Fast Three Dimensional Pharmacophore Virtual Screening of New Potent Non-Steroid Aromatase Inhibitors

Marco A. C. Neves,[†] Teresa C. P. Dinis,[‡] Giorgio Colombo,*,[§] and M. Luisa Sá e Melo*,[†]

Centro de Estudos Farmacêuticos, Laboratório de Química Farmacêutica, Faculdade de Farmácia, Universidade de Coimbra, 3000-295, Coimbra, Portugal, Centro de Neurociências, Laboratório de Bioquímica, Faculdade de Farmácia, Universidade de Coimbra, 3000-295, Coimbra, Portugal, and Istituto di Chimica del Riconoscimento Molecolare, CNR, 20131, Milano, Italy

Suppression of estrogen biosynthesis by aromatase inhibition is an effective approach for the treatment of hormone sensitive breast cancer. Third generation non-steroid aromatase inhibitors have shown important benefits in recent clinical trials with postmenopausal women. In this study we have developed a new ligand-based strategy combining important pharmacophoric and structural features according to the postulated aromatase binding mode, useful for the virtual screening of new potent non-steroid inhibitors. A small subset of promising drug candidates was identified from the large NCI database, and their antiaromatase activity was assessed on an in vitro biochemical assay with aromatase extracted from human term placenta. New potent aromatase inhibitors were discovered to be active in the low nanomolar range, and a common binding mode was proposed. These results confirm the potential of our methodology for a fast in silico high-throughput screening of potent non-steroid aromatase inhibitors.

Introduction

Aromatase, a member of the cytochrome P450 superfamily that catalyzes the final and key step of C19 steroids conversion to estrogens,[1] is a very attractive target for the endocrine treatment of breast cancer.[2] In this sense, attention has been focused on the discovery of aromatase inhibitors able to decrease the circulating levels of estrogens and therefore control the progression of hormone sensitive breast cancer.[3]

P450 enzymes are monooxygenases that catalyze a large number of reactions involved in drug metabolism and synthesis of steroids and lipids.[4] There are more than 7700 different P450 sequences identified to date.[5] The design of potent and selective inhibitors to a single member of this superfamily represents a very challenging task. In spite of these hurdles, selective aromatase inhibitors have been discovered, such as fadrozole and formestane, a second generation of aromatase inhibitors, or exemestane, letrozole, anastrozole, and vorozole, a third generation of aromatase inhibitors.[6] On the basis of their inhibition mechanisms, these molecules are divided into two classes: (1) irreversible steroid inactivators such as formestane or exemestane that bind covalently to the enzyme and (2) non-steroid inhibitors such as letrozole or anastrozole that act as reversible competitive inhibitors.

Improvements in the rational design of new aromatase inhibitors could be obtained once a high resolution 3D structure of the enzyme becomes available. However, to date, no X-ray structure of aromatase has been published in the Protein Data Bank, despite recent progress in the 3D determination of mammalian and human cytochrome P450 enzymes.[7−10] On this basis, several homology models for aromatase[11−13] have been

built and proved to be valuable in understanding the binding determinants of several classes of inhibitors.[14−16] The low sequence identity between members of different P450 families may, however, limit their potential in structure-based virtual screenings.

An alternative strategy for the rational design of new aromatase inhibitors relies on ligand-based virtual screening that can be developed on the basis of the vast information on strong aromatase inhibitors published in the literature. Pharmacophore-based approaches have been used for both steroid and non-steroid aromatase inhibitors, highlighting important physicochemical features and guiding the design of new potent compounds.[17−20] However, to date, only a few attempts to perform virtual screening based on such models have been performed. Langer and co-workers, for instance, have carried out a mixed qualitative and quantitative approach for non-steroid aromatase inhibitors, using the HipHop and HypoRefine algorithms of the Catalyst software.[21] The final pharmacophore model proposed by this group consisted of one aromatic ring feature, one hydrophobic group, two hydrogen bond acceptors, and one excluded volume sphere. Two possible antiaromatase candidates were identified on the basis of this approach, one of the compounds being able to decrease in 68% the activity of the enzyme at 36 μM and the other in 15% at the same concentration.

In this study we have used known non-steroid aromatase inhibitors to derive and validate a new pharmacophore model for fast in silico screening of antiaromatase compounds. The model combined information about the common pharmacophore features of azole non-steroid inhibitors, information about their binding to the aromatase active site, steric restrictions and a "druglikeness" filter. The large NCI[a] compound database was screened using this methodology, and new potent aromatase inhibitors were identified. The new compounds have in vitro

---

* To whom correspondence should be addressed. For G.C.: phone, (+39)0228500031; fax, (+39)0228901239; e-mail, g.colombo@icrm.cnr.it. For M.L.S.M.: phone, (+351)239859990; fax, (+351)239827126; e-mail, samelo@ci.uc.pt.

[†] Laboratório de Química Farmacêutica, Faculdade de Farmácia, Universidade de Coimbra.

[‡] Laboratório de Bioquímica, Faculdade de Farmácia, Universidade de Coimbra.

[§] Istituto di Chimica del Riconoscimento Molecolare.

[a] Abbreviations: NCI, National Cancer Institute; HYD, hydrophobic group; HBA, hydrogen bond acceptor; HYP, pharmacophoric hypothesis; C*x*-N-HBA, hydrogen bond acceptor from a *x*-membered nitrogen-containing heterocyclic ring; VS, virtual screening; PSA, polar surface area.
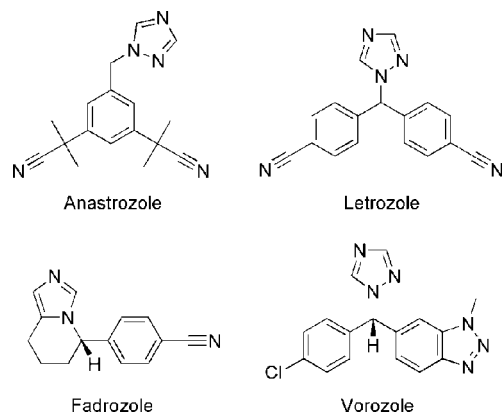
Figure 1. Training set of azole non-steroid aromatase inhibitors used for the common-features (HipHop) pharmacophore model generation.
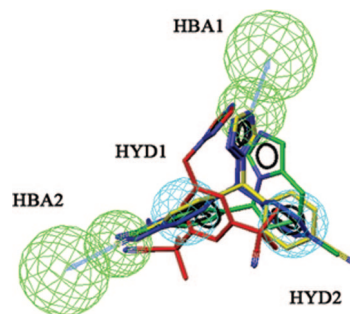

Figure 2. Common-features (Catalyst/HipHop) pharmacophore model of azole non-steroid aromatase inhibitors. The HYP2 pharmacophoric query had four features: two hydrogen bond acceptors (HBA1 and HBA2, green) and two hydrophobic groups (HYD1 and HYD2, cyan). The training set aromatase inhibitors (anastrozole, red; letrozole, yellow; fadrozole, green; vorozole, blue) are represented as the best fit alignment to the model.

activity comparable to that of second and third generation antiaromatase agents. Furthermore, the most active molecules were docked on the active site of an aromatase homology model[12] and these results were compared with the pharmacophore model.

