

Comparative molecular field analysis of non-steroidal aromatase inhibitors related to fadrozole

Maurizio Recanatini

Department of Pharmaceutical Sciences, University of Bologna, Via Belmeloro 6, I-40126 Bologna, Italy

Received 1 September 1995

Accepted 21 November 1995

Keywords: 3D QSAR; Cytochrome P-450; Antiestrogens

Summary

A series of non-steroidal inhibitors of aromatase, structurally related to fadrozole (**2**), was investigated with the aim of developing a 3D QSAR model using the Comparative Molecular Field Analysis (CoMFA) technique. The alignment of the molecules was performed following two approaches (atom-by-atom and field fit), both starting from an initial hypothesis of superimposition of fadrozole to a steroidal inhibitor (**3**). From a number of CoMFA models built with different characteristics, one was recognized as the most statistically relevant; this one is discussed in detail. The features of the 3D QSAR model are consistent with those of other 3D and QSAR models of aromatase and its inhibitors.

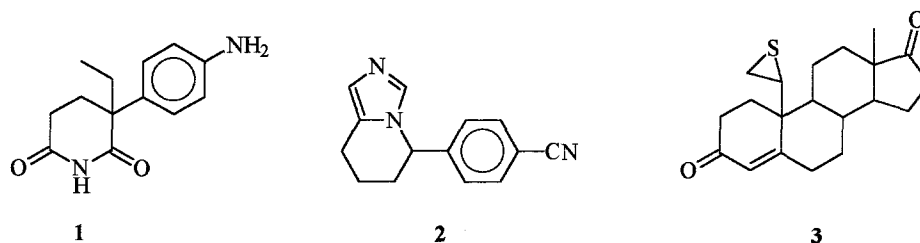
Introduction

The involvement of estrogens in the progression of several diseases, mainly typified by estrogen-dependent breast cancer, has led to the development of drugs that are able to counteract the action of these hormones. The so-called antiestrogens are divided into two classes: estrogen receptor antagonists and aromatase inhibitors. Aromatase is a cytochrome P-450 enzyme that catalyzes the terminal step in the biosynthesis of estrogens; its inhibition causes a lowering of the level of circulating hormones that is thought to be therapeutically useful. Although aromatase substrates are steroids, not all the inhibitors share the steroidal structure and, moreover, they are known to act through different mechanisms [1]. Aminoglutethimide (**1**) is a non-steroidal aromatase inhibitor presently in clinical use for the treatment of breast cancer; it is a member of the class of Type II binding aromatase inhibitors [1]. These compounds are competitive inhibitors that bind to the cytochrome P-450 enzyme in such a way that one heteroatom (N, S, O) coordinates to the iron of the porphyrin ring, causing a typical bathochromic shift in the UV absorption spectrum Soret band of the heme. The low potency and lack of selectivity of aminoglutethimide prompted the search of new, more potent and less toxic derivatives and, recently, fadrozole (**2**) was discovered after the optimization of a series of *N*-benzylimidazoles [2].

Subsequent work on fadrozole analogues led to the rationalization of the SAR of this class of compounds and to a hypothesis of interaction of the inhibitors with the enzyme based on a comparison between fadrozole and a steroidal inhibitor (19*R*)-10-thiiranylestr-4-ene-3,17-dione (**3**) [3,4]. Furet et al. observed that both molecules possess an iron-coordinating function (the unsubstituted imidazolic N atom of fadrozole and the thiiranylic S atom of **3**) and they proposed that the essential cyano group of fadrozole could mimic the carbonyl function at position 17 of the steroid [4]. If the iron-coordinating atoms of the two molecules are superimposed, the other two functions mentioned can be brought to overlap, as shown in Fig. 1. The same overlap was studied by these authors for a number of analogues of fadrozole [4], but no attempt was made to derive a quantitative three-dimensional model of the structure–activity relationships.

Since its first appearance [5], the comparative molecular field analysis (CoMFA) approach to studies of 3D QSARs for series of compounds has proved to be a valuable technique for building predictive models, and it has been applied to an ever increasing number of sets [6].

Considering the availability in the literature of a consistent set of structurally varied analogues of fadrozole tested in a single laboratory, it was thought of interest to perform a CoMFA on the series of compounds. The aim was to obtain a 3D QSAR model to be used as a guide in the design of new aromatase inhibitors. In this paper,



Scheme 1. Structures of aminogluthethimide (**1**), fadrozole (**2**) and (19*R*)-10-thiiranylestr-4-ene-3,17-dione (**3**).

the development of the model will be described and its main features will be compared with those of other aromatase inhibition models.

Methods

Biological data

The 29 compounds considered for this study are shown in Fig. 2 and their activity as *in vitro* inhibitors of human placental aromatase is reported in Table 1, expressed as

pIC_{50} . Biological data were taken from Refs. 3 and 4 and, from the series presented in these papers, only nonchiral compounds were chosen.

Molecular modeling

Molecular modeling was performed by means of the SYBYL software [7], running on a VAXStation 3100; the graphic models were studied on an IBM PC PS/VP 450DX2, using the program NITRO [8] as a graphic interface to SYBYL.

