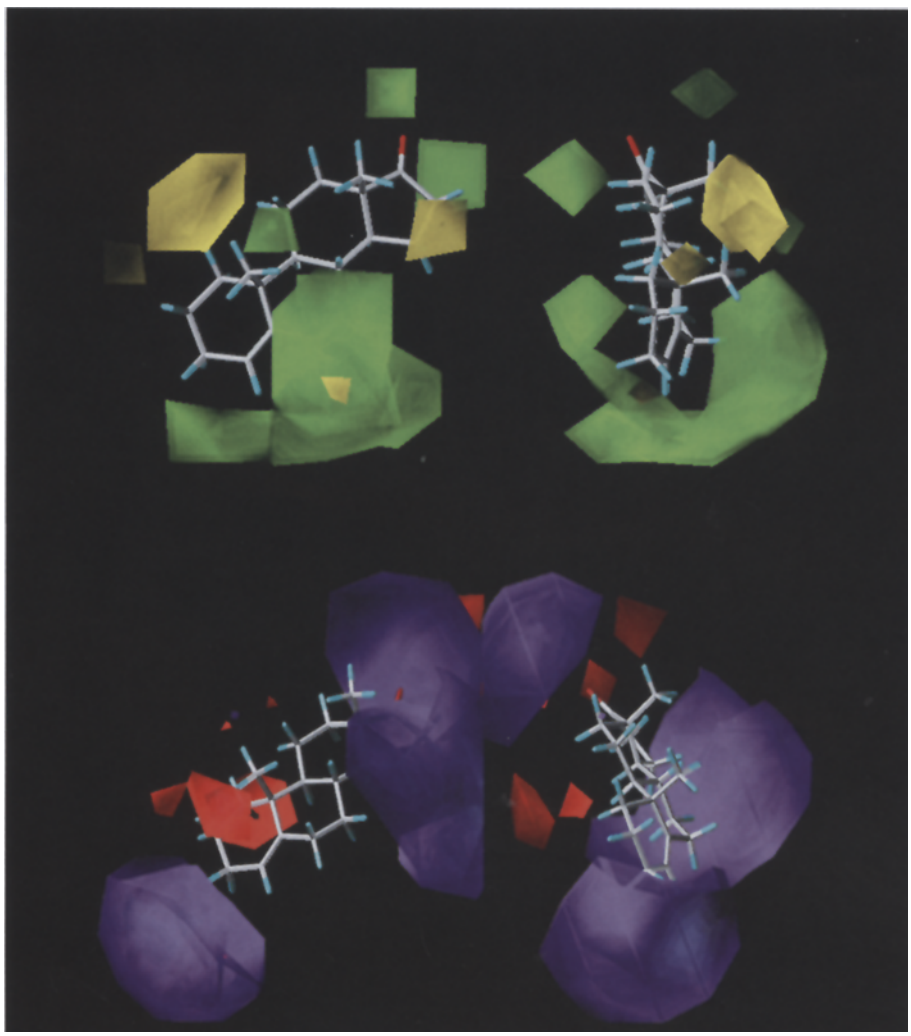

Fig. 6. Orthogonal views of the major steric (a) and electrostatic (b) features of the CoMFA for aromatase inhibition for the 50-compound model. CoMFA field conventions are explained in the legend to Fig. 4.