Results and Discussion

3D-Pharmacophore Generation. Several non-steroid aromatase inhibitors have been used for the treatment of breast cancer in postmenopausal women, with clinical efficacy.[2] Having progressed through a preclinical development and clinical trials, these compounds are expected to include information about strong antiaromatase potency and selectivity. In this study, we have built a training set with low energy conformations of second and third generation aromatase inhibitors, namely, anastrozole, letrozole, fadrozole, and vorozole (Figure 1) and used these molecules to generate a common-features pharmacophore model with the HipHop[22] algorithm of the Catalyst software.[23] Briefly, the program evaluates chemical features common to a set of active compounds and generates hypotheses for their activity. These hypotheses are spatial dispositions of pharmacophoric points providing the compounds' relative alignment in the binding site of the enzyme. Each point accounts for an important chemical feature such as hydrogen bond donors/acceptors, hydrophobic groups, negative/positive ionizable groups, and aromatic groups. Basic physicochemical features of known antiaromatase compounds include a high degree of hydrophobicity and the potential to establish hydrogen bonds as acceptors. These properties are related to the highly apolar binding pocket of the enzyme which is expected to have at least one hydrogen bond donor residue.[24] Therefore, hydrophobic (HYD) and hydrogen bond acceptor (HBA) pharmacophoric features were used in this study. The program found six possible alignment solutions displaying three or four pharmacophoric points. The first two top ranked solutions, HYP1 (rank = 31.9) and HYP2 (rank = 29.8), had four features, two hydrophobic groups, and two hydrogen bond acceptors. Analysis of the training set molecules aligned to these hypotheses revealed overlapping imidazole and triazole groups in HYP2 but not in HYP1, with one nitrogen atom acting as hydrogen bond acceptor (HBA1, Figure 2). This feature of HYP2 reflects the binding mechanism found for this type of molecule, which entails binding through heterocyclic aromatic coordination to the heme iron of the P450 active site.[6] Therefore, HYP2 was chosen for the subsequent steps. The second hydrogen bond acceptor feature (HBA2, Figure 2) matches either

a cyano group (anastrozole, letrozole, and fadrozole) or the N2 atom of a methylbenzotriazole (vorozole), and the hydrophobic features superimpose with phenyl rings of all the molecules (HYD1, Figure 2), the methyl of the anastrozole, or the piperidine moiety of the fadrozole (HYD2, Figure 2).

Pharmacophore models with a small number of points tend to be unspecific and return a large number of hits when used to screen large databases. In particular, hydrophobic and hydrogen bond acceptor features match a large number of chemical groups, increasing the number of false positives. Therefore, in an attempt to increase the specificity of HYP2, the HBA1 feature was substituted with new fragment/function pharmacophoric features: the C5-N-HBA and the C6-N-HBA features (Figure 3). C5-N-HBA matches hydrogen bond acceptors only if present in certain types of five-membered nitrogen-containing heterocyclic rings (imidazole, 1,2,4-triazole, 1,2,3-triazole, and tetrazole), whereas C6-N-HBA matches acceptors from six-membered nitrogen-containing heterocyclic rings (pyridine, pyridazine, pyrimidine, and pyrazine).

Cytochrome P450 enzymes have an inner binding pocket with well defined volume accessible from the outside by several channels.[25] Therefore, besides the physicochemical properties of the inhibitor, molecular shape is expected to be extremely important to the access and fit within the active site of the aromatase. Most aromatase inhibitors used to build the common-features model have a similar shape that is expected to be complementary to the volume of the aromatase active site. The shape of letrozole, a quite rigid third generation inhibitor with high degree of symmetry, was therefore converted into a set of inclusion volumes and combined with the pharmacophoric query (HYP2-HBA-Shape).

Therefore, the final pharmacophore model combined pharmacophoric information (two hydrogen bond acceptors and two hydrophobic groups) with fragment information (nitrogen-containing aromatic heterocycles) and steric restrictions. All aromatase inhibitors used on the training set matched the final hypothesis with best fit values ranging from 2.98 to the maximum value of 4.00 (letrozole = 4.00, vorozole = 3.46, anastrozole = 3.40, and fadrozole = 2.98).

Model Validation. To validate the new pharmacophore model, a test set with 82 non-steroid aromatase inhibitors collected from the literature was built.[17,26-31] Each molecule contained a five- or six-membered aromatic heterocyclic ring able to coordinate with the aromatase heme iron. The initial pharmacophore model (HYP2) returned 58 molecules (71% of the database, Table 1). The pharmacophore modified to match
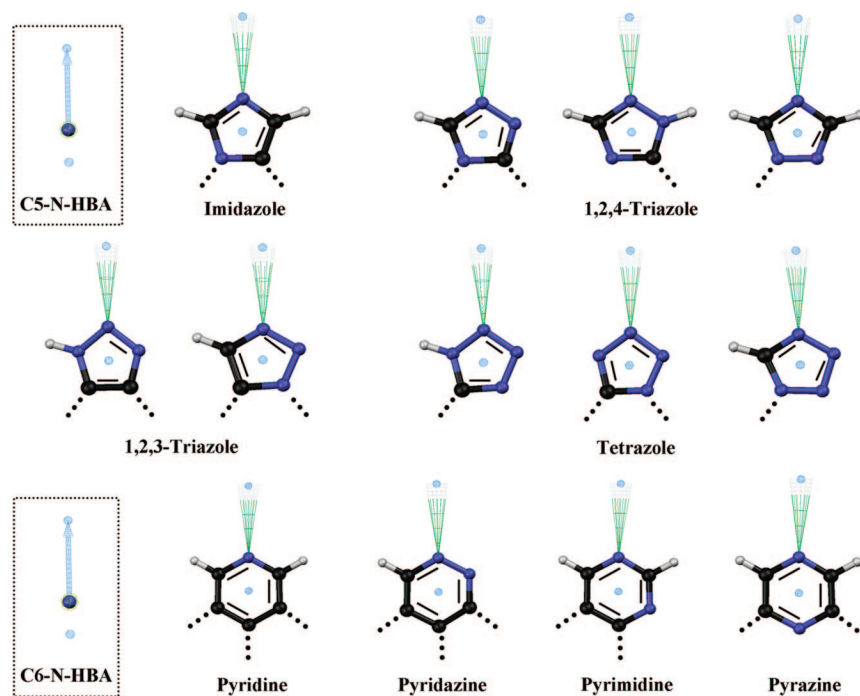
Figure 3. C5-N-HBA and C6-N-HBA fragment/function pharmacophore features. Imidazole, 1,2,4-triazole, 1,2,3-triazole, and tetrazole rings were combined into a single fragment feature using the centroids of the five membered ring moieties. Pyridine, pyridazine, pyrimidine, and pyrazine rings were combined into another fragment feature using the centroids of the six membered ring moieties. Anchoring points to the rest of the molecule are represented with doted lines. The nitrogen atom distal to these points defined a hydrogen bond acceptor and the projected point, the position in space from which the participating hydrogen will extend. These two positions are connected by a vector that indicates the direction from the heavy atom to the projected point of the hydrogen bond (C, black; N, blue; H, white).