TABLE 1
BIOLOGICAL AND CONFORMATIONAL PROPERTIES OF THE COMPOUNDS INVESTIGATED

Compound	$pIC_{50 \text{ obsd}}^a$	$pIC_{50 \text{ fit}}^b$	$\Delta pIC_{50 \text{ fit}}$	$pIC_{50 \text{ pred}}^c$	$\Delta pIC_{50 \text{ pred}}$	No. of conf. ^d	ΔE^e (kcal/mol)
2a	8.49	8.22	0.27	7.98	0.51	2	0.0
2b	6.17	5.99	0.18	6.31	-0.14	2	0.0
4	9.15	8.78	0.37	8.34	0.81	2	0.0
5	8.54	8.72	-0.18	8.75	-0.21	4	0.9
6	6.63	6.87	-0.24	7.56	-0.93	4	2.4
7	8.46	8.63	-0.17	8.84	-0.38	4	0.0
8	8.11	8.35	-0.24	8.46	-0.35	2	0.0
9	8.70	8.65	0.05	8.15	0.55	5	1.0
10	8.46	8.43	0.03	8.41	0.05	4	0.0
11	6.52	6.49	0.03	6.58	-0.06	7	0.0
12	6.00	6.04	-0.04	6.60	-0.60	8	0.5
13	7.07	7.13	-0.06	7.06	0.01	7	0.4
14	8.00	7.93	0.07	8.14	-0.14	19	2.0
15	8.26	8.15	0.11	8.17	0.09	3	0.0
16	8.00	8.21	-0.21	8.51	-0.51	3	0.0
17	7.68	7.81	-0.13	8.14	-0.46	2	0.0
18	8.12	7.99	0.13	7.74	0.38	6	1.1
19	7.96	7.87	0.09	7.80	0.16	11	0.0
20	7.92	8.05	-0.13	8.11	-0.19	13	0.5
21	9.00	8.84	0.16	8.52	0.48	13	0.4
22	7.15	7.09	0.06	7.11	0.04	9	0.0
23	7.91	8.04	-0.13	8.18	-0.27	12	0.1
24	7.67	7.90	-0.23	8.11	-0.44	10	0.0
25	8.05	7.80	0.25	7.45	0.60	14	0.6
26	7.74	7.52	0.22	7.61	0.13	13	1.4
27	7.89	7.78	0.11	7.85	0.04	12	0.6
28	7.68	7.73	-0.05	7.53	0.15	10	0.4
29	7.52	7.79	-0.27	7.96	-0.44	8	0.0
30	8.24	8.31	-0.07	8.32	-0.08	10	0.9

^a Original data from Refs. 3 and 4.

^b Calculated from the non-cross-validated CoMFA equation of model C.

^c Calculated from the cross-validated CoMFA equation of model C.

^d Number of conformers generated by the random search procedure.

^e Steric energy difference between the selected and the minimum energy conformer.

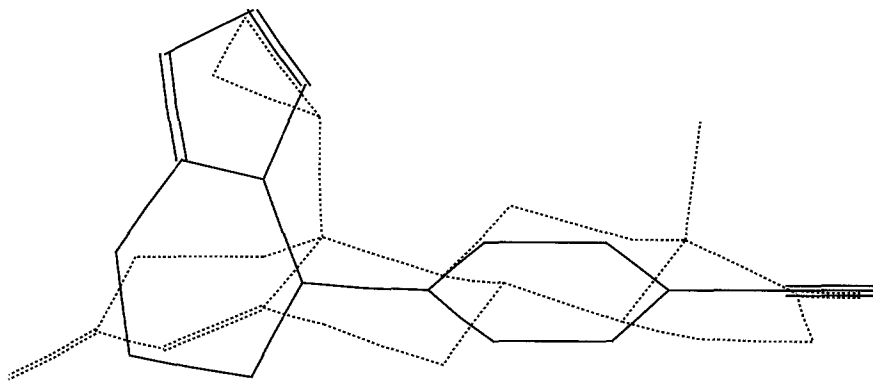


Fig. 1. Superimposition of *S*-fadrozole (**2**, solid) and (19*R*)-10-thiiranylest-4-ene-3,17-dione (**3**, dashed) as hypothesized by Furet et al. [4], showing the overlap of the iron-coordinating atoms (N1 of **2** and S of **3**) and of the 4'-CN (**2**) and 17-CO (**3**) functions.

Three-dimensional models of all molecules, except for **2a**, were built by assembling fragments from the SYBYL standard library. In the case of *S*-fadrozole (**2a**), the crystal structure [9] was used as the starting geometry. Energy minimizations were performed with MAXIMIN2 (Tripos force field [10]) without considering the electrostatic charges.

The conformational space of each molecule was sampled by means of the RANDOMSEARCH procedure of SYBYL. In this Monte Carlo-like method, a number of conformations is generated by randomly rotating selected bonds of the molecules. The degree of completeness of the

analysis can be checked by setting a number of parameters: energy cutoff, number of times one conformer is found and number of attempts to find a new conformer. The number of times that the same conformation is found was suggested as an acceptable measure of completeness [11]: it is estimated that, if a conformer is found *n* times, then there is a $(1 - (1/2)^n)$ probability that all possible conformations are found. In practice, the energy cutoff was set about 50 kcal/mol above the estimated total energy of the molecule, the maximum number of hits was 6 and the maximum number of attempts was 1000.