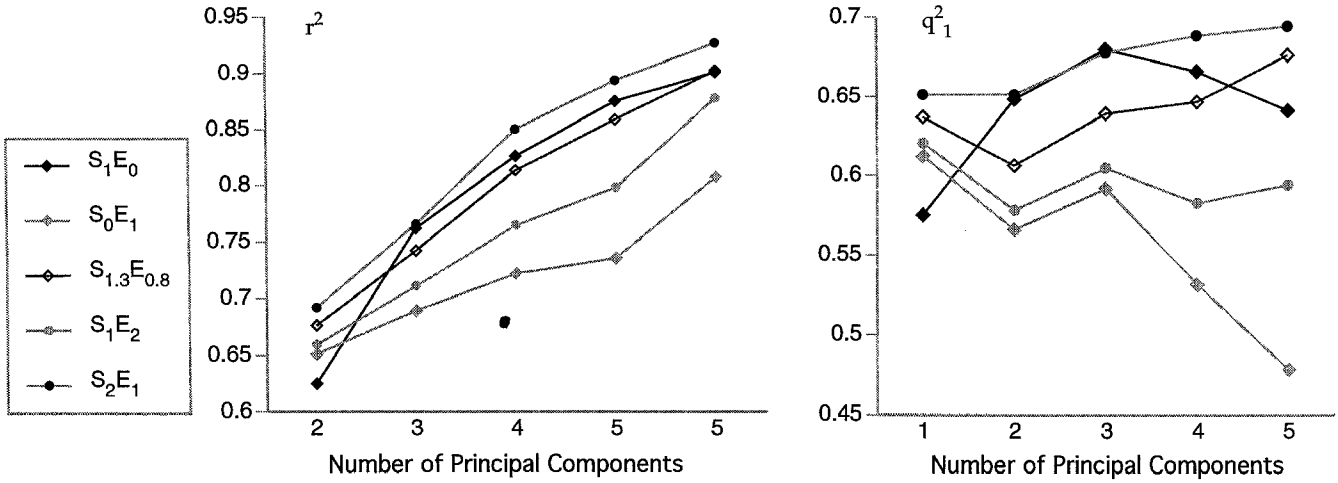

Fig. 7. Dependence of correlation results for the 50-compound model on the weighting coefficients applied to the steric and electrostatic fields. Only r^2 and q^2 (leave-one-out) statistics are presented. See text for details.

clinical trials [7]. Steric bulk substituents at C4 (e.g., 4-thiophenyl-androstenedione) can also yield competitive, potent aromatase inhibitors [60]. Some of the conformations adopted by these compounds (the phenyl group pointing towards the C6-C7 region) are consistent with our steric CoMFA fields. The presence of keto groups at C3 and C17 and of steric bulk at position 6 in potent training set inhibitors (e.g., compounds **34–41**, **47**, **48**, **50**) is consistent with these fields. We have discussed the detrimental electrostatic fields at the B-ring (consistent with the C6-C10 dimethyl-ether bridged steroid) and C17β in the previous model.

Several suicidal inhibitors (**1**, **3**, **4**, **6**, **7**, **14**, **15**) were included in this model, as previously discussed. The apparent K_I values, obtained by Kitz–Wilson plots in the kinetic assay, are numerically similar to the K_i values obtained by Dixon plots under initial velocity conditions: the compounds have the same ranking order in terms of potency, while the largest $(pK_i - pK_I)$ difference does not exceed 0.5 log units (see Table 2). The effect of removing these compounds from the training set over the CoMFA model was studied using GOLPE, under identical conditions as for the 50-compound model. The results from the one-component model ($r^2 = 0.663$, SDEC = 0.590, $q^2_1 = 0.607$, $SDEP_I = 0.637$, and $q^2_H = 0.571$, $SDEP_H = 0.668$) suggest that, in terms of robustness and predictivity, the qualities of the 50-compound model are preserved: the 43-compound model is robust, predictive, and simple. While examining the beneficial versus detrimental CoMFA fields, the only notable difference was a reduced importance of the detrimental electrostatic region above the A-ring (results not shown). We chose to maintain these compounds in our model because they contribute to the chemical diversity of the training set.

