Pharmacophore Modeling and in Silico Screening for New P450 19 (Aromatase) Inhibitors†

Daniela Schuster,[#,‡] Christian Laggner,[#,‡] Theodora M. Steindl,[#,‡] Anja Palusczak,[§] Rolf W. Hartmann,[§] and Thierry Langer*[,#,‡]

Institute of Pharmacy, Department of Pharmaceutical Chemistry, University of Innsbruck, Innrain 52, A-6020 Innsbruck, Austria, Center for Molecular Biosciences Innsbruck (CMBI), Peter-Mair-Str. 1a, A-6020 Innsbruck, Austria, and Pharmaceutical and Medicinal Chemistry, Saarland University, P.O. Box 151150, D-66041 Saarbrücken, Germany

Received June 13, 2005

Cytochrome P450 19 (P450 19, aromatase) constitutes a successful target for the treatment of breast cancer. This study analyzes chemical features common to P450 19 inhibitors to develop ligand-based, selective pharmacophore models for this enzyme. The HipHop and HypoRefine algorithms implemented in the Catalyst software package were employed to create both common feature and quantitative models. The common feature model for P450 19 includes two ring aromatic features in its core and two hydrogen bond acceptors at the ends. The models were used as database search queries to identify active compounds from the NCI database.

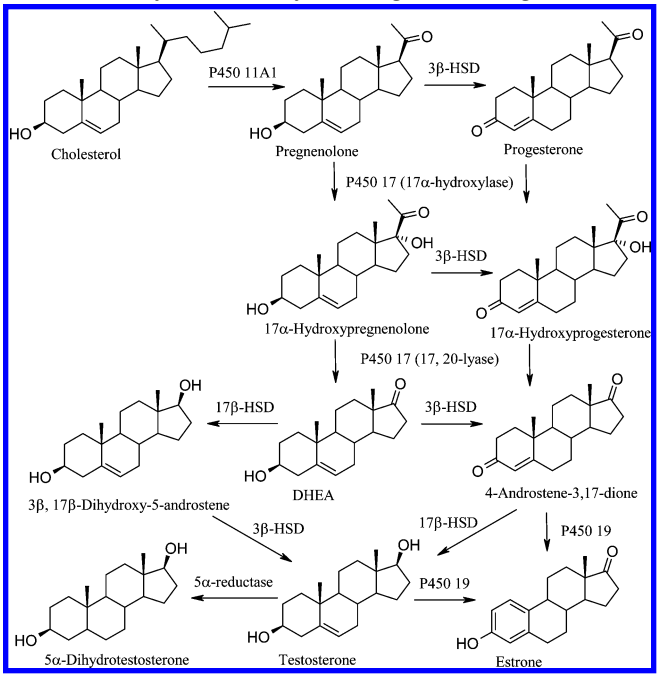
INTRODUCTION

Sex hormones play crucial roles in living organisms. They regulate growth, maturing, and reproduction of most life forms on earth. Structurally, the main part of sex hormones is composed of a steroidal body with different substituents. Chemically very similar sex hormones can produce totally different effects. Androgens promote e.g. muscle growth, alopecia, and the libido. Estrogens—synthesized from androgens by aromatization (Scheme 1)—produce e.g. water retention, dilatation of the capillary vessels, and a decrease of sebaceous gland activity.[1]

Apart from their physiological role, sex hormones can be involved in pathophysiological processes such as tumorgenesis and -growth. About 80% of all prostate cancers and 67% of all breast cancers are sex hormone dependent.[2,3] This allows the application of an endocrine therapy with a more favorable active profile and fewer side effects compared to unspecific chemotherapy.