The conformations generated by RANDOMSEARCH

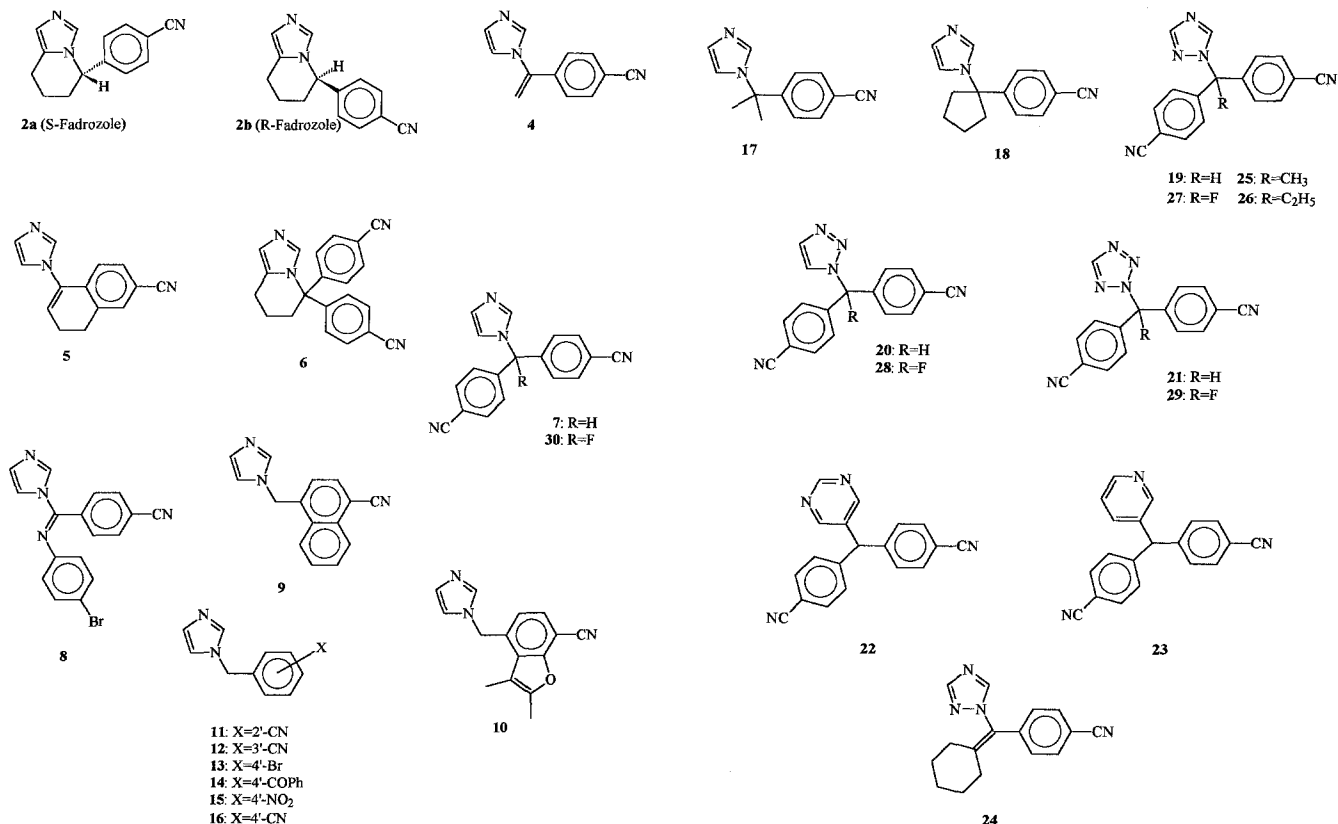


Fig. 2. Structural formulas of the compounds investigated.

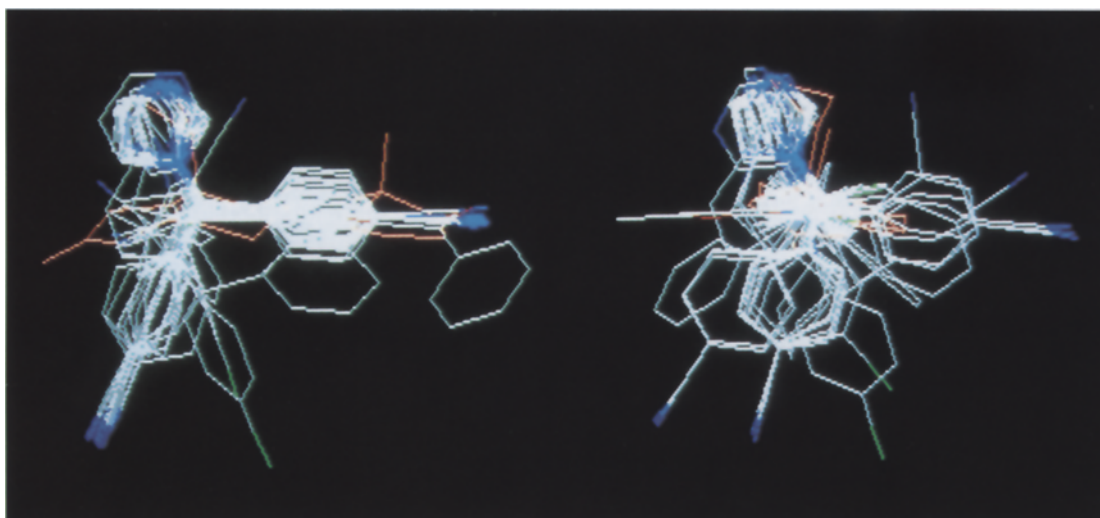


Fig. 3. Superimposition of all the non-steroidal aromatase inhibitors studied, showing the atom-by-atom alignment (orthogonal view). As a reference, the steroid (19*R*)-10-thiiranylestr-4-ene-3,17-dione (**3**, orange) is shown superimposed onto *S*-fadrozole (**2a**).

were minimized with MAXIMIN2 and were subsequently submitted to the selection of conformations for the CoMFA procedure. Electrostatic charges were calculated with the Gasteiger–Marsili method [12]. Further complete optimization of the conformations chosen for CoMFA was performed by means of AM1 [13].