Table 1. Screening Results of the Literature Database[17,26−31] (Model Validation) and NCI Database (Virtual Screening)

hypothesis	literature database hits	NCI database hits
HYP2	58	91050
HYP2-HBA[a]	40	676
HYP2-HBA-Shape[b]	26	89
postprocessing filtering[c]	26	76
visual inspection		29

[a] HYP2 modified to match acceptors from five- or six-membered nitrogen-containing heterocycles. [b] Letrozole was converted into a shape query. The minimum similarity tolerance was set to 0.48. [c] Filters applied: Lipinski rule of five, rotatable bonds ≤ 7, PSA < 150.

acceptors from five- or six-membered nitrogen-containing heterocyclic rings (HYP2-HBA) returned 40 molecules (49% of the database). By use of the inclusion volumes (HYP2-HBA-Shape), 26 molecules were retrieved (35% of the database). Furthermore, the final subset was enriched with the most potent aromatase inhibitors. These results demonstrate that the HYP2-HBA-Shape pharmacophore model is able to identify not only inhibitors from the training set but also potent non-steroid aromatase inhibitors with different structures.

NCI Database Virtual Screening. One initial virtual screening (VS) run performed on the NCI database,[32] a large library with more than 290 000 molecules, with HYP2 returned more than 90 000 molecules (Table 1). As expected, the reduced number of pharmacophoric points failed to filter a large database for aromatase inhibitors. The modified pharmacophore query (HYP2-HBA) returned 676 hits, less than 0.3% of the total number of molecules, and with the final hypothesis (HYP2-HBA-Shape) 89 molecules were found, 0.03% of the total number of molecules. Therefore, three different levels of filtering were able to reduce a large database with thousands of molecules to a small subset of promising antiaromatase candidates. Calculations took about 6 min (HYP2), 24 s (HYP2-HBA), and

50 s (HYP2-HBA-Shape) to be completed on an Intel Pentium D 3.4 GHz processor.

Postprocessing Filtering and Inspection of Molecular Properties. The 89 hits derived from the NCI database were superimposed with the HYP2-HBA-Shape pharmacophore query and visually inspected. Furthermore, a Lipinski rule of five filter[33] was applied using the Instant JChem software.[34] The maximum number of rotatable bonds was set to 7 per molecule, and the maximum polar surface area (PSA) was set to 150. Furthermore, compounds with ester groups were avoided because of possible in vivo hydrolysis. Twenty-nine molecules were selected from the VS hits, 15 of these being available to biological evaluation from the NCI open chemical repository collection (Figure 4).

Four different types of nitrogen-containing aromatic heterocycles were retrieved as hits, namely, imidazoles (compounds **1**−**6**), one triazole (compound **7**), pyridines (compounds **8**−**13**), and pyridazines (compounds **14** and **15**). These nitrogen-containing heterocycles are linked to several types of hydrophobic molecular scaffolds such as diphenylmethane (**2**, **3**, **4**, and **14**), arylbenzylsulfane (**1**), diphenylimidazole (**13**), pyridinylbenzoimidazole (**12**), phenylpyridine (**8**), benzyl-1*H*-indol (**5**), acridine (**9**), benzo/methylfuran (**7**), and quinoline (**6**) or in other cases extended with another phenyl ring, such as in two isoquinolines (**10** and **11**) and in one phthalazine (**15**). As hydrogen bond acceptors, methoxy groups (**1**, **7**, **9**, and **14**), nitro groups (**5**, **11**, **12**, and **15**), dioxol groups (**2**, **3**, and **4**), hydroxyl groups (**10** and **13**), imidazole (**6**), and pyridine (**8**) were found.

Before biochemical evaluation, these hits were inspected on a large electronic collection of organic chemistry (CrossFire Beilstein) using the MDL CrossFire Commander.[35] Additional searches were performed using the PubChem compound database,[36] a publicly available source of chemical and biological
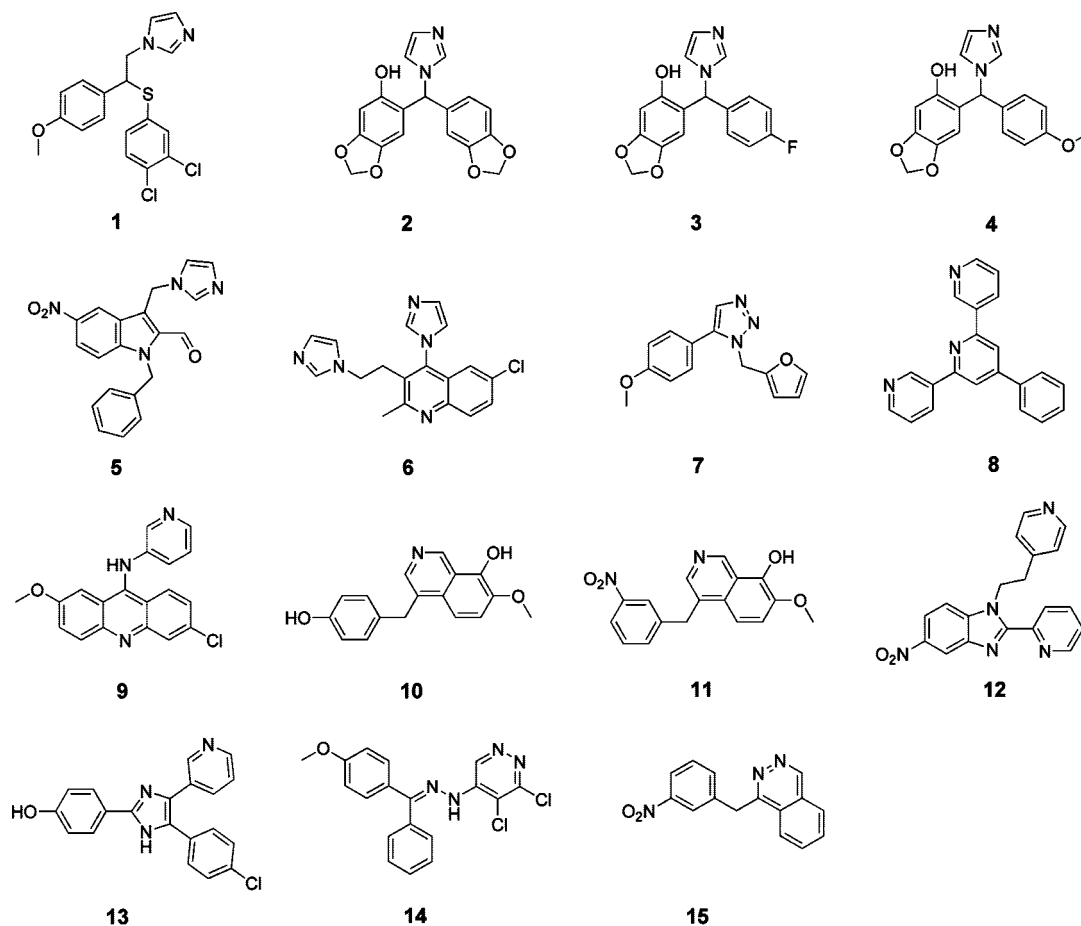
Figure 4. Selected NCI database hits.

