

Combining Computational and Biochemical Studies for a Rationale on the Anti-Aromatase Activity of Natural Polyphenols

Marco A. C. Neves,^[a] Teresa C. P. Dinis,^[b] Giorgio Colombo,^{*[c]} and M. Luisa Sá e Melo^{*[a]}

Aromatase, an enzyme of the cytochrome P450 family, is a very important pharmacological target, particularly for the treatment of breast cancer. The anti-aromatase activity of a set of natural polyphenolic compounds was evaluated in vitro. Strong aromatase inhibitors including flavones, flavanones, resveratrol, and oleuropein, with activities comparable to that of the reference anti-aromatase drug aminoglutethimide, were identified. Through the application of molecular modeling techniques based on grid-independent descriptors and molecular interaction fields, the major physicochemical features associated with inhibitory activi-

ty were disclosed, and a putative virtual active site of aromatase was proposed. Docking of the inhibitors into a 3D homology model structure of the enzyme defined a common binding mode for the small molecules under investigation. The good correlation between computational and biological results provides the first rationalization of the anti-aromatase activity of polyphenolic compounds. Moreover, the information generated in this approach should be further exploited for the design of new aromatase inhibitors.

Introduction

Cancer is one of the major causes of mortality in developed countries.^[1] Breast cancer, in particular, is the most frequent in postmenopausal women, and about 80% of breast cancers express hormone receptors that are responsive to the mitogenic effects of estrogens.^[2] In postmenopausal women, estrogens are no longer synthesized in the ovaries, but other tissues such as adipose and breast tissue have the ability to produce them.^[3] Thus, estrogen levels in breast tissue are maintained nearly at the same level as in premenopausal women, suggesting that estrogen production in or near the tumor can be determinant for its progression. Aromatase is a cytochrome P450 (CYP19) enzyme that catalyzes the conversion of androstenedione and testosterone to the aromatic estrogenic steroids estrone and estradiol, respectively. As such, suppression of estrogen biosynthesis by a mechanism of aromatase inhibition represents an effective approach for the treatment of hormone-sensitive breast cancer. The importance of this enzyme in breast cancer development has led to intensive research aiming for new compounds with the ability to inhibit aromatase, ranging from androstenedione analogues to nonsteroidal derivatives.^[4]

Besides these molecules, natural polyphenolic compounds have been recognized as aromatase inhibitors with moderate to strong potency.^[5] Flavonoids are the most abundant polyphenols in nature and share a common phenylchromanone structure.^[6] Other classes of phenolic compounds include coumarins, found in fruits and green tea, and stilbenes, like resveratrol, one of the main components of red wine. Oleuropein, the major polyphenolic constituent of olive oil, is another

structurally different phenolic compound.^[7] All these compounds have demonstrated a host of biological activities, displaying excellent antioxidant,^[8,9] antitumor,^[10,11] antiviral,^[12] antimicrobial,^[13] and anti-atherogenic^[14,15] properties.

The use of natural compounds as starting leads represents an attractive strategy for the discovery of new breast cancer drugs based on the aromatase inhibition mechanism. Several examples have recently been documented: derivatization with imidazole, triazole, and pyridinyl moieties markedly increased the anti-aromatase potency of flavones and isoflavones through coordination between the nitrogen-containing heterocycle and the Fe²⁺ ion of the aromatase heme group.^[16,17] However, to fully understand the pharmaceutical and thera-

[a] M. A. C. Neves, Prof. Dr. M. L. Sá e Melo
Centro de Estudos Farmacêuticos, Lab. Química Farmacêutica
Faculdade de Farmácia, Universidade de Coimbra
Rua do Norte, 3000-295 Coimbra (Portugal)
Fax: (+351) 239 827 126
E-mail: samelo@ci.uc.pt

[b] Prof. Dr. T. C. P. Dinis
Centro de Neurociências, Lab. Bioquímica
Faculdade de Farmácia, Universidade de Coimbra
Couraça dos Apóstolos 51, 3000-295 Coimbra (Portugal)

[c] Dr. G. Colombo
Istituto di Chimica del Riconoscimento Molecolare, CNR
Via Mario Bianco 9, 20131 Milano (Italy)
Fax: (+39) 022 890 1239
E-mail: g.colombo@icrm.cnr.it

Supporting information for this article is available on the WWW under <http://www.chemmedchem.org> or from the author.

peutic potential of these natural compounds in the treatment of breast cancer, a deeper rational study of the physicochemical determinants of their aromatase binding is needed.

To this end, we used a combination of experimental and theoretical approaches to correlate concentration–response curves of *in vitro* aromatase inhibition with calculated alignment-independent three-dimensional structure descriptors.^[18] Based on these preliminary data, molecular interaction fields (MIFs)^[19] were translated into the main relevant non-bonded interactions expected to occur in the process of enzyme–inhibitor recognition and binding, allowing us to obtain a representation of the putative active site. Finally, this virtual receptor site was compared with the active site structure of a homology model of the enzyme.^[20] The good agreement between hydrophobic and hydrophilic portions of the 3D QSAR virtual receptor site with the homology model prompted us to perform semi-flexible docking at the binding cavity using the GLUE^[21] v 1.0 docking software. With this strategy we were able to identify a common binding mode in all inhibitors tested and propose structure–activity rules that might be extrapolated to other types of compounds and used to optimize the activity of existing derivatives.

Compounds

For the evaluation of the anti-aromatase potency by a biochemical assay, we tested 12 flavones (Table 1), five flavanones (Table 2), one anthocyanin and one anthocyanidin (Table 3), four coumarins (Table 4), and *trans*-resveratrol and oleuropein (Table 5). The first-generation aromatase inhibitor, aminoglutethimide (AG), was also tested and used as a reference compound (Table 5).

Biochemical assays

The compounds in this study were tested for the ability to inhibit aromatase (CYP19), which performs the aromatization of ring A of androgens into estrogens. Aromatase inhibition was studied through an assay reported by Siiteri and Thompson^[22] by using human placental microsomes as the source of aromatase.^[23] Full concentration–response curves were obtained, allowing the determination of the half-maximal inhibitory concentration (IC₅₀). The concentration–response curves for the most active compound 5,7,3',4'-tetrahydroxyflavanone **14**, flavone **12**, and AG, the reference inhibitor, are shown in Figure 1. The IC₅₀ values obtained are shown in Tables 1–5.

Results and Discussion

Biochemical evaluation

Several natural compounds, with scaffolds ranging from low to high molecular weight and with various degrees of hydrophobicity, were considered in this study. Flavones inhibited aromatase with a potency between 7.2 and 100 μM (Table 1). Several structural variations such as hydroxylation and methoxylation at C5, C7, C3', C4', and C5' were considered in order to gener-

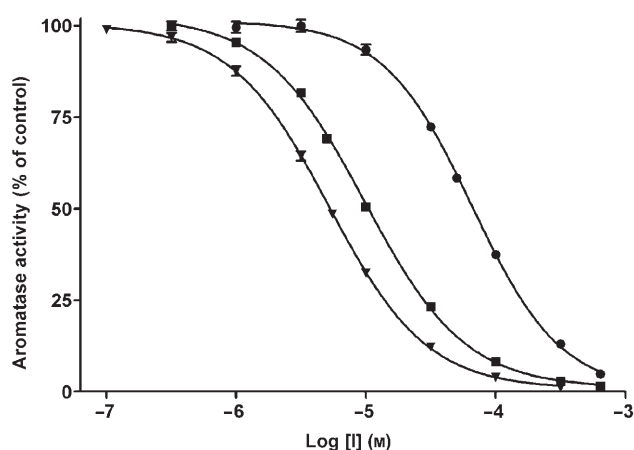


Figure 1. Concentration–response curves of compounds **14** (▼) and **12** (●). Aminoglutethimide (■) was tested as a reference aromatase inhibitor. Each point represents the mean of three assays performed in triplicate ($n=3$), and the vertical bars indicate standard error of the mean (SEM). The data were analyzed by nonlinear regression using a sigmoidal concentration–response curve with variable slope.

Table 1. Estimated binding energies (EBE) and aromatase inhibitory activities of flavones 1–12.

Compd						EBE [kcal mol ⁻¹] ^[a]	IC ₅₀ [μM] ^[b]
	R ¹	R ²	R ³	R ⁴	R ⁵		
1	OH	OH	H	H	H	-16.5	8.9 ± 0.1
2	OH	OH	H	OH	H	-16.2	15 ± 0.2
3	OH	OH	OH	OH	H	-16.1	8.6 ± 0.07
4	OH	OH	OH	OCH ₃	H	-15.6	27 ± 0.7
5	OH	OH	OCH ₃	OH	H	-15.8	7.2 ± 0.09
6	H	OH	H	H	H	-16.2	8.2 ± 0.1
7	H	OH	OH	OH	H	-15.4	38 ± 0.5
8	H	OH	OH	OH	OH	-16.1	45 ± 0.7
9	H	H	OH	OH	H	-15.1	100 ± 3.5
10	H	H	OCH ₃	OH	H	-16.0	73 ± 1.6
11	H	H	OCH ₃	OCH ₃	H	-15.8	42 ± 0.8
12	H	H	H	H	H	-16.2	67 ± 1.0

[a] Binding energies were estimated on a model of the aromatase active site using the GRID force field implemented in the GLUE^[21] docking software. [b] Results shown are the mean ± SEM.

ate a high number of hypotheses about the spatial distribution of hydrophilic, hydrogen bond donors and acceptors, and small cavities within the active site of the enzyme. Most of these molecules had stronger activity than the unsubstituted flavone **12**, confirming the potential of this type of inhibitor for lead optimization.