CoMFA

A table was generated containing the biological activity data of the aromatase inhibitors to be studied and the CoMFA columns. The latter were obtained using an sp^3 C atom with a +1 charge as the probe; the region was created automatically and the default grid spacing was used (2 Å).

With regard to the alignment, two different strategies were employed: an atom-by-atom approach and a field fit procedure. In the first case, the best fit was searched between some atoms of a template structure and the corresponding atoms of the molecule to be aligned. *S*-fadrozole (**2a**) was used as a template and each conformation of all remaining molecules was examined in order to find the one providing the best overlap of the following atoms: the two N atoms plus the C4 atom of the imidazole ring (or a corresponding one), and the C and N

atoms of the cyano group plus the C1' and C4' atoms of the *p*-cyanobenzyl moiety. Following Furet et al. [4], the molecules bearing a second *p*-cyanophenyl ring on the central C atom were aligned so that such a ring was oriented in different positions, depending on the activity displayed by the compound. Consequently, in the poorly active compound **6**, the second *p*-cyanophenyl ring was oriented to point outwards, almost perpendicularly to the plane of the tetrahydropyridinyl fragment of *S*-fadrozole, while the active compound **7**, allowing a number of different alignments, was oriented so as to place the second *p*-cyanophenyl ring in a different direction with respect to that of **6**. Other less active molecules (**2b**, **11**, **12**) were aligned in such a way that their *p*-cyanophenyl ring overlapped the second ring of **6**. A rather unique orientation was displayed by **8**, whose *p*-bromophenyl ring could fit neither of the previously described locations of the second *p*-cyanophenyl ring.

In the field fit method as implemented in SYBYL, the resemblance of a molecule to the template is considered in terms of both electrostatic and steric fields. Thus, the selected conformation is oriented so as to minimize the differences between its field values at the lattice points and those of the template field. The template molecule

TABLE 2
TORSION VALUES OF THE ANGLES Φ_1 AND Φ_2 IN THE CONFORMATIONS OF *S*-FADROZOLE (**2a**)

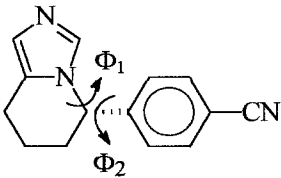
	Conformation	Origin	Φ_1	Φ_2
	Quasi-equatorial	MAXIMIN2	-46	138
	Quasi-equatorial	AM1	-46	130
	Quasi-equatorial	Furet et al. [4]	-36	120
	Quasi-equatorial	X-ray [9]	-41	123
	Quasi-axial	MAXIMIN2	-71	157
	Quasi-axial	AM1	-69	150
	Quasi-axial	Furet et al. [4]	-72	140

TABLE 3
RESULTS OF THE PLS ANALYSES

Model	Opt. no. of components	q^2	s_{cross}	r^2	s_{fit}
A	2	0.737	0.411	0.881	0.276
B	2	0.585	0.517	0.886	0.293
C	3	0.763	0.398	0.949	0.185
D	4	0.521	0.578	0.984	0.107
E	2	-0.684	1.041		
F	3	-0.896	1.127		

was again *S*-fadozole and the electrostatic and steric fields (identical to the CoMFA fields) were calculated using the Tripos force field. It has to be noted that in both alignments only rigid-body modifications of the orientations of the molecules were performed, without modifying the internal coordinates of the compounds.

The statistical analysis of the CoMFA table was performed by applying the PLS procedure to the appropriate columns and using the standard scaling method (COMFA_STD). Also, in order to reduce the actual number of columns considered, energy cutoff values of 30 kcal/mol were selected for both electrostatic and steric fields and the minimum sigma value was set to 2.0. Cross-validated PLS runs were carried out for establishing the optimum number of components to be used in the final fitting models. The number of cross-validation groups was always equal to the number of compounds (leave-one-out technique) and the optimum number of components was chosen by considering the lowest standard error of prediction (s_{cross}).

In order to check the predictive models against the appearance of spurious results (chance correlations), some analysis was done after randomization of the biological data and reassignment to the compounds.

Results

A first critical step in the construction of the CoMFA model was the attainment of a relevant conformation of *S*-fadozole (**2a**) on which all other derivatives had to be aligned. In their hypothesis of the binding mode of fadozole to the active site of aromatase, Furet et al. [4] established that the quasi-equatorial conformation of *S*-fadozole is the active one. We submitted the crystal structure of **2a** [9] to a random conformational search and obtained two conformations: one in which the *p*-cyanophenyl group is oriented quasi-equatorially and another in which this group shows a quasi-axial orientation. The geometries were optimized both with a molecular mechanics force field (MAXIMIN2) and with a semiempirical molecular orbital method (AM1). A comparison of the two torsion angles Φ_1 and Φ_2 of the resulting geometries can be made from the values reported in Table 2. It can be seen in the table that, while there is a substantial agreement between the two optimization methods, some differ-

ence exists between the values we obtained and those reported by Furet et al., particularly with regard to Φ_1 . However, the Φ_1 value measured in the crystal structure is rather close to the value obtained from both optimization procedures and the agreement of the latter values is particularly important, because it is Φ_1 that determines the distance between the imidazole N1 atom and the cyano N atom, i.e., between the two assumed pharmacophoric points of the molecule.

The conformational random search carried out on each of the remaining molecules generated variable numbers of conformers, as reported in Table 1. In many cases the conformer selected for the CoMFA was the minimum energy one, and in most cases the energy difference was less than 1 kcal/mol. In two cases (compounds **6** and **14**) this difference was higher (2.4 and 2.0 kcal/mol, respectively), but still within an acceptable range.