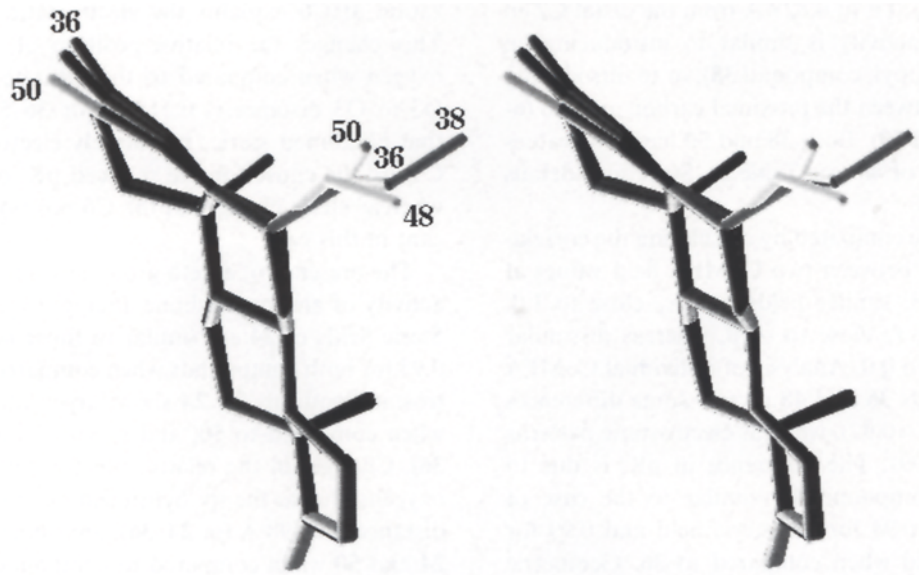

Fig. 8. Stereoview of structural models for active 6β-alkyl-substituted steroids. Hydrogen atoms were omitted for clarity. See text for details.

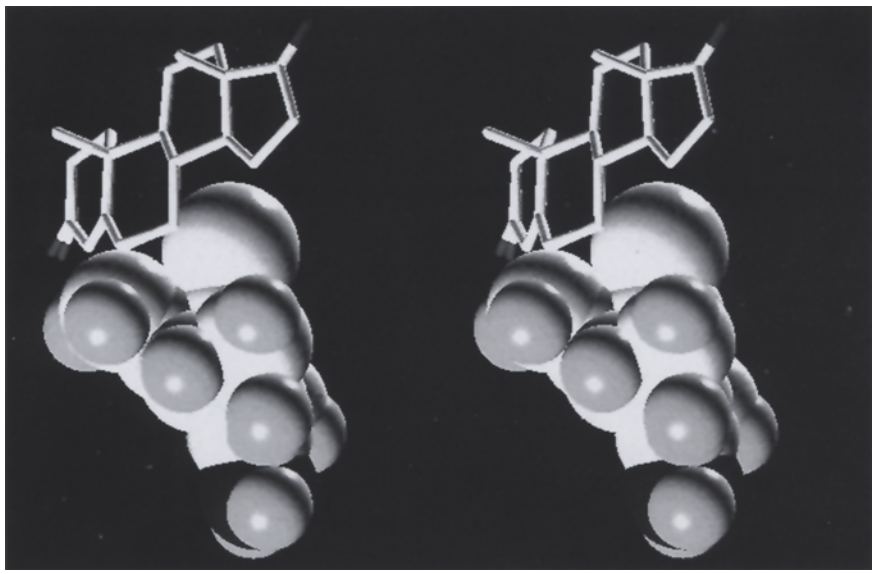

Fig. 9. Stereoview of a spacefilling model for 6α-benzyl (**47**) and 7α-thio-(*p*-amino)phenyl substituents of androstenedione. In our alignment, these moieties occupy the same region in space, suggesting a common hydrophobic binding pocket for 6α- and 7α-substituted aromatase inhibitors. The steroid framework is shown as capped sticks. Hydrogen atoms are shown for the steric bulk substituents only.

The hydrophobic binding pockets

Numazawa and Oshibe studied [27] structure–activity relationships for the 6-alkyl and 6-aryl derivatives. They observed that two saturated carbons (C6β-ethyl) represent the optimal steric bulk tolerance in this series (compounds **34**–**50**) [27]. To understand the differences in aromatase inhibition, we compared structure–activity features for those compounds containing two carbons substituted at C6 in the beta position (**36**, **48**) or in the ring plane (**50**). We used compound **36** (6β-ethyl), with the highest potency, as reference for these comparisons.

In our alignment, after rms fit of all steroid ring carbons, compound **48** (6β-vinyl) has the C1' (proximal carbon at C6) at 0.13 Å from the C1' carbon in **36**, whereas the distal carbon (C2') is at 0.829 Å from the distal C2' in **36**. The effect on activity is similar to introducing an extra carbon (6β-propyl, compound **38**), or to introducing the double bond between the proximal carbon and C6 (6-ethylene, compound **50**). Both **38** and **50** have pK_i values comparable to that of **48** (see Table 2). Structural details are shown in Fig. 8.