information for small molecules, including results from NCI anticancer drug screenings. Of utmost importance was the finding that none of them had been previously tested as aromatase inhibitor.

Biochemical Evaluation. The VS hits were experimentally evaluated for the ability to inhibit aromatase. Enzyme inhibition was studied in vitro in an enzymatic assay according to Siiteri and Thompson[37] using human placental microsomes as the source of aromatase.[38] The molecules were initially screened at 1 μM, followed by a full concentration−response study, allowing the determination of the IC_{50} value (Table 2). Letrozole, a third generation aromatase inhibitor, was used in the same assay conditions as a reference compound.[2]

Interestingly, all molecules in the study were able to inhibit the enzyme aromatase with IC_{50} potencies ranging from micromolar to the low nanomolar range. Compound **3** (IC_{50} = 5.3 nM) presented a better activity than letrozole (IC_{50} = 6.1 nM), whereas compounds **1** (IC_{50} = 25.0 nM), **2** (IC_{50} = 24.9 nM), **4** (IC_{50} = 55.0 nM), **9** (IC_{50} = 76.0 nM), and **13** (IC_{50} = 47.7 nM) had better potency than formestane, a second generation aromatase inhibitor previously studied in our group in the same assay conditions (IC_{50} = 92 nM).[39]

Kinetic analysis of the enzyme in the presence of the most potent inhibitors was also performed (Table 2), and Lineweaver−Burk plots were recorded. For all of them a mechanism of competitive inhibition was found and the K_i values indicate that they have a greater affinity to the active site of the enzyme than the natural substrate androstenedione. Control experiments in the absence of the inhibitors in the study allowed the determination of kinetic constants, Michaelis−Menten constant (K_m =

Table 2. Aromatase Inhibitory Activity of the Selected NCI Database Hits and Letrozole

compd	NCI code	IC_{50} $(nM)^a$	K_i $(nM)^{a,b}$
1	NSC289311	25.0 ± 0.2	6.5 ± 0.1
2	NSC368272	24.9 ± 0.2	5.6 ± 0.1
3	NSC368280	5.3 ± 0.02	1.9 ± 0.03
4	NSC369087	55.0 ± 0.7	13.1 ± 0.2
5	NSC625409	190.3 ± 2.1	52.7 ± 0.8
6	NSC666292	96.0 ± 1.3	33.4 ± 0.4
7	NSC356483	$3.1 \times 10^3 \pm 0.03 \times 10^3$	
8	NSC598	$46.7 \times 10^3 \pm 1.8 \times 10^3$	
9	NSC12999	76.0 ± 1.0	25.5 ± 0.3
10	NSC131735	$29.8 \times 10^3 \pm 0.7 \times 10^3$	
11	NSC131736	$1.4 \times 10^3 \pm 0.01 \times 10^3$	
12	NSC356781	$2.1 \times 10^3 \pm 0.02 \times 10^3$	
13	NSC683634	47.7 ± 0.5	15.1 ± 0.3
14	NSC75308	$5.6 \times 10^3 \pm 0.1 \times 10^3$	
15	NSC226644	$37.6 \times 10^3 \pm 0.8 \times 10^3$	
letrozole		6.1 ± 0.1	2.2 ± 0.03

^{a} Results shown are the mean ± SEM of three independent experiments. ^{b} All compounds with reported K_i values were competitive inhibitors.

92.3 nM), and maximum velocity of catalysis (V_{max} = 160.9 (pmol of substrate/min)/(mg of protein)).

Docking to the Active Site of an Aromatase Homology Model. The most active compounds (**1**−**6**, **9**, and **13**) were docked with the GOLD docking software[40] on the active site of an aromatase homology model built by Favia et al.[12] based on the structure of CYP2C9.[9] The distance between the coordinating nitrogen and the heme iron was constrained to a value between 1.9 and 2.5 Å in order to reproduce the binding mode found for non-steroid aromatase inhibitors.[41] This distance range was chosen on the basis of the analysis of P450 enzymes found in the PDB database with imidazole and pyridine ligands
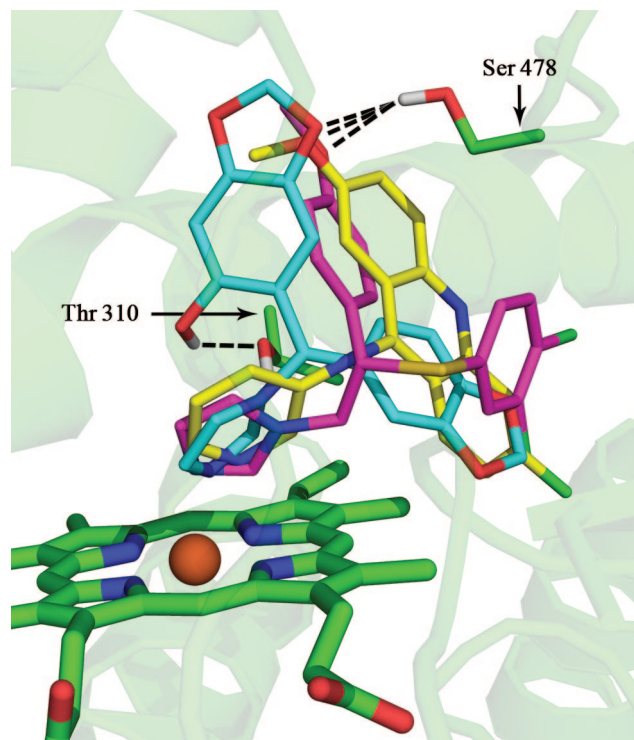
Figure 5. Docked poses of compounds **1** (magenta), **2** (cyan), and **9** (yellow) into the binding site of an aromatase homology model.[12] Dashed lines are drawn between Thr 310 and the hydroxy group of compound **2** (distance = 2.05 Å) and between Ser 478 and methoxy groups of compounds **1** (distance = 2.85 Å) and **9** (distance = 2.61 Å), and a dioxol group of compound **2** (distance = 2.42 Å). This figure was reproduced using Pymol.[55]

complexed with the heme group.[42–44] The GOLDScore fitness function was used to rank the binding modes to the enzyme. This scoring function is the negative of the sum of five energy terms, i.e., protein–ligand hydrogen bond energy, protein–ligand van der Waals energy, ligand internal van der Waals energy, ligand torsional strain energy, and an additional constraint scoring contribution.