In Table 3, the statistics of the CoMFA models developed under different conditions are reported. Models A and B were obtained with molecules optimized only with MAXIMIN2 and with charges calculated by means of the Gasteiger–Marsili method, while in models C and D, geometries and charges were obtained with AM1. In A and C, the alignment was done atom-by-atom, as previously described, and in B and D the field fit technique was used. Models E and F were derived by ‘scrambling’ the biological data and using the molecules minimized with MAXIMIN2 and AM1, respectively.

From the data in Table 3, it immediately appears that the best models in terms of predictivity (as evaluated by the cross-validation) are those obtained with the atom-by-atom alignment: q^2 and s_{cross} of models A and C are significantly more favorable than those of models B and D. This is particularly evident when the AM1-optimized molecules are considered, i.e., in models C and D, where the field fit model D involves a higher number of components and still gives poorer statistics than C. With respect to the energy optimization methods and the charges employed, models A and C look similar from a statistical point of view. The slightly better statistics of the cross-validated model C are obtained with one component more than used in model A, and this also affects the r^2 and s_{fit} values of the fitting model. Given the more accurate refinement of the geometry and the charges in model C, this can be assumed as the final reference model for

the discussion. The atom-by-atom alignment of all the molecules, together with the steroid inhibitor (19*R*)-10-thiiranylestr-4-ene-3,17-dione (**3**) superimposed onto *S*-fadrozole as in Fig. 1, is shown in Fig. 3.

Calculated activity values from both the non-cross-validated ($\text{pIC}_{50\text{fit}}$) and the cross-validated ($\text{pIC}_{50\text{pred}}$) models C are reported in Table 1, together with the residuals; the calculated versus the actual data are plotted in Fig. 4. It can be seen that the experimental data are fairly well reproduced over the entire activity range, except for two points that deviate more than $2 \times s$. The worst calculated points are compound **4**, which is underestimated by both the descriptive and the predictive models, and compound **6**, which is overestimated by the predictive model (see the ΔpIC_{50} values in Table 1).

The contour maps of the $\text{STDEV} \times \text{COEFF}$ terms of the CoMFA equation for the steric and electrostatic contributions to the model are shown in Fig. 5. In the steric map (Fig. 5a), the region where an increase in the volume leads to increased activity is colored green, and the region where addition of steric bulk is unfavorable is colored yellow. With regard to the electrostatic field map (Fig. 5b), the red zone indicates an increase of activity for the increased positive electrostatic features, while the blue contour indicates that adding negative charge increases activity. The electrostatic contour is less well defined than the steric one, which might be the consequence of the lower contribution of the electrostatic field to the model (steric contribution: 0.58, electrostatic contribution: 0.42). Another reason might be the low variance of the electrostatic characteristics of the series, which results in poorly defined coefficients associated with the electrostatic terms of the CoMFA equation.

Discussion

The comparative molecular field analysis performed on the data of Table 1 allowed us to establish a meaningful 3D QSAR for the compounds of Fig. 2. The statistics reported in Table 3 indicate that model C, built with AM1 geometries and charges and superimposing the molecules onto *S*-fadrozole (**2a**) on the basis of a selected group of atoms, is strongly significant. The cross-validated correlation coefficient ($q^2=0.763$) is good and was obtained with only three components. Actually, model A shows a lower q^2 (0.737), but this was obtained with only two components. We take this to mean that model A is not much worse than model C, as indicated also by the similar cross-validated standard errors (model C: $s_{\text{cross}}=0.398$; model A: $s_{\text{cross}}=0.411$). The fact that a low number of components can be used while still obtaining highly significant relationships strengthens the predictive power of the model, giving confidence in the use of the 3D QSAR for the design of new derivatives. From a descriptive point of view, model C is superior to model A; it shows better statistics, even considering the presence of three components instead of two (model C: $r^2=0.949$, $s_{\text{fit}}=0.185$; model A: $r^2=0.881$, $s_{\text{fit}}=0.276$).

The 3D QSARs obtained by aligning the molecules onto *S*-fadrozole (**2a**) on the basis of their electrostatic and steric fields are much poorer relationships (model B: $q^2=0.585$, $s_{\text{cross}}=0.517$, 2 components; model D: $q^2=0.521$, $s_{\text{cross}}=0.578$, 4 components). This demonstrates again the critical role of the alignment in the development of CoMFA models [14], and in this case it turns out that a subjective way of orienting the molecules works better than a more objective procedure based on calculated