Hydroxylation at C7 markedly increased the anti-aromatase activity, as observed in flavones **1**, **3**, **5**, and **6**, which have stronger activity than AG, the reference aromatase inhibitor tested. These molecules inhibited aromatase with 8- to 9-fold greater activity than the unsubstituted flavone **12**. Compound

6, with just a single hydroxy group at **C7**, was the least functionalized flavone with potent aromatase inhibition activity. Furthermore, comparing the activities of flavone **7** and flavone **9**, hydroxylation at **C7** provided a 3-fold increase in inhibitory potency.

Simultaneous hydroxylation at **C5** and **C7** was also important to the anti-aromatase activity, as observed with flavones **1**, **3**, and **5**. Comparison of the inhibitory potencies of compounds **3** and **5** (respectively 8.6 and 7.2 μM) with those of compounds **9** and **10** (respectively 100 and 73 μM), simultaneous hydroxylation at **C5** and **C7** appears to play an important role in aromatase inhibition. Hydroxylation at **C4'** is present in some of the most active flavones, compounds **2**, **3**, and **5**, and seems to contribute, at least partially, to the anti-aromatase activity of these molecules.

Flavanones lack the **C2** unsaturation present in flavones, which decreases the planarity at ring **C** and introduces new stereocenters in the molecule, making it chiral. Conformational analysis (described in more detail below in the Experimental Section) confirms, however, that both global minimum-energy conformations of flavone **12** and flavanone **17** are very well superimposed, particularly at the 4-oxo group and at **C5**, **C7**, and **C4'** (Figure 2).

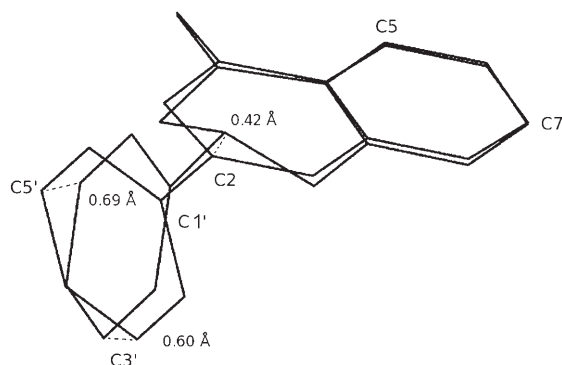
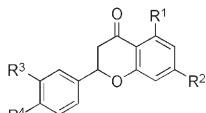


Figure 2. Superimposition of the global minimum-energy conformation found for flavone **12** and flavanone **17** using the Merck molecular force field (MMFF)^[34] and a water-generalized Born equation/surface area (GB/SA)^[35] continuum solvation. Distances separating equivalent **C2**, **C3'**, and **C5'** are displayed. The RMSD of both structures is equal to 0.35 Å.

As expected, substituted flavanones had greater activity than the nonfunctionalized flavanone **17**. The effect of substitutions is dependent on their positions (Table 2) and followed a similar pattern to that observed with flavones. However, compound **15** was less potent than flavone **5**, which might be the result of ring **B** rotation around the **C2**–**C1'** axis due to the transition from sp^2 to sp^3 geometry at **C2**, imposing a different spatial disposition for the methoxy group at **C3'** (Figure 2).

The anthocyanidin **18**, cyanidin (Table 3), shares some structural similarities with the most potent flavone and flavanone inhibitors tested, namely a similar scaffold and hydroxy groups at **C5**, **C7**, **C3'** and **C4'**. However, the anti-aromatase activity is much lower ($IC_{50} = 72 \mu\text{M}$). Two structural differences may explain this: the extended hydroxylations at positions **C3** and

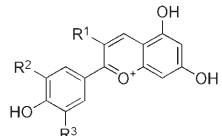
Table 2. Estimated binding energies (EBE) and aromatase inhibitory activities of flavanones **13**–**17**.



Compd	R ¹	R ²	R ³	R ⁴	EBE [kcal mol ⁻¹] ^[a]	IC ₅₀ [μM] ^[b]
13	OH	OH	H	H	-16.1	10 ± 0.1
14	OH	OH	OH	OH	-17.3	5.3 ± 0.06
15	OH	OH	OCH ₃	OH	-15.9	25 ± 0.4
16	H	OH	H	H	-16.9	10 ± 0.1
17	H	H	H	H	-16.5	32 ± 0.3

[a] Binding energies were estimated on a model of the aromatase active site using the GRID force field implemented in the GLUE^[21] docking software. [b] Results shown are the mean ± SEM.

Table 3. Estimated binding energies (EBE) and aromatase inhibitory activities of cyanidin **18** and malvidin-3-*O*-glucoside **19**.



Compd	R ¹	R ²	R ³	EBE [kcal mol ⁻¹] ^[a]	IC ₅₀ [μM] ^[b]
18	OH	OH	OH	-14.8	72 ± 0.7
19	OGlc	OCH ₃	OCH ₃	-11.6	299 ± 3.4

[a] Binding energies were estimated on a model of the aromatase active site using the GRID force field implemented in the GLUE^[21] docking software. [b] Results shown are the mean ± SEM.

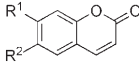
C5', which agree with the decreased activity of compound **8** relative to the less hydroxylated EBE compound **7**, and the absence of a 4-oxo acceptor group.

On the other hand, anthocyanin **19**, malvidin-3-*O*-glucoside, failed to show significant anti-aromatase activity in our assay conditions (Table 3). This compound shares the 5,7,4'-trihydroxy and 3'-methoxy groups found in the strongest flavone inhibitor tested, **5**, but a glucose moiety at **C3** might impose a different alignment at the aromatase active site. In contrast, oleuropein **25**, a large glycosylated molecule, shows surprisingly strong anti-aromatase activity (Table 5). Unlike malvidin-3-*O*-glucoside **19**, oleuropein **25** is a highly flexible molecule and may therefore align several hydrogen bond donor and acceptor groups in favorable positions on the active site of aromatase.

The coumarins studied showed greatly decreased anti-aromatase activity (Table 4). Low molecular volumes and simple scaffolds may be responsible for their poor interaction with the aromatase active site.

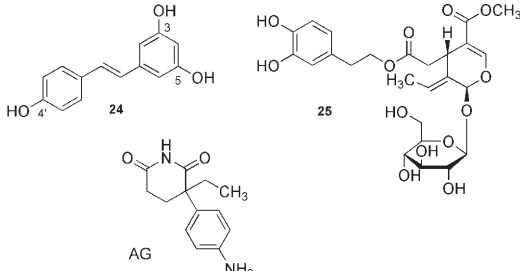
Although not equipped with the flavone phenylchromanone scaffold, *trans*-resveratrol **24** (Table 5) has a simple hydrophobic skeleton with two aromatic rings connected by an ethylene bridge, with hydroxy groups at **C3**, **C5**, and **C4'**, positions analogous to **C5**, **C7**, and **C4'** of the flavonoid inhibitors tested. In-

Table 4. Estimated binding energies (EBE) and aromatase inhibitory activities of coumarins **20–23**.

Compd			EBE [kcal mol ⁻¹] ^[a]	IC ₅₀ [μM]
	R ¹	R ²		
20	OH	OH	-11.4	NO ^[b]
21	OH	OCH ₃	-11.9	> 640
22	OCH ₃	OH	-12.7	> 640
23	OCH ₃	OCH ₃	-12.9	> 640

[a] Binding energies were estimated on a model of the aromatase active site using the GRID force field implemented in the GLUE^[21] docking software. [b] Inhibition was not observed at concentrations ≤ 640 μM.

Table 5. Estimated binding energies (EBE) and aromatase inhibitory activities of *trans*-resveratrol (**24**), oleuropein (**25**) and DL-aminoglutethimide (AG).

Compd			IC ₅₀ [μM] ^[b]
	EBE [kcal mol ⁻¹] ^[a]		
24	-13.0		12.8 ± 0.2
25	-16.5		27 ± 2.3
AG	ND ^[c]		10 ± 0.09

[a] Binding energies were estimated on a model of the aromatase active site using the GRID force field implemented in the GLUE^[21] docking software. [b] Results shown are the mean ± SEM. [c] Not determined.

Interestingly, *trans*-resveratrol is a good aromatase inhibitor (IC₅₀ = 12.8 μM) with potency similar to that of the related flavone **2**. This suggests that the lack of the 1-oxa and 4-oxo functionalities of the flavone scaffold is well balanced by the three hydrogen bond donor/acceptor groups with a good spatial disposition present in *trans*-resveratrol **24**.

3D QSAR with GRIND descriptors

The three-dimensional quantitative structure–activity relationships (3D QSAR) between the anti-aromatase activity of the flavones in this study and their 3D properties were analyzed using grid-independent descriptors (GRIND).^[18] The bioactive conformation of the flavones on the active site of aromatase is unknown. Therefore, the minimum-energy conformation, obtained by a conformational search described in detail in the Experimental Section, was chosen as representative. Although the quality of the model is critically dependent on the 3D conformation chosen, the extended aromatic conjugation makes these molecules rigid and therefore less prone to conformational changes.