Analysis of the docking results revealed that all the molecules in the study fit well within the binding site cavity, with fitness scores ranging from 53 to 68. With the exception of compounds **1** and **6**, the top ranked poses of all molecules have a hydrogen bond between an acceptor group and Ser 478. Slightly lower scored binding modes of compound **1** (62 versus 67, pose ranked number 12) and compound **6** (56 versus 57, pose ranked number 2) showed the same hydrogen bond interaction (Figure 5). These results are in agreement with previous mutagenesis and docking studies pointing out Ser 478 as a key residue to the catalysis mechanism of the enzyme and binding of some non-steroid aromatase inhibitors.[12,15,45] Compounds **2**, **3**, and **4** share the same binding mode, with another hydrogen bond between the hydroxyl group and Thr 310. This extra anchoring point might explain the strongest antiaromatase activity found for compound **3**. More detailed analysis of the docking mode of compounds **1**, **2**, and **9** (Figure 5) revealed a very close superimposition of hydrophobic and hydrogen bond acceptor pharmacophoric points similar to the pharmacophore model used for the VS, HYP2 (Figure 2). The HYD1 area is defined by three phenyl rings docked on an apolar region of the aromatase binding cavity, defined by hydrophobic residues Val 313 and Leu 477. The HYD2 area has either phenyl rings or a benzylsulfanyl group fitting a small hydrophobic pocket due to Pro 368 and Val 370.

Letrozole, the reference compound used in this study, fits into the binding pocket of the enzyme in a similar way to the other compounds. Good superimposition is observed between the hydrophobic and nitrogen-containing moieties of these molecules.

Furthermore, the docking poses of compounds **1–6**, **9**, and **13** were loaded into Catalyst and rigidly fitted into the pharmacophore model. With the exception of compound **5**, all the other docked conformations correctly fitted four groups into the pharmacophore model. The average fit value was 2.45.

These results validate our pharmacophore model as a screening methodology for potent non-steroid aromatase inhibitors.

Conclusions

In this study, a new virtual screening strategy for potent non-steroid aromatase inhibitors was proposed. The screening was based on the common features of a training set of second and third generation aromatase inhibitors (fadrozole, anastrozole, letrozole, and vorozole), combined with heme coordinating fragments (five- or six-membered nitrogen-containing aromatic heterocycles), inclusion volumes, and several "druglikeness" filters. The model was first validated theoretically on a literature database of known aromatase inhibitors and used subsequently to screen the NCI database, a large library of molecules focused on research for new antitumor and AIDS antiviral agents. A very small number of the NCI compounds (0.03%) was identified as interesting antiaromatase candidates. Finally, the model was validated experimentally by testing the most promising VS hits in an in vitro biochemical assay. New potent aromatase inhibitors with similar or better in vitro potency than the reference molecules were identified.

All tested compounds were able to inhibit the aromatase with IC_{50} potencies ranging from micromolar to the low nanomolar range. These results demonstrate that our VS strategy was very effective at filtering this large compound database with no false positives. Furthermore, several compounds had antiaromatase activities between second and third generation aromatase inhibitors (**1**, **2**, **4**, **9**, and **13**), and one of them (compound **3**) was stronger than the third generation reference. Kinetic studies confirmed a competitive inhibition mechanism. These molecules are therefore interesting targets for lead optimization and ADMET studies.

Although to the best of our knowledge these compounds had never been tested as aromatase inhibitors, molecules with the diphenylmethylimidazole or the 3-((1H-imidazol-1-yl)methyl)-1-benzyl-1H-indole motifs have already been reported as potent aromatase inhibitors with IC_{50} values in the nanomolar range.[46–48] Furthermore, compound **12** was tested previously for antimicrobial and antifungal properties with no activity on the assay conditions,[49] and compound **5** derivatives had antifungal activity only when tested at micromolar concentrations.[50] These results indicate that such compounds, used at very low nanomolar concentrations, might be selective for the aromatase enzyme. However, additional studies must be performed in order to confirm this hypothesis. Other interesting pharmacological activities have also been identified for some of these molecules and reported in the literature; for example, 9-anilinoacridines related to compound **9** have potential antitumor activity due to their DNA binding properties.[51] Large scale screenings using the NCI open chemical repository collection on the basis of the NCI anticancer drug discovery program identified potential new activities for compounds **1** and **6**. Compound **1** was active in vivo using a model for colon carcinoma in mice. Compound **6** was another interesting compound identified. It was found to inhibit the growth of a human leukemia cell line.

Docking of the most potent compounds into the active site of a homology model of the aromatase[12] led to a common binding mode that is in good agreement with the common features in the pharmacophore model. Besides the essential HBA1 feature involved in coordination with the heme iron, establishment of an additional hydrogen bond with Ser 478 appears to be important, therefore justifying the presence of the HBA2 feature. Several apolar active site residues explain the hydrophobic nature of the new aromatase inhibitors (HYD1 and HYD2), and Thr 310, involved in an additional hydrogen bond, might explain the strong antiaromatase properties of compound **3**.

These results, supported by extensive experimental validation, represent a clear improvement over previous ligand-based virtual screening approaches for new aromatase inhibitors,[21] providing an effective new tool for the identification of novel molecular entities acting on important cytochrome P450 targets. Common to all members of this large superfamily is a heme cofactor that should be carefully addressed when designing new pharmacophore models for such targets in order to achieve strong binding affinities. On the other hand, fine-tuning of the ligand potency and selectivity among different P450 targets can be addressed, exploring specific hydrophobic, hydrogen bonding, and shape complementarity interactions. Therefore, starting with potent inhibitors of different cytochrome P450 enzymes such as CYP17 and aldosterone synthase, this methodology can be easily applied to the design of new drugs for other relevant diseases.

The pharmacophore model built includes a broad range of nitrogen-containing heterocycles comprising imidazole, triazole, and pyridine, the most extensively used fragments to design potent non-steroid aromatase inhibitors. However, it might be extended to include other heterocycles such as rings containing both nitrogen and sulfur or oxygen heteroatoms (thiazole, oxazole).