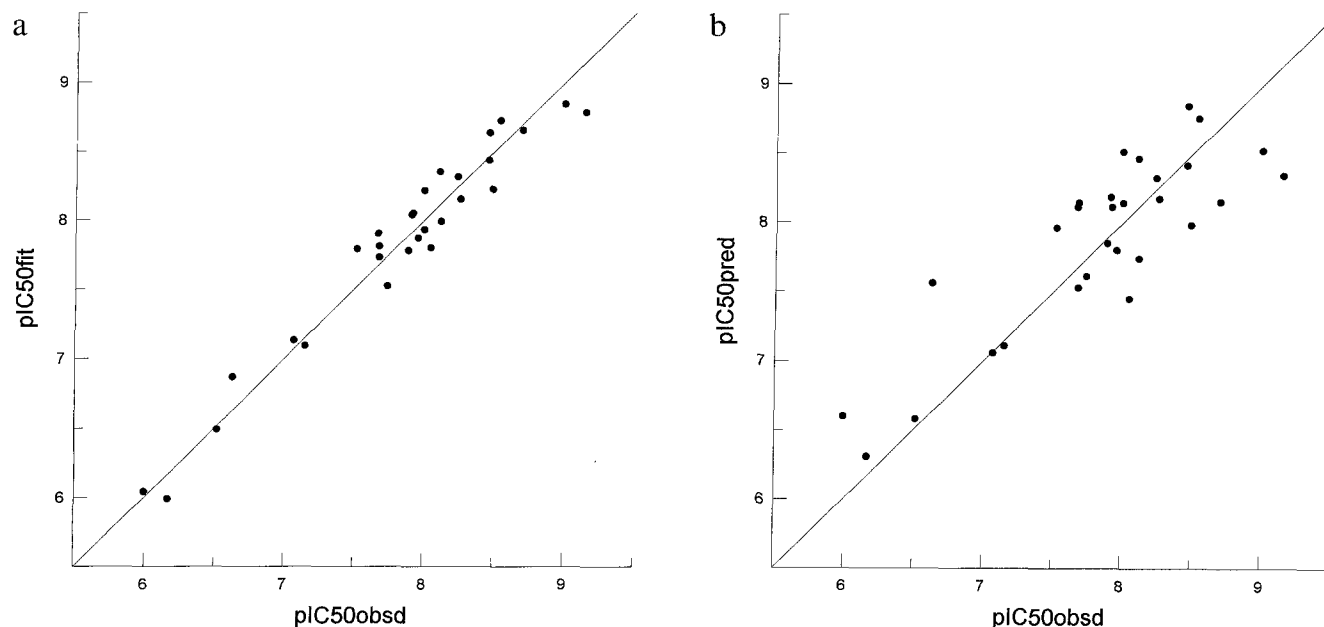


Fig. 4. Plots of calculated versus observed pIC_{50} values (see Table 1) from (a) the non-cross-validated model C and (b) the cross-validated model C.

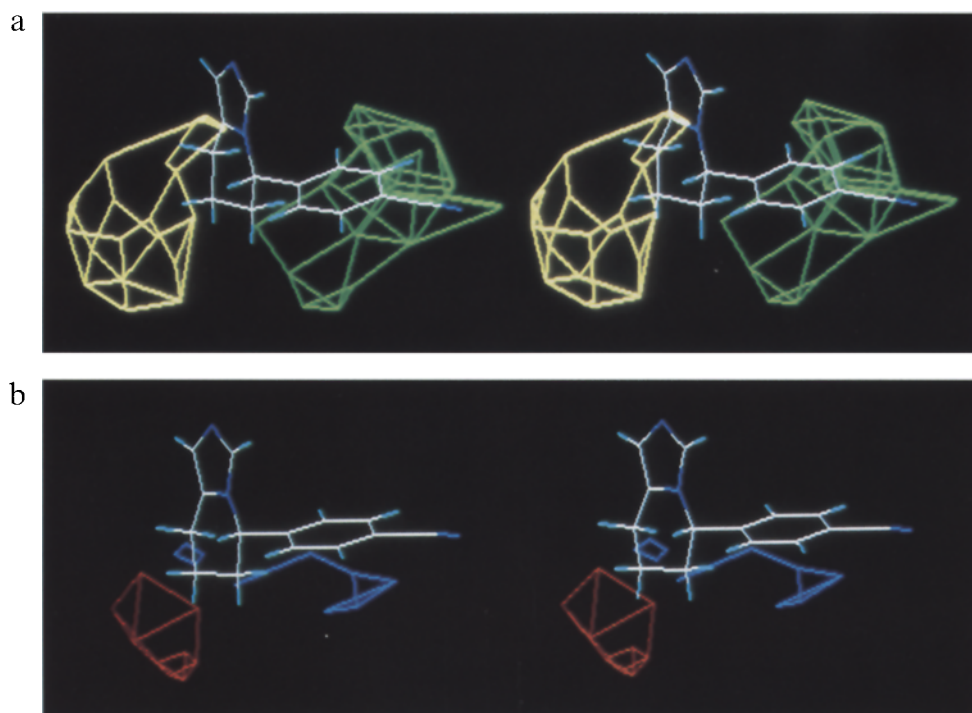


Fig. 5. Stereoview of the CoMFA contour maps from model C. (a) Steric contour: the region where increasing the volume increases activity is colored green (0.016 level) and the region where increasing the volume decreases activity is colored yellow (-0.023 level); (b) electrostatic contour: the red zone indicates an increase of activity with increasing positive charge (0.023 level) and the blue contour indicates an increase of activity with increasing negative charge (-0.030 level). *S*-fadrozole (**2a**) is shown as a reference.

fields. One reason for this behavior might be related to the rather strict congenerity of the data set. Almost all the derivatives present two constant substructures, an imidazole (or imidazole-like) ring and a *p*-cyanobenzyl moiety, and these fragments represent a great part of all molecules. An atom-by-atom alignment of these substructures might contribute to minimize the number of low-variance columns in the CoMFA table with respect to a field fit alignment, which considers the whole molecules. In the former case, the signal-to-noise ratio that is important for an efficient PLS analysis should be higher than in the latter and this might be at the basis of the higher q^2 values of models A and C.

The CoMFA contour plots illustrating the 3D QSAR (model C) of the set of aromatase inhibitors are shown in Fig. 5, with *S*-fadrozole (**2a**) as the reference structure. Both the green and the yellow steric regions, indicating the positive and negative contributions, respectively, are well defined and centered around well-identified features of the *S*-fadrozole molecule (Fig. 5a). The region where adding steric bulk leads to an increase of activity corresponds to a volume approximately around the meta and para positions of the *p*-cyanophenyl ring, while the unfavorable yellow zone is located mostly over the tetrahydropyridine moiety and is occupied by the fourth substituent (in this case an H atom) of the benzylic C atom of the benzylimidazole fragment.