GRIND descriptors are alignment-independent descriptors that represent important non-bonding interactions in the recognition and binding of the active site of a macromolecular target. A two-step procedure, described in detail in the Experimental Section, allowed the calculation of these descriptors from molecular interaction fields (MIFs)^[19] calculated with three chemical probes: DRY, O, and N1. These probes report the interaction between the molecules under study and a hydrophobic group, a hydrogen bond acceptor, and a hydrogen bond donor group, respectively. The MIFs chosen represent important types of interactions expected to guide the binding of flavones to the active site of aromatase. Besides chemical complementarity, good shape fit between the ligand and the protein is essential to binding. Therefore, the shape probe (TIP)^[24] was used to search for highly convex regions of the molecules. These regions interact most with the binding cavity. Therefore, 10 correlograms were obtained: four autocorrelograms (DRY–DRY, O–O, N1–N1, and TIP–TIP) and six cross-correlograms (DRY–O, DRY–N1, DRY–TIP, O–N1, O–TIP, and N1–TIP). Multivariate analysis with partial least-squares regression (PLS) was used to correlate the calculated descriptors with the activity, and the quality of the model was evaluated by the predictive correlation coefficient (q^2), obtained by leave-one-out (LOO) cross-validation or by a 3 random groups (3RG) procedure.

3D QSAR studies usually require the calculation of hundreds of descriptors, most of them not correlated with the activity. Therefore, the fractional factorial design (FFD),^[25] a variable selection procedure implemented in ALMOND^[21] v 3.3, was performed, excluding variables that increase the standard deviation of errors of prediction (SDEP). Variables decreasing the SDEP or those with an unclear effect were retained.

Optimal predictive ability for the anti-aromatase 3D QSAR model appears with a model dimensionality of three latent variables ($q^2_{\text{LOO}} = 0.85$ and $q^2_{\text{3RG}} = 0.53$) when it is able to explain 98% of the anti-aromatase activity variance ($r^2 = 0.98$). The residuals are shown in Table 6, and a plot of the predicted activi-

Table 6. LOO and 3RG cross-validation of the PLS 3D QSAR regression model.

Compd	pIC ₅₀ exptl ^[a]	pIC ₅₀ pred (LOO) ^[b]	δ (LOO) ^[d]	pIC ₅₀ pred (3RG) ^[c]	δ (3RG) ^[e]
1	5.05	5.10	-0.05	4.91	0.14
2	4.83	5.01	-0.18	4.94	-0.11
3	5.07	4.84	0.23	4.77	0.30
4	4.56	4.50	0.06	4.53	0.03
5	5.14	4.93	0.21	4.84	0.30
6	5.09	4.85	0.24	4.70	0.39
7	4.42	4.57	-0.15	4.62	-0.20
8	4.35	4.50	-0.15	4.64	-0.29
9	4.00	4.12	-0.12	4.24	-0.24
10	4.14	4.32	-0.18	4.40	-0.26
11	4.38	4.44	-0.06	4.46	-0.08
12	4.17	4.19	-0.02	4.33	-0.16

[a] pIC₅₀ determined experimentally. [b] pIC₅₀ predicted by leave-one-out (LOO) cross-validation. [c] pIC₅₀ predicted by 3 random groups (3RG) cross-validation. [d] Residuals (pIC₅₀exptl - pIC₅₀pred) from LOO cross-validation. [e] Residuals (pIC₅₀exptl - pIC₅₀pred) from 3RG cross-validation.

ty (LOO) versus the experimental activity is represented in Figure 3. All compounds were predicted with good confidence.

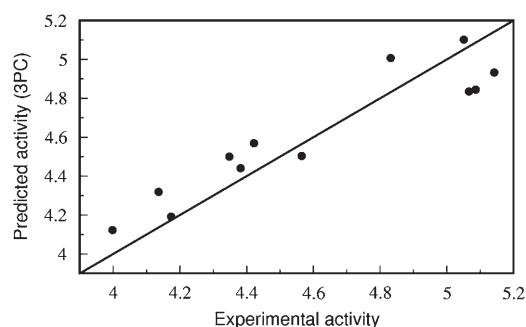


Figure 3. Plot of the predicted versus experimental pIC_{50} from the optimal PLS model with three latent variables ($q^2_{LOO} = 0.85$).

Further understanding of the PLS model was pursued using the PLS coefficients of the variables in study. Variables with high PLS coefficients are important for the biological activity fitting, whereas those with low coefficients are less important. Positive coefficients have a direct impact on the biological activity (that is, the greater the interaction energy, the stronger the inhibitor), whereas negative coefficients have an inverse impact. The PLS coefficients are plotted in Figure 4, and the most influential variables are indicated with an arrow. Interestingly, only a few variables have coefficients of high absolute value and therefore have critical importance to the model fitting. These arise from the O–O, O–N1, and O–TIP correlograms (direct

impact) and the N1–N1, DRY–O, O–N1, and N1–TIP correlograms (inverse impact), and are interpreted in Table 7.

Virtual receptor site derivation

Because GRIND variables are the energy product of MIF pairs of grid points, tracing back the original grid nodes around the molecules under study gives information about essential pharmacophoric groups. Maximum auto and cross-covariance (MACC2)^[18] energy products and grid-filtered MIFs are shown in Figure 5 for correlograms O–O, O–N1, and O–TIP obtained with the most active flavone **5** and flavone **12**, a weak inhibitor. The three most relevant variables with direct impact on biological activity (OO-9, ON-43, and OT-14) are indicated with an arrow in the MACC2 profile, and the pairs of points connected by a line in the filtered MIFs. As expected, these most relevant variables have higher values for the potent inhibitor, the 5,7,4'-trihydroxy-3'-methoxyflavone **5**, and lower values or even none for flavone **12**. This is explained by the lack of hydrogen bond donors and protruding groups in compound **12**. Contrarily, appropriate substitutions in compound **5** optimally

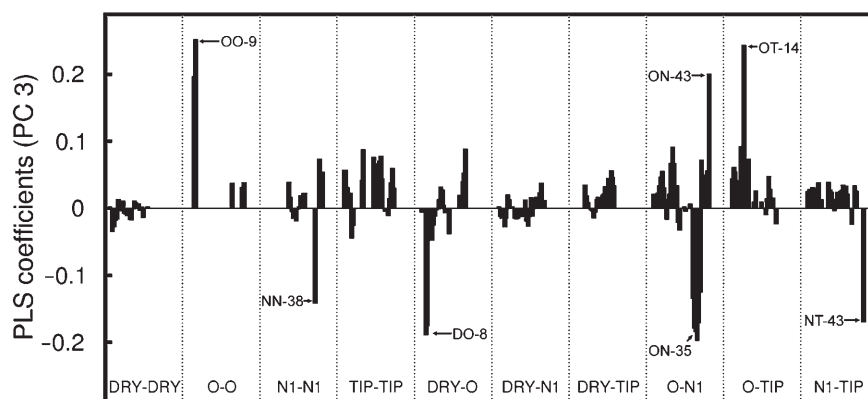


Figure 4. PLS coefficients plot of the GRIND variables used in the model. Different correlograms are separated by dotted lines and the pairs of probes are defined at the bottom. The most relevant variables are indicated by the variable number with an arrow.

Table 7. Relevant variables with high impact on the GRIND PLS model with three latent variables.

Probe pair	Variable number	Impact	Coeff	Interpretation ^[a]
O–O	OO-9	Direct	0.25	Interaction of the groups at C5 and C7 (ring A) with the probe O
N1–N1	NN-38	Inverse	–0.14	Interaction of the 4-oxo group (ring C) and the group at C4' (ring B) with the probe N1
DRY–O	DO-8	Inverse	–0.19	Interaction of the group at C3' (ring B) with the probe O and hydrophobic properties of ring C
O–N1	ON-43	Direct	0.20	Interaction of the group at C7 (ring A) with the probe O and the group at C4' (ring B) with the probe N1
O–N1	ON-35	Inverse	–0.20	Interaction of the group at C3' and C4' (ring B) with the probe O and the 4-oxo group (ring C) with the probe N1
O–TIP	OT-14	Direct	0.24	Interaction of the group at C4' (ring B) with the probe O and shape of the group at C3' (ring B)
N1–TIP	NT-43	Inverse	–0.17	Interaction of the group at C4' (ring B) with the probe N1 and shape of ring A

[a] Ring identification and carbon numbering are defined in Table 1.