In summary, we have developed and validated a new ligand-based strategy for virtual screening of non-steroid aromatase inhibitors. Besides the strong antiaromatase compounds identified on the NCI database, it should be possible to screen other large compound libraries and find new potent molecules.

Experimental Section

Materials and General Methods. The NCI selected compounds were obtained from the Drug Synthesis and Chemistry Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis of the National Cancer Institute. Letrozole was purchased from USP (Rockville, MD). NADPH was purchased from Sigma-Aldrich (St. Louis, MO). The [1β-³H]androstenedione (specific activity of 25.3 Ci/mmol) and the liquid scintillation cocktail Optiphase Hisafe 2 were purchased from PerkinElmer (Boston, MA), and the radioactive samples were counted on a Packard Tri-Carb 2900 TR liquid scintillation analyzer.

Pharmacophore Modeling. The common-features model for non-steroid aromatase inhibitors was performed using the Catalyst software.[23] Potent aromatase inhibitors were collected from the literature and submitted to a conformational search with internal energy minimization. A maximum of 250 conformers were saved within an energy window of 20 kcal/mol above the global minimum, using the best quality generation type. Four of these molecules (anastrozole, letrozole, S-vorozole, and S-fadrozole)[6] were included in the training set and used to build pharmacophore hypotheses using the HipHop[22] algorithm of Catalyst. The remaining molecules[17,26−31] were included in the test set. The most active enantiomers of vorozole[52] and letrozole[41] were preferred. For the initial hypothesis generation methodology, hydrophobic and hydrogen bond acceptor functions were used. The "Principal" value

was set to 2 for letrozole (all features in the molecule were considered to build the pharmacophore model) and 1 for the other compounds (at least one mapping for each generated hypothesis was found). The "Maximum Omitted Features" value was set to 1 for all the molecules (all but one feature was forced to map), and default settings were used for the other options.

The nitrogen-containing aromatic heterocyclic rings were built within Catalyst, converted into pharmacophore features using the "View Hypothesis" workbench, and combined into a new single fragment/function feature using the "Exclude/Or" tool. A similarity tolerance of 0.48 was used in the shape query. This value was chosen in order to match all the molecules of the training set.

Virtual Screening. The NCI database was downloaded from the 2007 release of the ZINC library[53] and converted into a multiconformer Catalyst database. A maximum of 100 conformations were generated for each molecule using the "FAST" conformational analysis model of the catDB utility program. Each pharmacophore hypothesis was screened using the "fast flexible database search" settings.

Postprocessing Filtering. Instant JChem[34] was used for management, search, and prediction of molecular descriptors for the NCI hits. A Lipinski rule of five[33] filter was applied (not more than 5 hydrogen bond donors, not more than 10 hydrogen bond acceptors, molecular weight under 500 g/mol, and calculated partition coefficient clogP of less than 5), as well as a filter based on the maximum number of rotatable bonds (not more than 7) and the maximum polar surface area (less than 150). Substructure searches were performed using the same software.

Human Placental Microsomes Isolation. Human term placental microsomes were obtained by differential centrifugation according to the method described by Ryan[38] and were used as a source of aromatase. The microsomes were resuspended in a 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 as previously described by Siiteri and Thompson.[37] Briefly, enzymatic assays were performed at 37 °C in medium containing 67 mM sodium phosphate buffer, pH 7.5, [1β-³H]androstenedione (6.6 × 10⁵ dpm), and 270 μM NADPH in a final volume of 500 μL. The inhibitors in the study were dissolved in DMSO. The amount of DMSO in the assay was always 2%. The reaction was started with the addition of 30 μg of microsomal protein and stopped after 20 min by adding 1 mL of chloroform and vortexing at 9000 rpm for 40 s. The tritiated water formed when the enzyme converts the tritiated substrate [1β-³H]androstenedione to estrone was quantified by liquid scintillation counting. Appropriate controls without inhibitor were performed in order to determine the maximum enzymatic activity, to which the relative percentage of inhibition was determined. All the molecules in the study were initially tested at 1 μM, followed by a full concentration−response experiment with at least eight concentrations ranging from 0.3162 nM to 160 μM. The assays were performed three times, each one in duplicate, and the results were treated by nonlinear regression analysis using a sigmoidal concentration−response curve with variable slope.

Kinetic Analysis. For the kinetic study, conditions similar to those of the concentration−response experiment were used. The concentration of [1β-³H]androstenedione was varied from 7.5 to 100 nM, three different concentrations of each inhibitor were tested, and the reaction time was 5 min. An assay without inhibitor was also performed. The assay was performed three times, each one in duplicate, and the data were fitted by nonlinear regression to the Michaelis−Menten equation. The kinetic constants, V_{max}, K_m, and K_i were estimated from nonlinear regression, and the type of inhibition was determined from Lineweaver−Burk plots.

Docking Calculations. The GOLD software[40] was used to perform flexible docking of the best compounds into the binding cavity of a homology model of the aromatase.[12] The binding site was defined as a 10 Å sphere centered approximately 5 Å above the heme iron and including all active site residues important to

the binding. An octahedral coordinating geometry was assigned to the heme iron, and the GOLDScore fitness function was used with metal parameters optimized for P450 enzymes, taking account of different H-bond acceptor types.[54] Since it has been reported that in some cases the GOLD program fails to reproduce the known binding mode of non-steroid aromatase inhibitors, i.e., binding through coordination between the heme iron and an acceptor nitrogen from an aromatic heterocycle,[15] substructure constraints were created for the pyridine and imidazole rings in order place the nitrogen within a distance between 1.9 and 2.5 Å from the heme iron. These upper and lower values were chosen on the basis of the analysis of PDB entries of cytochrome P450 enzymes complexed with imidazole and pyridine derivatives.[42−44] Twenty independent docking runs were performed with the default genetic algorithm search parameters. For comparison, unrestrained docking was also performed, being able to identify docking modes in which the nitrogen-containing heterocycle interacted correctly with the heme iron but, however, for most molecules, failing to score these poses in the top ranked solutions. As a consequence, the docking algorithm spent most of the computational time exploring unrealistic conformations.

Supporting Information Available: Results from elemental analysis and model validation. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Meunier, B.; de Visser, S. P.; Shaik, S. Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes. *Chem. Rev.* **2004**, *104*, 3947–3980.

(2) Smith, I. E.; Dowsett, M. Aromatase inhibitors in breast cancer. *N. Engl. J. Med.* **2003**, *348*, 2431–2442.

(3) Popat, S.; Smith, I. E. Breast cancer. *Update Cancer Ther.* **2006**, *1*, 187–210.

(4) Rendic, S. Summary of information on human CYP enzymes: human P450 metabolism data. *Drug Metab. Rev.* **2002**, *34*, 83–448.