It is noteworthy how the groups oriented under the

longitudinal axis of the *p*-cyanobenzyl fragment of *S*-fadrozole do not seem to contribute, neither favorably nor unfavorably, to the steric part of the 3D QSAR, despite characterizing rather precisely an area that, in the model proposed by Furet et al. [4], corresponds to the α -side of the steroid ligands (see Fig. 3).

The electrostatic contours in Fig. 5b are less clear than the steric ones and present some ambiguity with regard to the interpretation. The loose blue contours indicate areas where increasing the negative charge of a fragment increases the activity. Considering the fact that most of the blue zone lies just beneath the aromatic ring of *S*-fadrozole, it may be argued that similar electron-rich groups favorably occupy this area. The small blue rectangle on top of the benzylic H atom does not seem to be relevant. The apparently better defined red contour indicates an area that is favorable for positively polarized groups and unfavorable for negative ones. It extends in a region of space where the only identifiable fragments are the alkyl moieties bound to the central benzyl atom of compounds **18** and **24**. However, it may be questionable that alkyl groups would give rise to electrostatic CoMFA contours. Examination of a hydrophobic field could be useful in order to locate possible favorable interactions in that area.

A point deserving some comment is the poor performance of the model with regard to the two compounds **4** and **6**, which behave like outliers (see Results). Consider-

ing the fact that **4** is the most active derivative of the series, its large underestimation is suggestive of a limitation of the data set in terms of spanning the parameter space. On the other hand, other compounds (**5**, **8**, **24**) feature a double bond in the same position as **4** and do not deviate significantly. An explanation might be that the π electrons of the vinyl group of **4** are more available to contact the enzyme than those of the other mentioned compounds and that this factor is not accounted for by the model. Compound **6** is severely overestimated from the predictive model C (more than $2 \times s_{\text{cross}}$). The constrained geometry of this molecule is probably at the basis of the deviating behavior, which might originate from the impossibility of precisely locating both the phenyl rings, one in the favorable and one in the forbidden region of the model.

The 3D QSAR illustrated above for a series of analogues of fadrozole represents both a predictive and a descriptive model. The verification of the predictivity will be made using a test set of new derivatives presently under development. The descriptive model can be compared with other models for the interaction of inhibitors with aromatase that have recently appeared in the literature [4,15,18,19]. The alignment proposed in this paper and validated by the cross-validation procedure might be considered as a starting point for proposing an interaction model of aromatase inhibitors, to be integrated with the information coming from other studies.

The already mentioned binding model of Furet et al. [4] is at the basis of the atom-by-atom alignment carried out in this work, and the main features of that model are in agreement with the present one. Particularly, Furet et al. found an allowed region in space corresponding to the α -side of the steroid inhibitor and a forbidden region oriented perpendicularly to the longitudinal axis of the steroid. The superimposition picture of Fig. 3 shows that the first region is actually occupied by aromatic residues of fadrozole analogues, even if it is not particularly relevant from a quantitative point of view. However, it is interesting to note that these aromatic groups seem to individuate two different subregions, one characterized by the polar CN residues and one by the more hydrophobic *p*-bromophenyl groups. It might be assumed that a wider variance of the structures in this area would eventually provide a quantitative definition of the zone. On the other hand, the region protruding outwards perpendicularly to the long axis of the steroid can be identified with the yellow contour of Fig. 5a. This region is quantitatively relevant in the model presented here and the negative effect on the activity appears to be of a steric nature, at least for the set of Table 1.

The main difference between the two models is that in the alignment of the compounds we did not emphasize the eclipsing of the imidazole ring of the inhibitors with the N-Fe-coordinating bond of the heme group, thus

forcing the overlap of the imidazole rings of all molecules. This would have implied looking for higher energy conformations of some molecules and also locating the common *p*-cyano group in a variety of different positions in space, which would be inconsistent with the two-point pharmacophore assumed as the starting hypothesis.

A QSAR model for a set of substituted dichlorophenyl inhibitors of aromatase was recently published by Nagy et al. [15]. These authors studied one series of dichlorophenylmethanes and one series of dichlorophenylmethanols, both presenting different heterocycles as the fourth substituent on the central C atom. A number of descriptors were examined, among which were shape descriptors, electronic parameters (dipole moment, HOMO and LUMO energies, integrated potential field difference), as well as log P values and volume parameters. The most significant result for both series was a parabolic relationship in D, where D is the distance between the central C atom and an N atom of the heterocyclic ring (the β nitrogen, if present). The calculated optimum value of D was 3.6 ± 0.1 Å. Interestingly, the authors underlined the absence of any relationship with electronic parameters and with log P.

In the series of Table 1, less variance is present in the structure of the heterocycles (see Fig. 2); however, the distance between the benzyl C atom and the β nitrogen of the heterocycle ranges from 3.55 to 3.79 Å, with a mean value of 3.66 ± 0.05 Å. The possible contribution to the 3D QSAR of log P was explored by calculating the CLOGP values (not reported) of the compounds and inserting this new column in the CoMFA table. No reduction of the residual variance was found after considering the CLOGP column, even when the column was scaled 10, 100 or 1000 times as suggested by Greco et al. [16], in order to appropriately weight the hydrophobic parameter in the CoMFA table.