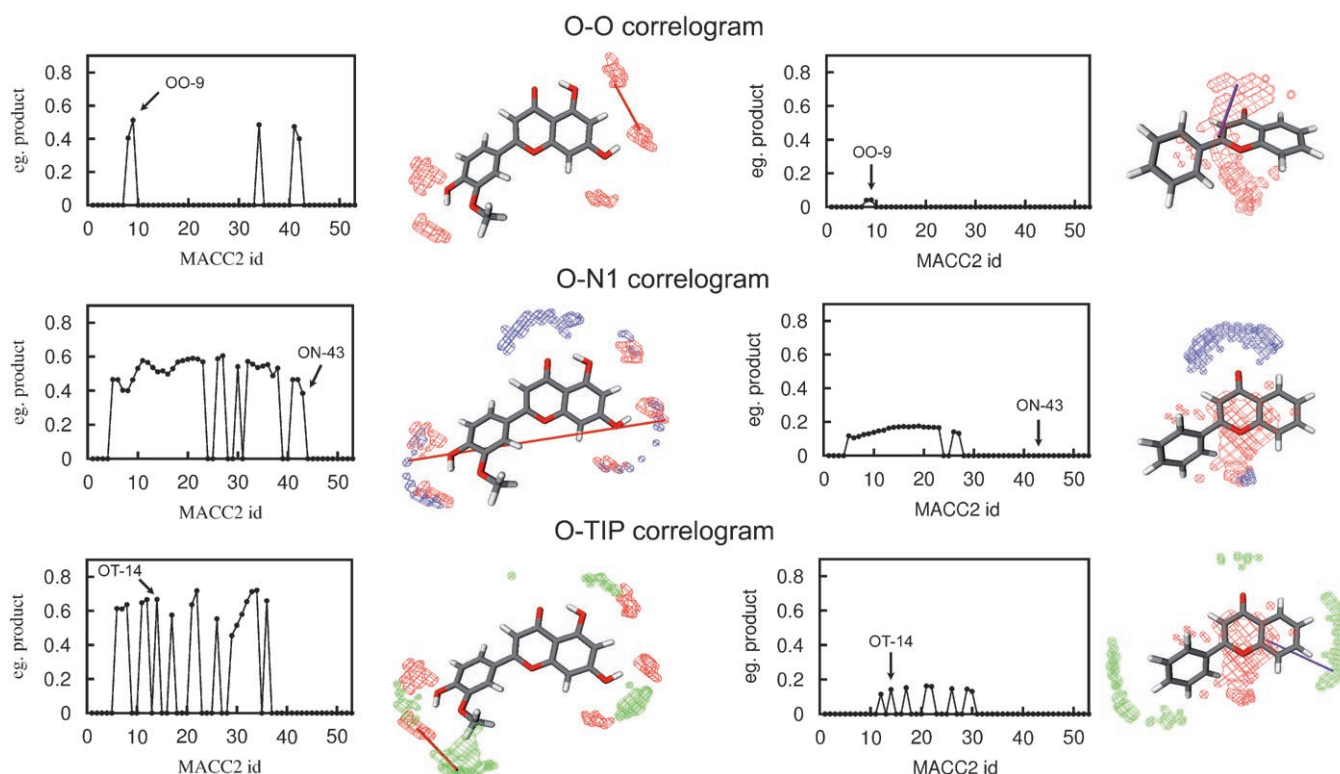


Figure 5. MACC2 interaction energy product and graphical representation of the three most relevant GRIND variables with direct impact obtained for the most active flavone, 5,7,4'-trihydroxy-3'-methoxyflavone (compound **5**, left), and flavone **12** (right), a moderate aromatase inhibitor. Arrows point to the most important variables OO-9, ON-43, and OT-14 in the MACC2 correlograms, and respectively correspond to the pairs of grid nodes O–O, O–N1, and O–TIP linked in the filtered MIFs shown. Molecules are shown (O, red; C, gray; H, white) at the minimum-energy conformation. The filtered MIFs are shown in colored contour (O probe, red; N1 probe, blue; shape probe, green).

place chemical functions that are able to increase the interaction energy products and therefore these are recognized by the model as important structural features.

On the other hand, variables with a high negative impact on the biological activity are increased for the weaker inhibitors found (for example, molecules **9** and **10**). These refer mostly to pairs of points between ring B (interaction of hydrogen bond donors/acceptors with groups at positions C3' and C4') and ring C (Table 7), which suggests that in the absence of hydroxylations on ring A (positions C5 and C7), the structure–activity relationships of ring B are different, favoring more hydrophobic groups (IC_{50} : **9** > **10** > **11**).

GRID probes use the GRID force field to calculate the interaction energy, which is the sum of several binding contributions ($E_{GRID} = E_{LJ} + E_{EL} + E_{HB} + E_{entropy}$), namely van der Waals interactions (E_{LJ}) given by the Lennard–Jones potential, electrostatic interactions (E_{EL}) given by the coulomb potential, and hydrogen bond interactions given by a distance-, angle-, and charge-dependent function (E_{HB}). The entropy decrease due to the freezing of degrees of freedom during the binding to the receptor is also considered with an entropic energy function ($E_{entropy}$). Therefore, GRIND descriptors are based on the idea of a virtual receptor site with which the inhibitors interact. When the three selected variables with direct impact on the biological activity (OO-9, ON-43, and OT-14) were applied to the most active flavone **5**, a virtual receptor site was obtained (Figure 6).

Two hydrogen bond acceptor regions are expected to be located in the vicinity of C5 and C7 of the flavone ring A. Hydrogen bonding with these hydroxy groups explains the increased activities of flavones **1**, **3**, and **5**. Position C4' on ring B is expected to lie close to a hydrogen bond donor/acceptor region. Furthermore, favorable steric and van der Waals interactions (good shape complementarity at the active site cavity) between the substitution at C3' (ring B) and the enzyme are also responsible for an increase in activity. In this virtual receptor model, hydrophobicity is important because of the apolar nature of the flavonoid scaffold, and one hydrogen bond donor region was also considered, close to the 4-oxo group.

Comparison with an aromatase homology model

Knowledge of the 3D structure of a target protein is important to understand its mechanism of action and inhibition. Unique information can be extracted from the atomic detail of the binding site and used in the identification of key interactions between the protein and the ligand.

Several cytochrome P450 enzymes have already been crystallized. Despite the low sequence identity (10–30%) between members of different P450 families, they all share the same overall topology and secondary structure elements.^[26] Furthermore, the internal heme binding pocket cavity shows a high percentage of sequence identity. Homology models have

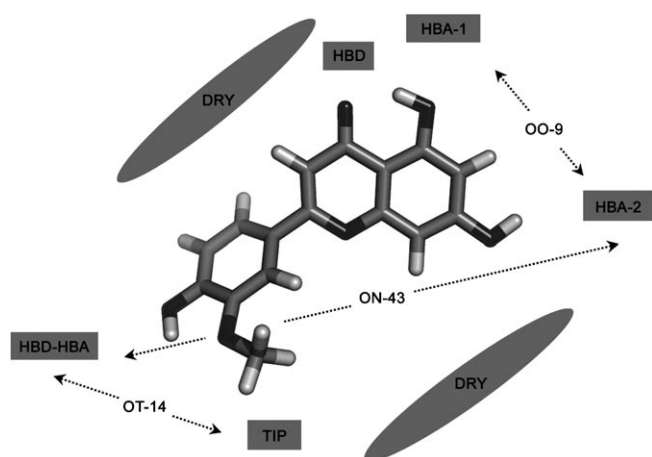


Figure 6. Virtual receptor site obtained when the most relevant descriptors with direct impact on the biological activity are coded into pharmacophoric regions around the strongest flavone tested, **5**. Two hydrogen bond acceptor regions (HBA-1 and HBA-2), a mixed hydrogen bond donor/acceptor region (HBD-HBA) and one cavity edge (TIP) were found. One additional large hydrophobic region (DRY) is responsible for the anchor of the flavone scaffold, and one hydrogen bond donor (HBD) stabilizes the binding to the active site through hydrogen bonding with the 4-oxo group. The flavone is shown (O, black; C, gray; H, white) at the minimum-energy conformation. The pharmacophoric regions are linked by an arrow showing the GRIND variable number.

therefore been proposed for a variety of cytochrome P450 binding sites, including aromatase, based on the known structures. One of these models was proposed by Favia et al.^[20] (PDB code 1TQA from the theoretical model section) using the human cytochrome P450 2C9 as template.^[27] This model was used in the current study to predict possible binding modes for the compounds under evaluation by using MIFs calculated within the active site with appropriate probes, and the results obtained were compared with the 3D QSAR virtual receptor

site. DRY (hydrophobic) and OH (aromatic hydroxy group) probes were initially computed. As expected, the active site cavity of the aromatase model is extremely hydrophobic owing to a large number of residues with apolar side chains, such as Leu120, Ile133, Phe134, Leu227, Ile305, Val373, Met374, and Leu477, shown in Figure 7. The probe DRY interacts with these residues, defining a hydrophobic contour represented in Figure 7 that is large enough to accommodate the inhibitors in study and place them close to the heme group, thus competing with the substrate for the active site of the enzyme. The OH probe accounts for hydrophilic portions of the active site that are prone to form hydrogen bonds with the ligands. This probe reveals several hydrophilic regions in the active site of aromatase, three of which contour the hydrophobic plane obtained with the DRY probe with a spatial disposition similar to the 3D QSAR model previously described (Figure 7). Contributions to these hydrophilic regions arise from backbone peptide bonds: the Ser118 amide nitrogen atom and the Phe116 and Met374 carbonyl oxygen atoms (zone 1); the Tyr244 hydroxy group, the Asp222 carboxylate, and the Ala226 carbonyl oxygen atom (zone 2); and finally, the Thr310 hydroxy group, the Ile305 and the Ala306 carbonyl oxygen atoms, and the heme iron (zone 3).