(5) Nelson, D. R.; Zeldin, D. C.; Hoffman, S. M. G.; Maltais, L. J.; Wain, H. M.; Nebert, D. W. Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants. *Pharmacogenetics* **2004**, *14*, 1–18.

(6) Recanatini, M.; Cavalli, A.; Valenti, P. Nonsteroidal aromatase inhibitors: recent advances. *Med. Res. Rev.* **2002**, *22*, 282–304.

(7) Williams, P. A.; Cosme, J.; Sridhar, V.; Johnson, E. F.; Mcree, D. E. Microsomal cytochrome P4502C5: comparison to microbial P450s and unique features. *J. Inorg. Biochem.* **2000**, *81*, 183–190.

(8) Williams, P. A.; Cosme, J.; Vinkovic, D. M.; Ward, A.; Angove, H. C.; Day, P. J.; Vonrhein, C.; Tickle, I. J.; Jhoti, H. Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone. *Science* **2004**, *305*, 683–686.

(9) Williams, P. A.; Cosme, J.; Ward, A.; Angove, H. C.; Matak, V. D.; Jhoti, H. Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* **2003**, *424*, 464–468.

(10) Rowland, P.; Blaney, F. E.; Smyth, M. G.; Jones, J. J.; Leydon, V. R.; Oxbrow, A. K.; Lewis, C. J.; Tennant, M. G.; Modi, S.; Eggleston, D. S.; Chenery, R. J.; Bridges, A. M. Crystal structure of human cytochrome P450 2D6. *J. Biol. Chem.* **2006**, *281*, 7614–7622.

(11) Loge, C.; Le Borgne, M.; Marchand, P.; Robert, J. M.; Le Baut, G.; Palzer, M.; Hartmann, R. W. Three-dimensional model of cytochrome P450 human aromatase. *J. Enzyme Inhib. Med. Chem.* **2005**, *20*, 581–585.

(12) Favia, A. D.; Cavalli, A.; Masetti, M.; Carotti, A.; Recanatini, M. Three-dimensional model of the human aromatase enzyme and density functional parameterization of the iron-containing protoporphyrin IX for a molecular dynamics study of heme-cysteinato cytochromes. *Proteins* **2006**, *62*, 1074–1087.

(13) Karkola, S.; Holtje, H. D.; Wahala, K. A three-dimensional model of CYP19 aromatase for structure-based drug design. *J. Steroid Biochem. Mol. Biol.* **2007**, *105*, 63–70.

(14) Neves, M. A. C.; Dinis, T. C. P.; Colombo, G.; Melo, M. L. S. Combining computational and biochemical studies for a rationale on the anti-aromatase activity of natural polyphenols. *ChemMedChem* **2007**, *2*, 1750–1762.

(15) Jackson, T.; Woo, L. W. L.; Trusselle, M. N.; Purohit, A.; Reed, M. J.; Potter, B. V. L. Non-steroidal aromatase inhibitors based on a biphenyl scaffold: synthesis, in vitro SAR, and molecular modelling. *ChemMedChem* **2008**, *3*, 603–618.

(16) Cavalli, A.; Greco, G.; Novellino, E.; Recanatini, M. Linking CoMFA and protein homology models of enzyme−inhibitor interactions: an application to non-steroidal aromatase inhibitors. *Bioorg. Med. Chem.* **2000**, *8*, 2771–2780.

(17) Gobbi, S.; Cavalli, A.; Rampa, A.; Belluti, F.; Piazzi, L.; Paluszcak, A.; Hartmann, R. W.; Recanatini, M.; Bisi, A. Lead optimization providing a series of flavone derivatives as potent nonsteroidal inhibitors of the cytochrome P450 aromatase enzyme. *J. Med. Chem.* **2006**, *49*, 4777–4780.

(18) Bak, A.; Polanski, J. Modeling robust QSAR 3: SOM-4D-QSAR with iterative variable elimination IVE-PLS: application to steroid, azo dye, and benzoic acid series. *J. Chem Inf. Model.* **2007**, *47*, 1469–1480.

(19) Su, B.; Tian, R.; Darby, M. V.; Brueggemeier, R. W. Novel sulfonanilide analogs decrease aromatase activity in breast cancer cells: synthesis, biological evaluation, and ligand-based pharmacophore identification. *J. Med. Chem.* **2008**, *51*, 1126–1135.

(20) Oprea, T. I.; Garcia, A. E. Three-dimensional quantitative structure−activity relationships of steroid aromatase inhibitors. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 186–200.

(21) Schuster, D.; Laggner, C.; Steindl, T. M.; Paluszcak, A.; Hartmann, R. W.; Langer, T. Pharmacophore modeling and in silico screening for new P450 19 (aromatase) inhibitors. *J. Chem. Inf. Model.* **2006**, *46*, 1301–1311.

(22) Clement, O. O.; Mehl, A. T. HipHop: Pharmacophores Based on Multiple Common-Feature Alignments. In *Pharmacophore Perception, Development, and Use in Drug Design*; International University Line: La Jolla, CA, 2000; pp 69−84.

(23) *Catalyst*, release version 4.11; Accelrys Software Inc. (www.accelrys.com).

(24) Cavalli, A.; Recanatini, M. Looking for selectivity among cytochrome P450s inhibitors. *J. Med. Chem.* **2002**, *45*, 251–254.

(25) Schleinkofer, K.; Sudarko; Winn, P. J.; Ludemann, S. K.; Wade, R. C. Do mammalian cytochrome P450s show multiple ligand access pathways and ligand channelling? *EMBO Rep.* **2005**, *6*, 584–589.

(26) Jacobs, C.; Frotscher, M.; Dannhardt, G.; Hartmann, R. W. 1-Imidazolyl(alkyl)-substituted di- and tetrahydroquinolines and analogues: syntheses and evaluation of dual inhibitors of thromboxane A(2) synthase and aromatase. *J. Med. Chem.* **2000**, *43*, 1841–1851.

(27) Recanatini, M.; Bisi, A.; Cavalli, A.; Belluti, F.; Gobbi, S.; Rampa, A.; Valenti, P.; Palzer, M.; Paluszcak, A.; Hartmann, R. W. A new class of nonsteroidal aromatase inhibitors: design and synthesis of chromone and xanthone derivatives and inhibition of the P450 enzymes aromatase and 17 alpha-hydroxylase/C17,20-lyase. *J. Med. Chem.* **2001**, *44*, 672–680.

(28) Leonetti, F.; Favia, A.; Rao, A.; Aliano, R.; Paluszcak, A.; Hartmann, R. W.; Carotti, A. Design, synthesis, and 3D QSAR of novel potent and selective aromatase inhibitors. *J. Med. Chem.* **2004**, *47*, 6792–6803.