CoMFA fields are compared by calculating the correlation coefficient (r_F) between two CoMFA field values at the same grid point; similar fields have r_F close to 1.0, opposite fields have r_F close to −1.0, whereas dissimilar fields have r_F close to 0.0. Analysis of individual CoMFA fields for compounds **36** and **48** reveals some differences in the steric field ($r_F = 0.92$), whereas electrostatic patterns are similar ($r_F = 0.99$). The difference in pK_i is due to steric effects for compound **48**, similar to the case of compound **38**: $r_F = 0.94$ for the steric field and 0.99 for the electrostatic field when compared to **36**. Geometric arguments further support the possibility of a steric, not electrostatic effect: the C6-C1'-C2' angle is 109.5° in one case (6β-ethyl) and 120° in the other case (6β-vinyl), while the C1'-C2' distance in the latter compound is approximately 0.3 Å shorter.

The role of the steric effect in this region is further supported by structure–activity relationships for compound **50** (6-ethylene). In our alignment, compound **50** has the C1' at 0.59 Å from the C1' carbon in **36**, and the distal carbon (C2') at 0.363 Å from the corresponding C2' in **36**. By placing C2' in close proximity for both compounds, the variation in the steric field is considerably reduced; $r_F = 0.97$ for the steric field and 0.91 for the electrostatic field, respectively, when compared to **36**. The distortion of the B-ring introduced by the sp^2 hybridization at C6 explains the electrostatic field differences. This changes the relative position of the C3 carbonyl oxygen when compared to the saturated congeners (the O3 to O3 distance is 0.518 Å for **36**–**50**). We conclude that combined steric, but largely electrostatic effects at C3 are the cause for the reduced pK_i of compound **50**, whereas steric changes at the C6 position are not significant in this case.

The presence of a keto group at C6 further reduces the activity of androstenedione analogs (e.g., compound **24**). Steric fields of **24** are similar to those of **36** and **50** ($r_F = 0.95$ for both compounds when compared to **24**), but electrostatic patterns for **24** show larger variations ($r_F = 0.81$ when compared to **50**, and $r_F = 0.77$ when compared to **36**). Changes in the relative position of the C3 carbonyl oxygens, due to the sp^2 hybridization at C6 (the O3 to O3 distance is 0.498 Å for **24**–**36**), are similar for compounds **24** and **50** when compared to compound **36**. Changes at the C6 position reduce the similarity in the electrostatic

field more in both cases (comparing **24** to **36** and to **50**). The effect of the C6 carbonyl oxygen is twofold: the absence of the beneficial distal carbon at C6β (present in **36** and **50**), and the presence of a negative charge in the same region. This combined effect has a detrimental contribution to the compound's activity: **24** has a smaller pK_i compared to **50**. Therefore, the lack of steric bulk and the presence of negatively charged atoms in the C6 region are detrimental to aromatase inhibition potency.

The reduced tolerance for steric bulk in the C6β binding pocket can be deduced by comparing pK_i values for compounds **34**, **36**, **38**, **40**, **42**, **46**, **48** and **50**, whereas comparison of compounds **24**, **36** and **50** supports the detrimental effect of electrostatic interactions. These observations are not in agreement with the PM3-optimized model system presented by Numazawa and Oshibe [27]; they report: 'The 6β-vinyl moiety superimposes very well with the 6β-ethyl function in our modeling. This strongly suggests that an electrostatic effect rather than steric reasons would principally be involved in the change of the affinity.' The above geometric and structure–activity considerations, however, support the existence of a small binding pocket, favoring steric, and not electrostatic, interactions, in the 6β-region of the steroid framework.

On the basis of the studied compounds, and by comparing structural flexibility with that of 7- and 7α-substituted steroids [53–55], we hypothesize that steric bulk substituents at the 7α- and 6α-positions occupy parts of the same hydrophobic binding pocket. For example, 6α-benzyl (**47**) and 7α-thio-(*p*-amino)phenyl substituents of androstenedione can occupy the same region of space with three of the phenyl ring atoms (see Fig. 9). The predicted pK_i of the 7α-thio-(*p*-amino)phenyl, 1.985, is close to the actual value [56], 2.004, which is identical to that of compound **47** (2.000). Since other structural features are identical in the two compounds, it can be hypothesized that steric bulk substituents in the C6 and C7 alpha region have the same effect on aromatase inhibition.