Docking calculations

The similarities between the 3D QSAR virtual receptor site and the hydrophobic/hydrophilic regions inside the active site prompted us to search for possible binding modes for the aromatase inhibitors studied. GLUE^[21] is a docking software that uses the GRID MIFs to generate binding poses for the ligands, and the GRID force field to calculate the respective binding energy. The aromatase binding cavity was initially explored with eight GRID probes that represent the most important non-bonded binding interactions: H₂O recognizes hydrophilic

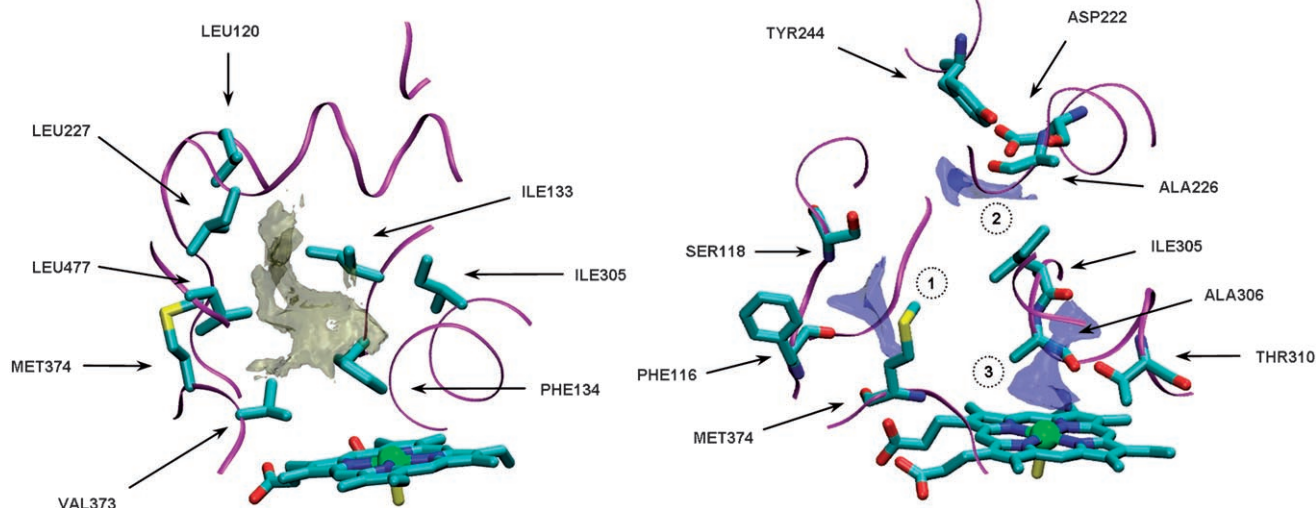


Figure 7. MIF isosurfaces obtained with the probe DRY (left) and OH (right) on the homology model of the aromatase active site.^[20] The protein is shown in ribbons with the residues responsible for the interactions labeled with residue name and number; the heme group appears at the bottom (C, cyan; O, red; N, blue; S, yellow; Fe, green). The DRY (hydrophobic) MIF is represented in brown at an energy level of $-0.4 \text{ kcal mol}^{-1}$, and the OH (aromatic hydroxy group) MIF is represented in blue at an energy level of $-7.0 \text{ kcal mol}^{-1}$. The three hydrophilic zones are numbered 1–3.

regions and evaluates competition with water molecules; DRY recognizes hydrophobic regions; H evaluates protein shape; N1 hydrogen bond donor to evaluate hydrogen bond acceptor groups; O hydrogen bond acceptor to evaluate hydrogen bond donor groups; O⁻ partially charged carboxylate group to evaluate positively charged regions; N⁺ positively charged nitrogen to evaluate negatively charged regions; OH donor/acceptor probe to evaluate mixed groups (see Experimental Section for further details). The MIFs local energy minima calculated with the probes DRY, N1, O, O⁻, N⁺, and OH represent favorable interaction points between the probes and the protein and correspond to pharmacophoric locations where similar groups of a ligand might be placed. A set of these site points was recorded for each probe and together combined in a large number of pharmacophoric quadruplets. The large number of possible pharmacophoric alignments in the binding site was then confronted with the inhibitors under study, and only similar pharmacophoric quadruplets found in the molecules were accepted. The conformational flexibility of the ligands was taken into account as described in the Experimental Section.

The estimated binding energy (EBE) of each best pose in the active site is shown in Tables 1–5. These results show that the docking program was able to discriminate between active and inactive compounds with good qualitative correlation between EBE and experimental activity, scoring coumarins **20**, **21**, **22**, **23**, and malvidin-3-*O*-glucoside **19**, weak aromatase inhibitors, with less-favorable binding energies.

Best-docked poses (that is, minimal EBE found) were further analyzed. Active site bound conformations found for flavone **5** (Figure 8), malvidin-3-*O*-glucoside **19** (Supporting Information figure 1), oleuropein **25** (Figure 9), coumarin **20** (Supporting In-

formation figure 2), and *trans*-resveratrol **24** (Figure 9) are represented.

The most active flavone **5** fits the active site cavity with the apolar scaffold in the hydrophobic core (Figure 8, left side). This apolar region extends across the whole skeleton of the molecule and is relatively planar. A bulky residue (Phe134) close to ring B of the flavone imposes planarity for this type of aromatase inhibitor (Figure 8, right side). In agreement with the virtual receptor site, the three hydrophilic zones are occupied by the 4-oxo and hydroxy groups. The C4' hydroxy group lies close to the hydrophilic zone 3 on the vicinity of the heme group, establishing a hydrogen bond with the hydroxy group of Thr310. The 4-oxo and the C5 hydroxy functions are located close to zone 1. The C5 hydroxy accepts a hydrogen bond from the amide N atom of Ser118, and the C7 hydroxy group lies close to zone 2, hydrogen bonding Ala226. Also consistent with the 3D QSAR model, the presence of a C3' methoxy group allows better shape complementary between the flavonoid and Val373, filling a small hydrophobic cavity. Other flavones and flavanones share similar docking conformations. Particularly interesting is the anti-aromatase activity ($IC_{50} = 8.2 \mu\text{M}$) of 7-hydroxyflavone **6**, the least functionalized strong inhibitor. Despite having only one hydroxy group, this flavone places one hydrogen bond acceptor, the 4-oxo group, in one hydrophilic region and one hydrogen bond donor/acceptor, the C7 OH group, in another, and therefore retains most of the aromatase inhibitory potency.

Cyanidin **18** and malvidin-3-*O*-glucoside **19** form quinoidal bases at pH 7.5 due to double bond conjugation (Supporting Information figure 3) and hydroxy groups at positions C5, C7, and C4'. Therefore, the three different resonance forms were conformationally analyzed and docked. Docking results with malvidin-3-*O*-glucoside **19** agree with the experimental data,

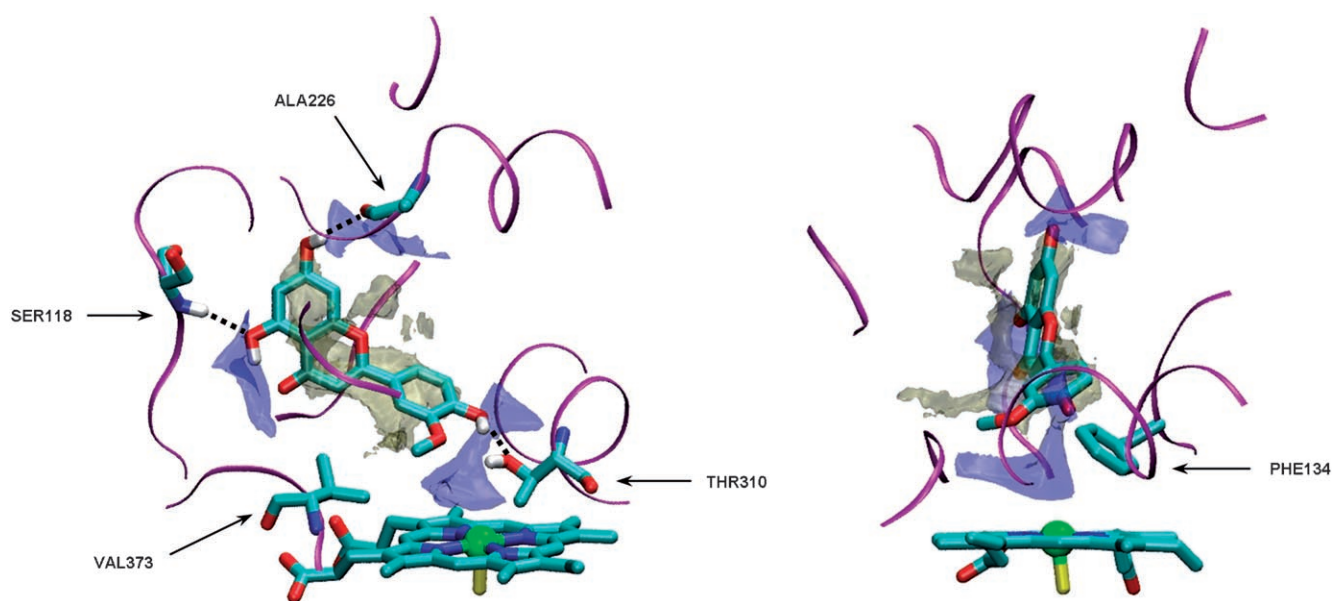


Figure 8. Docked pose of 5,7,4'-trihydroxy-3'-methoxyflavone **5**. MIF isosurfaces for the probes DRY (hydrophobic) and OH (aromatic hydroxy group) are shown at an energy level of $-0.4 \text{ kcal mol}^{-1}$ (brown) and $-7.0 \text{ kcal mol}^{-1}$ (blue), respectively. Val373 is displayed on the left, and dashed lines are drawn between the molecule and the residues involved in hydrogen bonding. Phe134 is displayed on the right.