In recent years, some papers have been published aimed at studying the structure and functions of aromatase by means of molecular modeling and site-directed mutagenesis [17–19]. These studies provided a description of the hypothetical active site of the enzyme and of the possible interactions of steroid substrates with active site residues. Considering that the X-ray crystal structure of aromatase is still lacking, computer-generated models might be tentatively used to interpret SAR studies and an attempt can be made to relate one of these models to the 3D QSAR reported here.

Laughton et al. [18] derived a detailed model of aromatase on the basis of the cytochrome P-450_{cam} X-ray structure and of the known sequence of the cytochrome P-450_{arom} (aromatase). Some groups of active site residues were identified and their interactions with a steroid ligand were examined. Some of the observations reported in their paper are consistent with our 3D QSAR. For example, the Laughton model reveals the presence of two

phenylalanine residues (Phe²³⁴ and Phe²³⁵) in the α region of the steroid. Considering the orientation of the fadrozole analogues with respect to steroid **3** (see Fig. 3), the phenylalanine phenyl rings might be involved in some interaction with the *p*-cyanophenyl groups or the *p*-bromophenyl groups supposed to occupy the same region. Another interesting observation is the presence of a histidine residue (His⁴⁷⁵) in the area corresponding to the lower edge of the steroid A and B rings. This residue might represent the steric limitation of the hydrophobic site hosting the substrate and might be the cause of the steric hindrance revealed by the CoMFA model (yellow contour, Fig. 5a). A subsequent refinement of the aromatase model, assisted by site-directed mutagenesis [19], confirmed the position of the histidine residue, placing it even closer to the C4 position of the steroid. Finally, Laughton et al. [18] hypothesized that Tyr²⁴⁴ on the β face and Ile³⁰⁵ on the α face of the steroid severely restrict the space available to the D ring. This might support the possibility suggested by the green CoMFA contour of Fig. 5a of increasing the volume correspondingly to the meta positions of the *p*-cyanophenyl ring of *S*-fadrozole; the plane of this ring is almost aligned to that of the C and D rings of the steroid ligand (see Fig. 1).

Conclusions

A 3D QSAR model was derived for a set of aromatase inhibitors related to fadrozole (**2**), which allowed to individuate regions in the space surrounding the molecule available for favorable (and unfavorable) steric and electrostatic modifications. Based on a previously hypothesized superimposition of *S*-fadrozole on a steroid inhibitor [4], the model was tentatively used as a starting point for a more general description of the interactions of aromatase with inhibitors consistent with other reports recently published. Further work is in progress for the validation of the predictive power of the model and for its extension to other classes of inhibitors.

Acknowledgements

This work was supported by a grant of MURST. The author thanks A. Cavalli for technical assistance.

References

- 1 Cole, A.P. and Robinson, C.H., *J. Med. Chem.*, 33 (1990) 2933.
- 2 Browne, L.J., Gude, C., Rodriguez, H., Steele, R.E. and Bhatnagar, A., *J. Med. Chem.*, 34 (1991) 725.
- 3 Lang, M., Batzl, C., Furet, P., Bowman, R., Häusler, A. and Bhatnagar, A.S., *J. Steroid Biochem. Mol. Biol.*, 44 (1993) 421.
- 4 Furet, P., Batzl, C., Bhatnagar, A., Francotte, E., Rihs, G. and Lang, M., *J. Med. Chem.*, 36 (1993) 1393.
- 5 Cramer III, R.D., Patterson, D.E. and Bunce, J.D., *J. Am. Chem. Soc.*, 110 (1988) 5959.
- 6 Thibaut, U., In Kubinyi, H. (Ed.) *3D QSAR in Drug Design: Theory, Methods and Applications*, ESCOM, Leiden, 1993, pp. 661–696.
- 7 SYBYL Molecular Modeling System (v. 5.5), Tripos Associates, St. Louis, MO, 1992.
- 8 NITRO (v. 6.0), Tripos Associates, St. Louis, MO, 1992.
- 9 Van Roey, P., Bullion, K.A., Osawa, Y., Browne, L.J., Bowman, R.M. and Braun, D.G., *J. Enzyme Inhib.*, 5 (1991) 119.
- 10 Clark, M., Cramer III, R.D. and Van Opdenbosch, N., *J. Comput. Chem.*, 10 (1989) 982.
- 11 Saunders, M., *J. Am. Chem. Soc.*, 109 (1987) 3150.
- 12 Gasteiger, J. and Marsili, M., *Tetrahedron*, 36 (1980) 3219.
- 13 Dewar, M.J.S., Zoebisch, E.G., Healy, E.F. and Stewart, J.J.P., *J. Am. Chem. Soc.*, 107 (1985) 3902.
- 14 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, 1993, pp. 443–485.
- 15 Nagy, P.I., Tokarski, J. and Hopfinger, A.J., *J. Chem. Inf. Comput. Sci.*, 34 (1994) 1190.
- 16 Greco, G., Novellino, E., Silipo, C. and Vittoria, A., *Quant. Struct.-Act. Relatsh.*, 11 (1992) 461.
- 17 Graham-Lorence, S., Khalil, M.W., Lorence, M.C., Mendelson, C.R. and Simpson, E.R., *J. Biol. Chem.*, 266 (1991) 11939.
- 18 Laughton, C.A., Zvelebil, M.J.J.M. and Neidle, S., *J. Steroid Biochem. Mol. Biol.*, 44 (1993) 399.
- 19 Zhou, D., Cam, L.L., Laughton, C.A., Korzekwa, K.R. and Chen, S., *J. Biol. Chem.*, 269 (1994) 19501.