A similar effect is observed by introducing the thiophenyl group at the C4 position in androstenedione [52]. This compound can access the same hydrophobic pocket as C6 α-substituted steroids. In our 50-compound model, the C7 *p*-amino-phenyl moiety and the C4 thiophenyl occupy beneficial CoMFA steric fields. Furthermore, the 6α,7α-cyclopropane derivative **26** ($pK_i = 2.301$) of compound **21** ($pK_i = 1.886$) shows improved potency over its congener, likely because it introduces steric bulk in the 6α,7α-hydrophobic binding pocket. This evidence supports our hypothesis.

Our model is in agreement with the 6β-region hydrophobic binding pocket proposed by Numazawa and Oshibe [27]. Both models found no overlap between steric bulk substituents at the 6α- and 6β- positions. The limited bulk tolerance for 6β-substituted steroids compared with the high degree of bulk tolerance of 6α- and 7α-substituted compounds suggests the existence of two different binding pockets for aromatase: a smaller one located near the 6β-region, and a larger one at the C6-C7 α-face, accessible to C4-thioaryl substituents.

In summary, the 50-compound model reveals structure–activity relationships that explain differences in the binding affinity of studied compounds. Chemical structures not included in this study overlap with CoMFA field features of the 50-compound model. Our model suggests two different hydrophobic binding pockets, one located at the α-face of the C6-C7 (and possibly, C4) region, and a smaller one at the 6β-region.

Conclusions

In the analysis of steroid aromatase inhibition, we have derived a robust and simple one-component CoMFA model. The unique features of this model were confirmed by comparison of CoMFA fields with potent inhibitors, not included in the training set (explanatory power analysis). Variable selection methods (D-optimal design, fractional factorial design) have no effect on the external predictivity of the 33-compound model. These methods are useful in selecting regions for chemical optimization and explaining biological activity. The training set Y_m-based predictive r^2 parameter can be misleading for external predictivity evaluation, when the test set has a different activity distribution, and should be combined with the test set Y_m-based predictive r^2. In the absence of an external set, the two-random-groups cross-validation procedure is a good estimate of model predictivity.

Using our CoMFA model we rationalized, at the atomic level, the observed differences in activity among modeled compounds. These observations permitted us to suggest two hydrophobic binding pockets in the C6 region. A unique, large binding pocket is located at C4 and the α-face of the C6-C7 steroid framework, favoring mostly hydrophobic interactions. A smaller hydrophobic binding pocket is located in the 6β-region. CoMFA field specificity, as demonstrated by field analysis and structure–activity relationships, makes this model useful for designing new aromatase inhibitors.

Acknowledgements

We thank Drs. Denise Beusen and Douglas F. Covey (both from Washington University, St. Louis, MO) and Dr. Norton P. Peet (Marion Merrell Dow Research Institute, Cincinnati, OH) for helpful discussions on aromatase inhibition. Prof. Sergio Clementi, Drs. Gabriele Cruciani and Daniela Riganelli (all from the University of Perugia, Perugia, Italy) are acknowledged for discussions on the predictive r^2 index and for the GOLPE program. We thank Tripos Inc. (St. Louis, MO) for SYBYL. This work was supported by the U.S. Department of Energy.