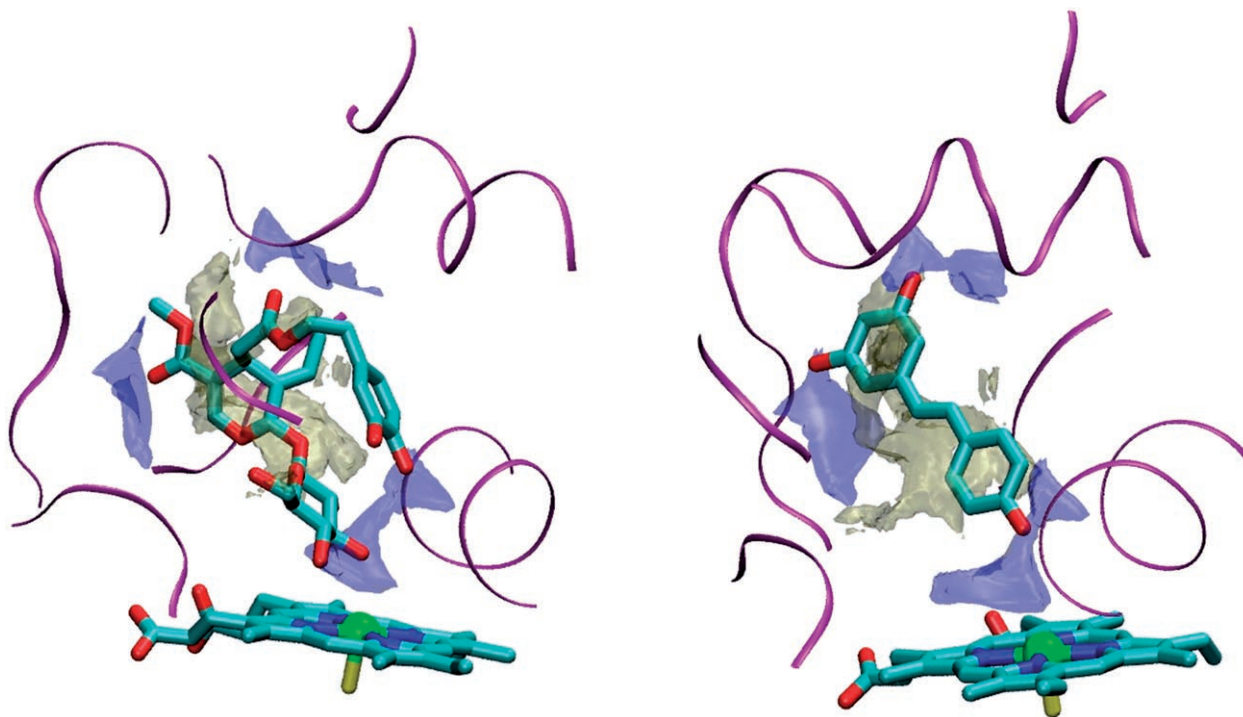


Figure 9. Docked pose of oleuropein (compound **25**, left) and *trans*-resveratrol (compound **24**, right). MIF isosurfaces for the probes DRY (hydrophobic) and OH (aromatic hydroxy group) are shown at an energy level of $-0.4 \text{ kcal mol}^{-1}$ (brown) and $-7.0 \text{ kcal mol}^{-1}$ (blue), respectively.

as the EBE is higher than the most active compounds. Detailed analysis of the EBE reveals a large steric penalty due to close contacts between the molecule and active site residues, namely Phe134, Val373, Met374, and Leu477 (Supporting Information figure 1). These contacts are related to the decreased flexibility of this molecule, and the presence of two bulky groups, a dimethoxy hydroxyphenyl group and a glucose moiety, linked at adjacent positions. This Y-shaped rigid structure does not fit well in the active core of the enzyme.

Despite the large molecular volume of oleuropein **25**, several rotatable bonds connecting the phenol moiety to the heterocyclic ring are determinant to a good fit within the active site of aromatase. The docked molecule reveals that the glucose lies at the bottom of hydrophilic zone 3, close to the heme group, which allows anchoring of the phenol moiety at the top of the same area (Figure 9, left side). The two remaining hydrophilic areas are occupied by carboxy groups. Therefore, superimposition of oleuropein **25** with 5,7,4'-trihydroxy-3'-methoxyflavone **5** in the active site conformation predicted with GLUE^[21] makes it clear that although apparently different, these two molecules share structural and pharmacophoric features (Figure 10). The heterocyclic carboxy group lies in an intermediate position between the flavone 4-oxo and the C5 hydroxy groups. The carboxy group in the phenol heterocyclic linker lies close to the flavone C7 hydroxy group, and both the sugar and the phenol point in the same direction of the flavone ring B. The length of both molecules is very similar, and the hydrophobic scaffolds are well overlapped. Different binding modes of the flexible molecule **25** could also be hypothesized in which a dynamic fitting of the substrate and the re-

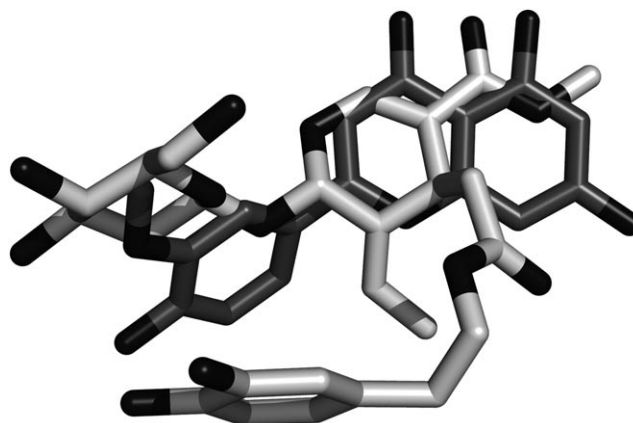


Figure 10. Predicted superimposition of the 5,7,4'-trihydroxy-3'-methoxyflavone (compound **5**, dark-gray C atoms) and the oleuropein (compound **25**, white C atoms) in the active site of the aromatase model. Oxygen atoms are represented in black.

ceptor lead to conformational changes that eventually determine a complex in which the sugar moiety is exposed to the solvent. P450 enzymes have actually been shown to undergo conformational changes upon substrate binding, characterized by the opening of solvent access channels.^[28,29] Moving the hydrophilic glucose moiety to a different orientation might determine a different energetically favorable binding configuration represented by an open conformation with a solvent-exposed sugar moiety. This hypothesis cannot be ruled out at this stage of calculations. It should be confirmed, however, with the use of molecular dynamics simulations allowing the full flexibility

of the whole system, which were outside of the scopes of this paper.

6,7-Dihydroxycoumarin **20** (Supporting Information figure 2) aligns well in the same hydrophobic region as the other polyphenolic compounds. However, this short scaffold fails to put the hydroxy groups in the vicinity of the hydrophilic areas. Together with the decreased hydrophobicity, this explains the experimentally observed low anti-aromatase potency and the higher EBE.

trans-Resveratrol **24**, a strong aromatase inhibitor, had an unfavorable EBE relative to the other active molecules. Because this molecule docks within the active site, perfectly aligned with the hydrophobic region, with three hydroxy groups oriented toward polar regions (Figure 9, right side), the EBE value might be related to the lower hydrophobicity of the stilbene scaffold and the absence of the 4-oxo group. However, this is well balanced by a perfect planar conformation and a good disposition of the phenyl and hydroxyl groups, increasing the strength of the interaction. Superimposition of resveratrol **24** and the strongest flavone inhibitor **5** in the active site docked conformation shows a good pharmacophoric alignment between these two molecules (Figure 11). This simple molecule therefore gives important information about the molecular determinants for the binding of polyphenolic inhibitors to the active site of aromatase.

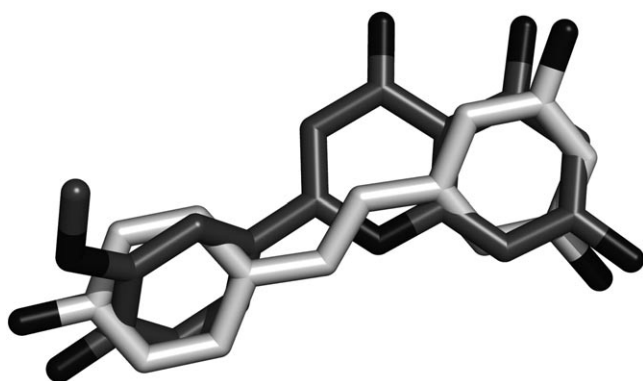


Figure 11. Predicted superimposition of the 5,7,4'-trihydroxy-3'-methoxyflavone (compound **5**, dark-gray C atoms) and the *trans*-resveratrol (compound **24**, white C atoms) in the active site of the aromatase model. Oxygen atoms are represented in black.

In summary, the anti-aromatase activity of a set of natural polyphenolic compounds was rationalized through the combination of experimental data and molecular modeling. In the absence of an experimental 3D structure of the enzyme, a virtual receptor site was defined and validated. The binding site cavity of the enzyme is expected to be strongly hydrophobic. Three hydrophilic regions were predicted based on the models and linked to the presence of the hydroxy functions at C5, C7, and C4' found in the strongest inhibitors tested. The C3' position was shown to be tolerant of bulky substituents, protruding towards a pocket near the active site. Docking of the compounds into a 3D structure of aromatase based on homology modeling is positively consistent with the results reported

above. The docking experiments were able to distinguish between active and inactive molecules, and could define a common binding mode for all compounds in the study.