Breast Cancer. Estrogens can be involved in the genesis of breast cancer as well as in the growth of established tumors. The exact mechanisms by which estrogens achieve these effects are still hypothesized. They may increase proliferation within the breast that in turn may increase susceptibility to genetic mistakes. Estrogens could increase the growth rate of occult breast cancers so that they become detectable. At last, estrogen metabolites—reactive quinones which can directly interact with DNA and cause mutations—could account for these effects.[4] More than 100 years ago, Beatson et al. showed that ovariectomy in premenopausal patients with breast cancer induced tumor remission.[5] After the advantageous effect of estrogen level reduction on breast cancer growth was established, antihormonal approaches were conducted to develop drugs. Estrogen receptor antagonists prevent estrogen from activating its receptor, while cytochrome P450 19 (aromatase, P450 19) inhibitors directly inhibit estrogen biosynthesis (Scheme 1). Today, selective estrogen receptor antagonists (SERMs) and P450 19 inhibitors constitute important therapeutical weapons in the fight against breast cancer deaths.[6] P450 19 inhibitors currently in clinical use include the steroidal enzyme inactivators formestane and exemestane as well as the nonsteroidal inhibitors aminoglutethimide, anastrozole, and letrozole.[4]

Scheme 1. Biosynthesis Pathway of Androgens and Estrogens



* Corresponding author phone: (+43) 512/507-5252; fax: (+43) 512/507-5269; e-mail: Thierry.Langer@uibk.ac.at.

† Parts of this study were presented as talks at the WATOC'05 conference in Cape Town and at the 19. Darmstädter Molecular Modeling Workshop in Erlangen.

[#] University of Innsbruck.

[‡] Center for Molecular Biosciences.

[§] Saarland University.

10.1021/ci050237k CCC: $33.50
Published on Web 04/18/2006

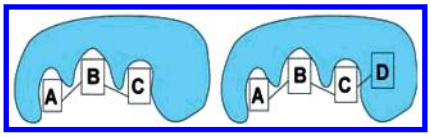


Figure 1. ABC and ABCD share the same pharmacophore. ABC is an active compound; ABCD is inactive. By placing an excluded volume sphere at position D, both compounds can be assessed correctly as being active or inactive, respectively.[15]

Previous Approaches. A QSAR analysis for a family of substituted dichlorodiphenyl compounds suggested an intermolecular hydrogen bond or strong electrostatic interaction between a nitrogen atom in the inhibitor heterocycle and P450 19.[7] Other groups applied molecular dynamic studies for P450 19.[8,9] Cavalli and Recanatini were the first to propose general physicochemical ligand features for strong and selective P450 17 or P450 19 inhibition.[10] Apart from an electron rich heterocycle (preferably an imidazole ring) to interact with the iron atom of the heme group, a general high hydrophobic character is required for the inhibition of both enzymes. Site-directed mutagenesis experiments and docking studies indicate that the presence of Ser478—comprising a hydrogen bond donor (HBD) group—in the active site of P450 19 is the main difference between both enzyme ligand binding domains (LBDs). In the LBD of P450 17, no such HBD group was located; however, a further hydrophobic amino acid residue (Val483) is reported to be located in the P450 17 active site and contributes to the high binding affinity of hydrophobic compounds.[10,11]

We here present pharmacophore models for P450 19 inhibitors. After generation of a druglike virtual molecular database, a sophisticated in silico compound selection process can lead to promising candidates for drug development.

MATERIALS AND METHODS

The computational molecular modeling studies were carried out using a SGI Octane R12000 workstation running IRIX 6.5.10 or on a Athlon-1900 PC running MS Windows 2000Pro.

Pharmacophore Modeling with Catalyst. As there is no X-ray crystal structure of P450 19 published, a ligand-based approach was used to generate pharmacophore models derived from highly active and selective P450 19 inhibitors. The compounds of the training and test set were submitted to internal strain energy minimization and conformational analysis (max. number of conformers = 250, generation type: best quality, energy range = 20 kcal/mol above the calculated global minimum). Using the HipHop algorithm of Catalyst,[12,13] common feature pharmacophore models for highly active compounds were built. For each training set, 10 hypotheses were returned by the hypotheses generation process. The resulting hypotheses consist of three-dimensional arrangements of several default feature types (e.g. hydrogen bond acceptor, hydrogen bond donor, hydrophobic, ring aromatic, positive/negative ionizable) located at defined positions (location constraints). These are surrounded by certain spatial tolerance spheres, assessing the area in space that should be assigned by the corresponding chemical functions of the matched molecule. Hydrogen bond acceptors, donors, and ring aromatic features additionally include a vector, indicating the direction of the interaction.