(29) Cavalli, A.; Bisi, A.; Bertucci, C.; Rosini, C.; Paluszcak, A.; Gobbi, S.; Giorgio, E.; Rampa, A.; Belluti, F.; Piazzi, L.; Valenti, P.; Hartmann, R. W.; Recanatini, M. Enantioselective nonsteroidal aromatase inhibitors identified through a multidisciplinary medicinal chemistry approach. *J. Med. Chem.* **2005**, *48*, 7282–7289.

(30) Kim, Y. W.; Hackett, J. C.; Brueggemeier, R. W. Synthesis and aromatase inhibitory activity of novel pyridine-containing isoflavones. *J. Med. Chem.* **2004**, *47*, 4032–4040.

(31) Hackett, J. C.; Kim, Y. W.; Su, B.; Brueggemeier, R. W. Synthesis and characterization of azole isoflavone inhibitors of aromatase. *Bioorg. Med. Chem.* **2005**, *13*, 4063–4070.

(32) Developmental Therapeutics Program NCI/NIH. http://dtp.nci.nih.gov/.

(33) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **1997**, *23*, 3–25.

(34) *Instant JChem*, release version 2.0.0; ChemAxon (www.chemaxon.com).

(35) *MDL CrossFire Commander*, release version 7.0SP2; Elsevier MDL (www.mdli.com).

(36) PubChem Compound database available at http://www.ncbi.nlm.nih.gov/sites/entrez/.

(37) Siiteri, P. K.; Thompson, E. A. Studies of human placental aromatase. *J. Steroid Biochem.* **1975**, *6*, 317–322.

(38) Ryan, K. J. Biological aromatization of steroids. *J. Biol. Chem.* **1959**, *234*, 268–272.

(39) Neves, M. A.; Dinis, T. C.; Colombo, G.; Sa E Melo, L. Biochemical and computational insights into the anti-aromatase activity of natural catechol estrogens. *J. Steroid Biochem. Mol. Biol.* **2008**, *110*, 10–17.

(40) *GOLD*, release version 3.2; The Cambridge Crystallographic Data Centre (http://www.ccdc.cam.ac.uk).

(41) Furet, P.; Batzl, C.; Bhatnagar, A.; Francotte, E.; Rihs, G.; Lang, M. Aromatase inhibitors: synthesis, biological activity, and binding mode of azole-type compounds. *J. Med. Chem.* **1993**, *36*, 1393–1400.

(42) Cupp-Vickery, J. R.; Garcia, C.; Hofacre, A.; McGee-Estrada, K. Ketoconazole-induced conformational changes in the active site of cytochrome P450eryF. *J. Mol. Biol.* **2001**, *311*, 101–110.

(43) Podust, L. M.; von Kries, J. P.; Eddine, A. N.; Kim, Y.; Yermalitskaya, L. V.; Kuehne, R.; Ouellet, H.; Warrier, T.; Altekoster, M.; Lee, J. S.; Rademann, J.; Oschkinat, H.; Kaufmann, S. H. E.; Waterman, M. R. Small-molecule scaffolds for CYP51 inhibitors identified by high-throughput screening and defined by X-ray crystallography. *Antimicrob. Agents Chemother.* **2007**, *51*, 3915–3923.

(44) Raag, R.; Li, H. Y.; Jones, B. C.; Poulos, T. L. Inhibitor-induced conformational change in cytochrome-P-450(Cam). *Biochemistry* **1993**, *32*, 4571–4578.

(45) Kao, Y. C.; Korzekwa, K. R.; Laughton, C. A.; Chen, S. Evaluation of the mechanism of aromatase cytochrome P450. A site-directed mutagenesis study. *Eur. J. Biochem.* **2001**, *268*, 243–251.

(46) Jones, C. D.; Winter, M. A.; Hirsch, K. S.; Stamm, N.; Taylor, H. M.; Holden, H. E.; Davenport, J. D.; Vonkrumkalns, E.; Suhr, R. G. Estrogen synthetase inhibitors. 2. Comparison of the in vitro aromatase inhibitory activity for a variety of nitrogen-heterocycles substituted with diarylmethane or diarylmethanol groups. *J. Med. Chem.* **1990**, *33*, 416–429.

(47) Bhatnagar, A. S.; Hausler, A.; Schieweck, K.; Browne, L. J.; Bowman, R.; Steele, R. E. Novel aromatase inhibitors. *J. Steroid Biochem. Mol. Biol.* **1990**, *37*, 363–367.

(48) Le Borgne, M.; Marchand, P.; Duflos, M.; Delevoye-Seiller, B.; Piessard-Robert, S.; Le Baut, G.; Hartmann, R. W.; Palzer, M. Synthesis and in vitro evaluation of 3-(1-azolylmethyl)-1*H*-indoles and 3-(1-azolyl-1-phenylmethyl)-1*H*-indoles as inhibitors of P450 arom. *Arch. Pharm.* **1997**, *330*, 141–145.

(49) Hisano, T.; Ichikawa, M.; Tsumoto, K.; Tasaki, M. Synthesis of benzoxazoles, benzothiazoles and benzimidazoles and evaluation of their antifungal, insecticidal and herbicidal activities. *Chem. Pharm. Bull.* **1982**, *30*, 2996–3004.

(50) Na, Y. M.; Le Borgne, M.; Pagniez, F.; Le Baut, G.; Le Pape, P. Synthesis and antifungal activity of new 1-halogenobenzyl-3-imidazolylmethylindole derivatives. *Eur. J. Med. Chem.* **2003**, *38*, 75–87.

(51) Baguley, B. C.; Denny, W. A.; Atwell, G. J.; Cain, B. F. Potential anti-tumor agents. 34. Quantitative relationships between DNA-binding and molecular-structure for 9-anilinoacridines substituted in the anilino ring. *J. Med. Chem.* **1981**, *24*, 170–177.

(52) Vanden Bossche, H.; Willemsens, G.; Roels, I.; Bellens, D.; Moereels, H.; Coene, M. C.; Le Jeune, L.; Lauwers, W.; Janssen, P. A. R 76713 and enantiomers: selective, nonsteroidal inhibitors of the cytochrome P450-dependent oestrogen synthesis. *Biochem. Pharmacol.* **1990**, *40*, 1707–1718.

(53) Irwin, J. J.; Shoichet, B. K. ZINC—a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45*, 177–182.

(54) Kirton, S. B.; Murray, C. W.; Verdonk, M. L.; Taylor, R. D. Prediction of binding modes for ligands in the cytochromes p450 and other heme-containing proteins. *Proteins* **2005**, *58*, 836–844.

(55) DeLano, W. L. *The PyMOL Molecular Graphics System*, release version 0.99; DeLano Scientific: Palo Alto, CA (http://www.pymol.org).