Conclusions

In the study reported herein, the anti-aromatase potency of a set of 25 polyphenols was tested. Seven strong inhibitors, compounds **1**, **3**, **5**, **6**, **13**, **14**, and **16**, with equal or better activity than the reference compound, AG, were identified. All these molecules are flavone or flavanone derivatives, which confirms the potential of natural polyphenols for lead optimization into strong aromatase inhibitors. Flavone hydroxylation at position C7 of ring A proved to be determinant for good anti-aromatase potency, as well as the double hydroxylation at C5 and C7. Further improvement by C4' hydroxylation and either C3' hydroxylation or methoxylation on ring B, as with compounds **3** and **5**, was observed. Flavanones gave results in good agreement with the rules described above and led to the identification of compound **14**, the best aromatase inhibitor studied.

The 3D QSAR study performed with the subset of flavones explained the variance of the data and showed a good predictive ability even with the use of a less conservative cross-validation procedure with random groups. The use of GRIND^[18] descriptors calculated with the GRID force field and extracted from relevant MIFs (hydrophobicity, hydrogen bonding, and molecular shape) was a valid strategy to characterize a virtual receptor site and gain further insight into the molecular environment surrounding polyphenol inhibitors on the active site of aromatase. Variables with a direct high impact on the inhibitory activity (Figure 4) defined a putative virtual receptor site (Figure 6): two hydrogen bond acceptor regions are expected to be close to C5 and C7 of flavone ring A, and a hydrogen bond donor/acceptor area close to C4' of ring B. C3' was found to be permissive to bulky substitutions, allowing good shape complementarity with the active site. A hydrophobic scaffold is common to all flavones, and a 4-oxo group should also be important for binding.

Analysis of an aromatase homology model with molecular interaction probes that account for hydrophobicity and hydrogen bonding agreed with the virtual receptor site and showed a large hydrophobic region within the active site surrounded by three hydrogen bond donor/acceptor regions (Figure 7). Docking calculations with the GLUE^[21] software aligned the inhibitors under study on the hydrophobic plane, with polar groups correctly oriented to the hydrophilic areas (Figure 8). Therefore, hydroxylations at C5, C7, and C4' strengthens the affinity for the active site and increases anti-aromatase potency. The binding cavity is large enough to accommodate most of the compounds studied without steric clashes, and it was found that Val373 might be responsible for a better shape complementarity between bulky substitutions at flavone C3' and aromatase.

The higher EBE calculated for coumarins agrees with their weak anti-aromatase potency, and the docked poses confirm that the benzopyrone hydrophobic scaffold is too short for a

strong interaction with the active site of the enzyme (Supporting Information figure 2). Extension of this scaffold with apolar groups might increase the hydrophobicity and lead to strong aromatase inhibitors, as shown recently.^[30]

The simple and conformationally restrained *trans*-resveratrol **24** was another strong aromatase inhibitor tested (IC_{50} = 12.8 μ M). Despite the absence of the 4-oxo group found in flavones and flavanones, the two aromatic rings and the three hydrogen bond donor/acceptor groups agree with the virtual receptor site and the MIFs calculated on the active site model (Figure 9, right side), leading to an increase in binding strength to aromatase.

Herein we found that oleuropein **25**, the major polyphenolic constituent of olive oil^[7] is a good aromatase inhibitor, with an IC_{50} value of 27 μ M. This strong inhibitory potency was rationalized based on a proper pharmacophoric superimposition with the best flavone inhibitor tested (compound **5**, Figure 10). A good fit was found within the hydrophilic/hydrophobic regions of the aromatase active site without steric clashes. The flexibility introduced by the ethyl acetate chain that connects the catechol to the glycosylated heterocyclic moiety is determinant for the adoption of an U-shaped conformation that fits well in the active site (Figure 9, left side). In contrast, the strong conformational barriers of malvidin-3-*O*-glucoside **19** do not allow a suitable conformation. Our docking protocol, with an extensive conformational search of flexible molecules and the use of adequate MIFs calculated with GLUE,^[21] is therefore a valid method to perform docking and molecular superimposition at the active site of a macromolecular target.

In conclusion, we have described a new methodology that combines biochemical evaluation, 3D QSAR with alignment-independent descriptors, and screening for energetically favorable binding sites for the rationalization of the anti-aromatase activity of a set of diverse natural polyphenolic compounds. Starting from experimentally determined inhibitory profiles, our approach provided novel insight into the chemical and steric environment within the active site of aromatase and on the essential determinants for binding of polyphenolic inhibitors. Based on this information, a virtual receptor site was built, and a possible common binding mode was proposed. This was further corroborated by the comparison with the results of docking experiments into a 3D homology model of aromatase. The good agreement between experimental and predicted activity, and between the properties of the virtual receptor model and the 3D homology model suggests that this approach could be extended to other cases for which the structure of the receptor is unknown. The use of the aromatase virtual receptor model to perform 3D database searching and molecular design of new strong aromatase inhibitors is under investigation.

Experimental Section

Materials and general methods. 7-Hydroxyflavone, 5,7,4'-trihydroxy-3'-methoxyflavone, 5,7,4'-trihydroxyflavone, 3',4'-dihydroxyflavone, 5,7,3'-trihydroxy-4'-methoxyflavone, flavanone, 7-hydroxyflavanone, 5,7-dihydroxyflavanone, 5,7,3',4'-tetrahydroxyflavanone,

5,7,4'-trihydroxy-3'-methoxyflavone, cyanidin chloride, resveratrol, and oleuropein were purchased from Extrasynthese (Genay, France). Flavone, 5,7-dihydroxyflavone, 7,3',4'-trihydroxyflavone, 5,7,3',4'-tetrahydroxyflavone, 7,3',4',5'-tetrahydroxyflavone, 4'-hydroxy-3'-methoxyflavone, 3',4'-dimethoxyflavone, 6,7-dimethoxycoumarin, 6,7-dihydroxycoumarin, 7-hydroxy-6-methoxycoumarin, and 6-hydroxy-7-methoxycoumarin were purchased from Indofine Chemical (Hillsborough, NJ, USA). Malvidin-3-*O*-glucoside was obtained from Polyphenols (Hanaveien, Norway). NADPH and DL-aminoglutethimide were purchased from Sigma-Aldrich (St. Louis, MO, USA). [1β - 3 H]Androstenedione (specific activity: 25.3 Ci mmol⁻¹) and the liquid scintillation cocktail Optiphase Hisafe 2 were purchased from PerkinElmer (Boston, MA, USA), and the radioactive samples were counted on a Packard Tri-Carb 2000 CA liquid scintillation analyzer.

Isolation of human placental microsomes. Human term placental microsomes were obtained by differential centrifugation according to the method described by Ryan,^[23] and were used as a source of aromatase. The microsomes were resuspended in buffer containing 0.1 M sodium phosphate, 0.25 M sucrose, 20% glycerol, and 0.5 mM dithiothreitol, pH 7.4, and stored in aliquots at -80°C until needed.

Concentration–response study. Aromatase activity was evaluated by quantifying the amount of $^3\text{H}_2\text{O}$ released upon enzymatic conversion of the tritiated substrate, [1β - 3 H]androstenedione, into estrone. This method was first described by Siiteri and Thompson.^[22] Briefly, enzymatic assays were performed at 37°C in a sodium phosphate buffer (67 mM, pH 7.5) at a final volume of 500 μ L, containing [1β - 3 H]androstenedione (6.6×10^5 dpm) and 270 μ M NADPH. The inhibitors studied were dissolved in DMSO and added to the assay in concentrations ranging from 316.2 to 640×10^3 nM. The amount of DMSO in the assay was always < 2%. The reactions were started with the addition of microsomal protein (30 μ g) and stopped after 20 min with the addition of chloroform (1 mL) and vortexing at 9000 rpm for 40 s. The tritiated water formed during the aromatization reaction and released into the incubation medium was quantified by liquid scintillation counting. Appropriate controls without inhibitor were performed to determine the maximum enzymatic activity to which the relative percentage of inhibition was determined. The assays were performed three times in triplicate, and the results were treated by nonlinear regression analysis using a sigmoidal concentration–response curve with variable slope.

Conformational search and geometry optimization. All compounds were constructed with building fragments from the standard libraries of MAESTRO^[31] v 5.1.016. Flexibility was taken into account by conformational search using the systematic unbounded multiple minimum (SUMM)^[32] routine implemented in MACROMODEL^[33] v 8.1 with the Merck molecular force field (MMFF)^[34] and the Polak–Ribiere conjugate gradient (PRCG) minimization method, with an energy convergence criterion of 0.05 kJ mol⁻¹. The generalized Born equation/surface area (GB/SA)^[35] continuum solvation model was used with parameters for water, with a dielectric constant (ϵ) of 78. A maximum of 30 000 conformations were generated and saved if within an energy window of 50 kJ mol⁻¹ over the global minimum. Similar structures were excluded based on heavy atom superimposition. All other settings were used as default.

3D QSAR using ALMOND. GRIND descriptors^[18] were calculated for the flavones studied with ALMOND^[21] v 3.3 using MIFs^[19] computed with the chemical probes DRY (hydrophobic probe), O (hydrogen bond acceptor interactions), and N1 (hydrogen bond donor inter-

actions), and the molecular shape probe TIP,^[24] with the grid spacing set to 0.5 Å. These descriptors encode geometrical relationships between pairs of non-bonded interactions; therefore, to represent only the highly favorable interactions, one initial filtering step was performed, and 100 low-energy interaction points, separated from each other as much as possible, were extracted from each MIF. To this step, the relative importance of the field values was set to 50, and all other settings were set to default. Finally, the filtered MIFs were encoded into GRIND variables through a maximum auto and cross-covariance (MACC) transformation performed with the width of the smoothing window set to 0.8. These variables represent the product of MIF energy node pairs separated by a certain distance around each molecule.