Another pharmacophore generation algorithm implemented in Catalyst—HypoGen/HypoRefine[14,15]—was used to create hypotheses that can estimate the activity of a compound by calculating the geometric fit of the molecular structure to the given pharmacophore model.

By using the Shape algorithm of Catalyst,[13] spatial information of highly active compounds can be converted into an additional steric constraint model. Such a shape model excludes compounds that do not fit in the same space as the template molecule(s) and helps to reduce abundant hitlists derived from database mining.

A so-called stockroom database search was conducted with all pharmacophore models. The stockroom database consists of all compounds of the training and test set used for model generation and validation.

Overlays of the pharmacophore models were performed by using the Compare algorithm of Catalyst.[13]

The HypoGen and HypoRefine Algorithms. HypoGen tries to find hypotheses that are common among the active compounds of the training set but do not reflect the inactive ones. Pharmacophores that correlate best the three-dimensional arrangement of features in a given set of training compounds with the corresponding pharmacological activities (IC_{50} or K_i) are constructed and ranked. To ensure the statistic relevance of the calculated model, the training set should contain at least 16 compounds together with their activity values derived from comparable binding assays. Activities should spread equally over at least 4 orders of magnitude.[14] Each feature of the resulting hypotheses occupies a certain weight that is proportional to its relative contribution to biological activity. HypoGen therefore constructs pharmacophore models that correlate best with measured activities and that consist of as few features as possible. The HypoGen hypothesis generation process is executed in three steps: the constructive phase, the subtractive phase, and the optimization phase.[16]

In the constructive phase, hypotheses that are common to the most active set of compounds are identified. HypoGen enumerates all possible pharmacophore configurations using all combinations of pharmacophore features for each of the conformations of the two most active compounds. Additionally, the hypotheses must fit a minimum subset of features of the remaining most active compounds in order to be considered. At the end of the constructive phase, a large database of pharmacophore configurations is generated.

In the subtractive phase, all pharmacophore configurations that are also present in the least active set of molecules are removed. All compounds whose activity is by default 3.5 orders of magnitude less than that of the most active compound are considered to represent the least active molecules. The value 3.5 is adjustable depending on the activity range of the training set.

During the optimization phase, the hypothesis score is improved. Hypotheses are scored based on errors in activity estimates from regression and complexity. The optimization involves a variation of features and/or locations to optimize activity prediction via a simulated annealing approach. After the geometric fit values of the training set compounds are calculated, a linear regression of −log(activity) versus the geometric fit is performed. The "total cost" parameter—which will be explained later—is calculated for each new hypothesis. When the optimization process no longer improves the score,

HypoGen stops and reports the top scoring 10 unique pharmacophores.

The hypotheses are scored according to their "cost analysis".[17] The cost is a quantitative extension of Occam's razor (everything else being equivalent, the simplest model is the best). This simplicity is defined using the minimum description length principle from information theory. The overall costs of a model consist of three cost components: the weight cost, the error cost, and the configuration cost. The weight component is a value that increases in a Gaussian form as this function weights in a model deviate from the ideal value of two. The error cost represents the root-mean-squared difference between estimated and measured activities of the training set. The configuration cost quantifies the entropy of the hypothesis space. If the input information is too multiplex, e.g. due to too flexible training set molecules, this would lead to an effusive number of hypotheses as outcome of the subtractive phase. This configuration cost should not exceed a maximum value of 17 which corresponds to a number of 2^{17} pharmacophore models. It has been empirically determined that higher values often lead to a good correlation by chance. Catalyst cannot consider a higher number of models in the optimization phase, and so the rest of generated hypotheses from the constructive and subtractive phase are left out of the process.