References

1 Ganong, W.F., Review of Medical Physiology, 16th ed., Appleton and Lange, Norwalk, CT, 1993, p. 780.
2 Chen, S., Besman, M.J., Shively, J.E., Yanagibashi, K. and Hall, P.F., Drug Metab. Rev., 20 (1989) 511.
3 Kellis, J.J. and Vickery, L.E., J. Biol. Chem., 262 (1987) 4413.
4 Thompson, E.A.J. and Siiteri, P.K., J. Biol. Chem., 249 (1974) 5373.
5 Brodie, A.M.H., Brodie, H.B., Callard, G., Robinson, C., Roselli, C. and Santen, R., J. Steroid Biochem. Mol. Biol., 44 (1993) 321.
6 Brodie, A.M.H. and Santen, R.J., Breast Cancer Res. Treat., 30 (1994) 1.
7 Brodie, A.M.H., J. Steroid Biochem. Mol. Biol., 49 (1994) 281.
8 Hervey, H.A., Lipton, A. and Santen, R.J., Cancer Res. (Suppl.), 42 (1982) 3261s.
9 Reed, M.J., Breast Cancer Res. Treat., 30 (1994) 7.
10 Strobl, J.S., In Craig, C.R. and Stitzel, R.E. (Eds.) Modern Pharmacology, 4th ed., Little, Brown and Co., Boston, MA, 1994, pp. 747–759.
11 Hartmann, R.W., Bayer, H. and Grun, G., J. Med. Chem., 37 (1994) 1275.
12 Chen, S. and Zhou, D., J. Biol. Chem., 267 (1992) 22587.
13 Chen, S., Zhou, D., Swiderek, K., Kadohama, N., Osawa, Y. and Hall, P.F., J. Steroid Biochem. Mol. Biol., 44 (1993) 347.
14 Amarneh, B., Corbin, C.J., Peterson, J.A., Simpson, E.R. and Graham-Lorence, S., Mol. Endocrinol., 7 (1993) 1617.
15 Kadohama, N., Yarborough, C., Zhou, D., Chen, S. and Osawa, Y., J. Steroid Biochem. Mol. Biol., 43 (1992) 693.
16 Zhou, D., Pompon, D. and Chen, S., Proc. Natl. Acad. Sci. USA, 88 (1991) 410.
17 Zhou, D., Korzekwa, K.R., Poulos, T. and Chen, S., J. Biol. Chem., 267 (1992) 762.
18 Zhou, D., Cam, L.L., Laughton, C.A., Korzekwa, K.R. and Chen, S.U., J. Biol. Chem., 269 (1994) 19501.
19 Laughton, C.A., Zvelebil, M.J.J.M. and Neidle, S., J. Steroid Biochem. Mol. Biol., 44 (1993) 399.
20 Poulos, T.L., Finzel, B.C. and Howard, A.J., J. Mol. Biol., 195 (1987) 687.
21 Oprea, T.I., Ho, C.M.W. and Marshall, G.R., In Reynolds, C.H., Holloway, M.K. and Cox, H.K. (Eds.) Computer-Aided Molecular Design, ACS Symposium Series Vol. 589, Washington, DC, 1995, pp. 64–81.
22 Green, S.M. and Marshall, G.R., Trends Pharmacol. Sci., 16 (1995) 285.
23 Cramer III, R.D., Patterson, D.E. and Bunce, J.D., J. Am. Chem. Soc., 110 (1988) 5959.
24 Numazawa, M., Mutsumi, A., Hoshi, K. and Koike, R., Biochem. Biophys. Res. Commun., 160 (1989) 1009.
25 Numazawa, M., Mutsumi, A., Hoshi, K., Oshibe, M., Ishikawa, E. and Kigawa, H., J. Med. Chem., 34 (1991) 2496.
26 Numazawa, M. and Mutsumi, A., Biochem. Biophys. Res., Commun., 177 (1991) 401.
27 Numazawa, M. and Oshibe, M., J. Med. Chem., 37 (1994) 1312.
28 Numazawa, M., Mutsumi, A., Tachibana, M. and Hoshi, K., J. Med. Chem., 37 (1994) 2198.
29 Baroni, M., Costantino, G., Cruciani, G., Riganelli, D., Valigi, R. and Clementi, S., Quant. Struct.–Act. Relatsh., 12 (1993) 9.
30 Topliss, J.G. and Costello, R.J., J. Med. Chem., 15 (1972) 1066.
31 Cramer III, R.D. and Bunce, J.D., In Hadzi, D. and Jerman-Blazic, B. (Eds.) QSAR in Drug Design and Toxicology, Elsevier, Amsterdam, The Netherlands, 1987, pp. 3–12.