Statistical analysis. Partial least-squares regression (PLS) was used to derive the 3D QSAR model by using the statistical tools included in ALMOND^[21] v 3.3. No scaling was applied to the variables. The optimal dimensionality of the model was selected by cross-validation using either LOO or 3RG, recalculating the weights in both cases. The random groups validation was repeated 200 times to get a stable predictive correlation coefficient. The fractional factorial design (FFD),^[25] a variable selection methodology implemented in ALMOND^[21] v 3.3, allowed the removal of descriptors not correlated with activity, introducing only noise to the model. Uncertain variables were kept. The FFD was forced to “fold over”, and the combinations/variables ratio was set to 5 in order to better estimate the effect of each variable in the model predictive ability. The model noise level was evaluated by 20% of “dummy variables”.

Active site MIF calculations. GREATER, the graphical user interface for the GRID^[21] v 22a package of programs, was used to calculate MIFs with the probes DRY (hydrophobic) and OH (aromatic hydroxy group) on the active site of a homology model of aromatase (PDB entry 1TQA from the theoretical model section).^[20] The protein was considered rigid, and a grid box was built within the active site, including the heme group and all interacting residues. This resulted in a cubic box of 8000 Å³ (20 Å × 20 Å × 20 Å) centered at the aromatase binding cavity and containing the heme group at the bottom. The MIFs were calculated with the grid spacing set to 0.25 Å and the directive ALMD set to 1.

Binding mode prediction. GLUE^[21] v 1.0 is a docking program that uses the GRID MIFs to generate ligand poses within the active site of a macromolecular target and the GRID force field to calculate binding energies. The aromatase binding cavity was initially explored with eight GRID probes that mimic most of the chemical groups present in the ligands: H₂O (recognizes hydrophilic regions and evaluates competition with water molecules), DRY (recognizes hydrophobic regions), H (evaluates protein shape), N1 (hydrogen bond donor to evaluate hydrogen bond acceptor groups), O (hydrogen bond acceptor to evaluate hydrogen bond donor groups), O⁻ (partially charged carboxylate to evaluate positively charged regions), N⁺ (positively charged nitrogen to evaluate negatively charged regions), and O1 (donor/acceptor probe to evaluate mixed groups). The MIFs obtained from the interaction of these probes with the active site were calculated in the same cubic box as described previously. Local energy minima extracted from the DRY, N1, O, O⁻, N⁺ and O1 MIFs were combined into a large number of pharmacophoric quadruplets. For this step the distance tolerance was set to 2.5, and the cavity expansion was set to 5 in order to maximize the number of combinations saved, available to the next steps. This large number of possible alignments was then filtered through comparison with the inhibitors in study, with the steric tolerance set to 1. When the number of rotatable bonds was three or fewer, several conformers were generated. When greater, a more

refined conformational search was performed with MACROMODEL^[33] v 8.1, and conformations found within an energy window of 15 kJ mol⁻¹ from the global minimum were submitted to GLUE.^[21] Both *R* and *S* flavanone stereoisomers were considered, and all cyanidin **18** and malvine-3-*O*-glucoside **19** equilibrium quinoidal bases were studied independently (Supporting Information figure 3). The filtered pharmacophoric quadruplets were then docked within the active site of aromatase. Redundancy was avoided with a minimum RMS set to 2 Å, and a maximum of 2000 minimization iterations were allowed in order to relax the ligands in the binding pocket. The GLUE^[21] scoring function was then applied to the binding poses, with contributions for steric repulsion, hydrogen bonding, and hydrophobicity. Top-ranked orientations were extracted and further analyzed.

Acknowledgements

Thanks are due to Fundação para a Ciência e a Tecnologia (FCT) through POCTI (FEDER) for financial support. We also thank the Maternidade Dr. Daniel de Matos, Coimbra, Portugal for kindly providing human placental tissue. M.A.C.N. thanks Dr. M. Meli and Dr. G. M. S. de Mori for help on the computational methods, and FCT for a PhD grant (SFRH/BD/17624/2004).

Keywords: antitumor agents • aromatase inhibitors • natural products • polyphenols • structure–activity relationships

- [1] A. Jemal, R. Siegel, E. Ward, T. Murray, J. Xu, C. Smigal, M. J. Thun, *CA Cancer J. Clin.* **2006**, *56*, 106–130.
- [2] C. Andreatta, I. Smith, *Cancer Lett.* **2007**, *251*, 17–27.
- [3] H. Geisler, *J. Steroid Biochem. Mol. Biol.* **2003**, *86*, 245–253.
- [4] I. E. Smith, M. Dowsett, *N. Engl. J. Med.* **2003**, *348*, 2431–2442.
- [5] J. T. Kellis, Jr., L. E. Vickery, *Science* **1984**, *225*, 1032–1034.
- [6] L. Bravo, *Nutr. Rev.* **1998**, *56*, 317–333.
- [7] K. L. Tuck, P. J. Hayball, *J. Nutr. Biochem.* **2002**, *13*, 636–644.
- [8] C. A. Rice-Evans, N. J. Miller, G. Paganga, *Free Radical Biol. Med.* **1996**, *20*, 933–956.
- [9] B. Frei, J. V. Higdon, *J. Nutr.* **2003**, *133*, 3275S–3284S.
- [10] H. K. Hamdi, R. Castellon, *Biochem. Biophys. Res. Commun.* **2005**, *334*, 769–778.
- [11] B. B. Aggarwal, S. Shishodia, *Biochem. Pharmacol.* **2006**, *71*, 1397–1421.
- [12] V. Micol, N. Caturla, L. Perez-Fons, V. Mas, L. Perez, A. Estepa, *Antiviral Res.* **2005**, *66*, 129–136.
- [13] N. H. Aziz, S. E. Farag, L. A. Mousa, M. A. Abo-Zaid, *Microbios* **1998**, *93*, 43–54.
- [14] M. A. Carluccio, L. Siculella, M. A. Ancora, M. Massaro, E. Scoditti, C. Storelli, F. Visioli, A. Distanto, R. De Caterina, *Arterioscler. Thromb. Vasc. Biol.* **2003**, *23*, 622–629.
- [15] J. A. Vita, *Am. J. Clin. Nutr.* **2005**, *81*, 292S–297S.
- [16] S. Gobbi, A. Cavalli, A. Rampa, F. Belluti, L. Piazzi, A. Paluszczak, R. W. Hartmann, M. Recanatini, A. Bisi, *J. Med. Chem.* **2006**, *49*, 4777–4780.
- [17] Y. W. Kim, J. C. Hackett, R. W. Brueggemeier, *J. Med. Chem.* **2004**, *47*, 4032–4040.
- [18] M. Pastor, G. Cruciani, I. McLay, S. Pickett, S. Clementi, *J. Med. Chem.* **2000**, *43*, 3233–3243.
- [19] P. J. Goodford, *J. Med. Chem.* **1985**, *28*, 849–857.
- [20] A. D. Favia, A. Cavalli, M. Masetti, A. Carotti, M. Recanatini, *Proteins: Struct. Funct. Genet.* **2006**, *62*, 1074–1087.
- [21] Molecular Discovery Ltd., <http://www.moldiscovery.com>.
- [22] P. K. Siiteri, E. A. Thompson, *J. Steroid Biochem.* **1975**, *6*, 317–322.
- [23] K. J. Ryan, *J. Biol. Chem.* **1959**, *234*, 268–272.
- [24] F. Fontaine, M. Pastor, F. Sanz, *J. Med. Chem.* **2004**, *47*, 2805–2815.
- [25] M. Baroni, G. Costantino, G. Cruciani, D. Riganelli, R. Valigi, S. Clementi, *Quant. Struct.-Act. Relat.* **1993**, *12*, 9–20.

- [26] C. A. Hasemann, R. G. Kurumbail, S. S. Boddupalli, J. A. Peterson, J. Deisenhofer, *Structure* **1995**, *3*, 41–62.
- [27] P. A. Williams, J. Cosme, A. Ward, H. C. Angove, V. D. Matak, H. Jhoti, *Nature* **2003**, *424*, 464–468.
- [28] H. Li, T. L. Poulos, *Biochimie* **1996**, *78*, 695–699.
- [29] A. R. Dunn, I. J. Dmochowski, A. M. Bilwes, H. B. Gray, B. R. Crane, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12420–12425.
- [30] S. U. Chen, M. Cho, K. Karlsberg, D. J. Zhou, Y. C. Yuan, *J. Biol. Chem.* **2004**, *279*, 48071–48078.
- [31] Schrödinger LLC, <http://www.schrodinger.com>.
- [32] J. M. Goodman, W. C. Still, *J. Comput. Chem.* **1991**, *12*, 1110–1117.
- [33] F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson, W. C. Still, *J. Comput. Chem.* **1990**, *11*, 440–467.
- [34] T. A. Halgren, *J. Comput. Chem.* **1996**, *17*, 490–519.
- [35] W. C. Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.

Received: June 21, 2007

Revised: September 10, 2007

Published online on October 1, 2007