Two additional cost calculations are performed by HypoGen. The fixed cost is the lowest possible cost representing a hypothetical simplest model that fits all data perfectly. Fixed costs are calculated by adding the minimum achievable error and weight cost and the constant configuration cost. Another cost parameter—the null cost—represents the maximum cost of a pharmacophore with no features and estimates activity to be the average of the training set molecules' activity data. The null cost value is equal to the maximum occurring error cost. The greater the difference between these two cost values and the closer the total cost of the generated hypothesis is to the fixed cost, the more statistically significant the hypothesis is supposed to be. According to randomized studies, a cost difference of 40−60 between the total cost and the null cost indicates a 75−90% chance of representing a true correlation in the data.

HypoRefine is an extension of the HypoGen algorithm that uses strategically placed excluded volume spheres (XVOLs)—"forbidden" spheres a compound must not map—where inactivity of compounds with the same pharmacophore features is due to steric clashes (Figure 1).

The constructive phase of HypoRefine is identical to the HypoGen algorithm. The subtractive phase is not operated which results in a pharmacophore configuration space populated with the most active compounds in the training set. Subsequently, XVOLs are included in the simulated annealing optimization process. The automated identification of excluded volume spheres is performed in four steps: (i) alignment of active molecules, (ii) alignment of inactive molecules with the highest fit score, (iii) identification of atoms in the aligned inactive compounds that are far away from those in the aligned actives, and (iv) random selection from those points and assignment as excluded volume sphere. Pharmacophore models that also fit inactive molecules are automatically penalized. An explicit specification of inactive compounds as well as the maximum number of excluded volume spheres, excluded volume spheres move types, and softening excluded volume spheres during fitting can be adjusted by the user.

Definition of a New Pharmacophore Feature. Pharmacophore modeling for P450 19 inhibitors required the definition of a new Catalyst feature. Although organic fluorine and chlorine are able to interact with electrophilic structure elements such as hydroxyl hydrogen atoms, they hardly ever participate in real hydrogen bonds[18,19] It is indicated that fluorine and chloride substructures contained in azole inhibitors of P450 19 interact with the HBD group of Ser478 and so contribute to the strong inhibitory potency of some azole compounds.[20] No predefined feature of Catalyst can describe this polar interaction. Generally, a hydrogen bond constitutes a special case of a polar interaction. HBAs are formed by electron rich substructures and can therefore be characterized as Lewis bases. Fluorine and chloride also exhibit Lewis bases attributes. To comprise all known interactions between Ser478 and P450 19 inhibitors, we added fluorine and chlorine to the standard HBA feature of catalyst and defined the new feature "HBA+F+Cl". Together with this new feature, H, hydrophobic aromatic (Har), and RA features were considered for P450 19 inhibition modeling.

Generation of a Druglike Virtual Library. The Java-based software ilib diverse is a tool for the fast generation of both diverse and focused virtual libraries.[21,22] The user can select fragments from different chemical and functional groups, which are then connected randomly by Monte Carlo-based methods, which ensures broad coverage of molecular diversity. User-adjustable parameters include the number of used fragments, the definition of atom reactivities, the favored MW range, the minimum and maximum repetition of a particular fragment, the obligatory occurrence of fragments of a special group, import of new fragments, additional chemical modifications and stereochemistry, and special filters for druglike and leadlike properties as well as the desired number of generated molecules. The user can select from two different modes of library generation: In the random generation mode, ilib diverse selects fragments of all selected groups randomly, applying the selected weight values for each fragment which provides the highest diversity at minimal setup expenditure. In the ordered groups mode, several so-called flasks containing user-adaptable fragments or fragment groups can be defined. The program now selects one fragment from each flask and successively connects the fragments. With the additional possibility of defining higher or lower reactivities at different atoms of these fragments, the user can control at which sites the new connections are made, thus being able to generate highly specified compound libraries. Two recent application examples of ilib diverse are given in refs 23 and 24. Different definitions of drug- and leadlikeness have been given in the literature, most often based on physical properties such as molecular weight and logP and on the number and type of allowed atoms or functional groups. Furthermore, neuronal networks and kernel-based learning methods have been applied to distinguish between drugs and nondrugs.[25−27] Ilib diverse addresses this question by providing a set of predefined filters based on different published procedures as well as the possibility of defining your own constraints based on molecular weight distribution, flexibility, size, lipophilicity, reactivity, chemical features, and the number of occurrences of oxygen, nitrogen,