32 Wold, S., Johansson, E. and Cocchi, M., In Kubinyi, H. (Ed.) 3D QSAR in Drug Design: Theory, Methods and Applications, ESCOM, Leiden, The Netherlands, 1993, pp. 523–550.
33 Wold, S., Esbensen, K. and Geladi, P., Chemometrics Intelligent Lab. Syst., 2 (1987) 37.
34 Cramer III, R.D., Bunce, J.D., Patterson, D.E. and Frank, I.E., Quant. Struct.–Act. Relatsh., 7 (1988) 18.
35 Cruciani, G., Baroni, M., Clementi, S., Costantino, G., Riganelli, D. and Skagerberg, B., J. Chemometrics, 6 (1992) 335.
36 Cramer III, R.D., DePriest, S.A., Patterson, D.E. and Hecht, P., In Kubinyi, H. (Ed.) 3D QSAR in Drug Design: Theory, Methods and Applications, ESCOM, Leiden, The Netherlands, 1993, pp. 443–485.
37 Waller, C.L., Oprea, T.I., Giolitti, A. and Marshall, G.R., J. Med. Chem., 36 (1993) 4152.
38 Oprea, T.I., Waller, C.L. and Marshall, G.R., J. Med. Chem., 37 (1994) 2206.
39 Goodford, P.J., J. Am. Chem. Soc., 28 (1985) 849.
40 Cruciani, G., Clementi, S. and Baroni, M., In Kubinyi, H. (Ed.) 3D QSAR in Drug Design: Theory, Methods and Applications, ESCOM, Leiden, The Netherlands, 1993, pp. 551–564.
41 Cruciani, G. and Watson, K.A., J. Med. Chem., 37 (1994) 2589.
42 Martin, Y., Bures, M., Danaher, E. and DeLazzer, J., In Wermuth, C.G. (Ed.) Trends in QSAR and Molecular Modelling 92 (Proceedings of the 9th European Symposium on Structure–Activity Relationships: QSAR and Molecular Modelling), ESCOM, Leiden, The Netherlands, 1993, pp. 20–27.
43 Jain, A.N., Koile, K. and Chapman, D., J. Med. Chem., 37 (1994) 2315.
44 Vinter, J., J. Comput.-Aided Mol. Design, 8 (1994) 653.
45 Oprea, T.I. and Vinter, J.G., In Sanz, F. (Ed.) Proceedings of the 10th European QSAR Symposium, J.R. Prous, Barcelona, Spain, 1995, in press.
46 Appelt, K., Perspect. Drug Discov. Design, 1 (1993) 23.
47 Arevalo, J.H., Taussig, M.J. and Wilson, I.A., Nature, 365 (1993) 859.
48 Stewart, J.J.P., J. Comput.-Aided Mol. Design, 4 (1990) 1.
49 Weisberg, S., Applied Linear Regression, 2nd ed., Wiley, New York, NY, 1985, pp. 196–239.
50 Oprea, T.I., Waller, C.L. and Marshall, G.R., Drug Des. Discov., 12 (1994) 29.
51 Brueggemeier, R.W., Li, P.-K., Moh, P.P. and Katlic, N.E., J. Steroid Biochem. Mol. Biol., 37 (1990) 379.
52 Li, P.-K. and Brueggemeier, R.W., J. Med. Chem., 33 (1990) 101.
53 Li, P.-K. and Brueggemeier, R.W., J. Steroid Biochem. Mol. Biol., 36 (1990) 533.
54 Brueggemeier, R.W., Moh, P.P., Ebrahimian, S. and Darby, M.V., J. Steroid Biochem. Mol. Biol., 44 (1993) 357.
55 Burkhart, J.P., Peet, N.P., Wright, C.L. and Johnston, J.O., J. Med. Chem., 34 (1991) 1748.
56 Peet, N.P., Johnston, J.O., Burkhart, J.P. and Wright, C.L., J. Steroid Biochem. Mol. Biol., 44 (1993) 409.
57 Johnston, J.O., Wright, C.L., Burkhart, J.P. and Peet, N.P., J. Steroid Biochem. Mol. Biol., 44 (1993) 623.
58 Di Salle, E., Briatico, G., Giudici, D., Ornati, G., Zaccheo, T., Buzzetti, F., Nesi, M. and Panzeri, A., J. Steroid Biochem. Mol. Biol., 49 (1994) 289.
59 Di Salle, E., Giudici, D., Ornati, G., Briatico, G., D'Alessio, R., Villa, V. and Lombardi, P., J. Steroid Biochem. Mol. Biol., 37 (1990) 369.
60 Abul-Hajj, Y.J., J. Steroid Biochem., 35 (1990) 139.