Table 1. Search Patterns that Were Used for Filtering the Virtual Library of 20 000 Compounds Generated by ilib Diverse and the Number of Compounds Mapped by These Patterns

search pattern	no. of hits
has no bond between C and O, N, or S	82
Michael acceptor (only conjugated C═C bonds, no heteroatom at β carbon)	976
β-heterosubstituted carbonyl compound with H at α carbon	2128
1,2-dicarbonyl and derivatives in chain	296
α-halocarbonyl and derivatives	130
acetyl and trifluoroacetyl esters	222
allylic and benzylic halides	154
enol-form of aldehydes	2
enamines (not in ring, not in conjugated double bond)	278
enol and thioenol ethers (not in conjugated double bond)	4
acetals, hemiacetals and similar (hetero-C-hetero not in ring which is annelated to an aromatic ring; does not filter x,x-dihalogenalkanes)	2700
β-carbonyl carboxylic acids and derivatives (esters, but not amides and 1,3-diketones)	353
haloamidines	5
imidoacid and some derivatives, not in ring (also isoureas, but no amidines and guanidines)	1
ketenacetals and derivatives, central structure not in ring	9
orthocarbonic acid derivatives and trihalomethyl-hetero except for trifluoromethyl-hetero	0
cyanides connected to anything but carbon	26
cyanides connected to carbonyl and similar	1
carbonic acid derivatives except ureas and guanidines, not in ring	211
1,2,3-, 1,2,4-, or 1,3,5-trihydroxyaryl	0
anilines with two H (not heterocyclic NH_2 groups)	75
arylhalides *o* or *p* to nitro	0
arylhalides *o* or *p* to aromatic nitrogen in six-membered rings	54
N,P,S-mustards	190
nitrogen-N/O/S single bond (bond is not in ring, includes oximes and hydrazones, no charges as in nitro)	420
nitro on anything but aromatic C	3
p-nitrophenylesters and derivatives	0
o-, *m*-, or *p*-dinitro	0
nontautomerizable thiols	82
highly conjugated, nonaromatic systems (four or more conjugated bonds, may be in ring, but not aromatic)	1
aliphatic chain, unbranched, seven or more atom long	282
allenes, isocyanates, and similar	5
tripeptides or longer peptides	0
4 or more halogen on one ring	1
total compounds filtered	7215

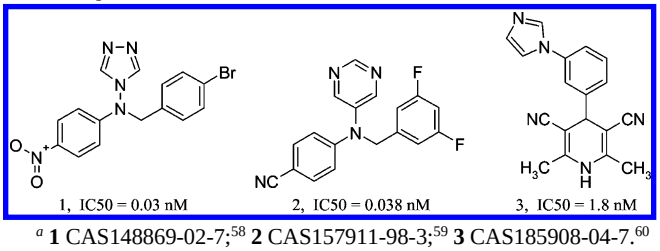
Table 2. Classification of P450 19 Inhibitors Depending on Their Mode of Action and Chemical Structure[57]

classification	characteristics	examples
type I	competitive inhibitors steroidal structure	7α-APTA 7α-phenylethylandrostenedione
type II	competitive inhibitors nonsteroidal structure	aminoglutethimide anastrozole
inactivators	mechanism-based irreversible inactivators	exemestane formestane

and halogens. The predefined set includes filters for high druglikeness,[28−32] for orally bioavailable drugs,[33−37] for blood brain barrier permeable drugs,[38,39] and for leadlikeness.[40] Furthermore, a default filter is provided, which requires a molecular weight of more than 100, a relative topological surface area value between 0 and 10, and which applies some basic reactivity filters (acyl halides; anhydrides; acetals, hemiacetals, and derivatives which are not part of a ring; formyl groups such as aldehydes and formic acid derivatives; halogen connected to anything but carbon).

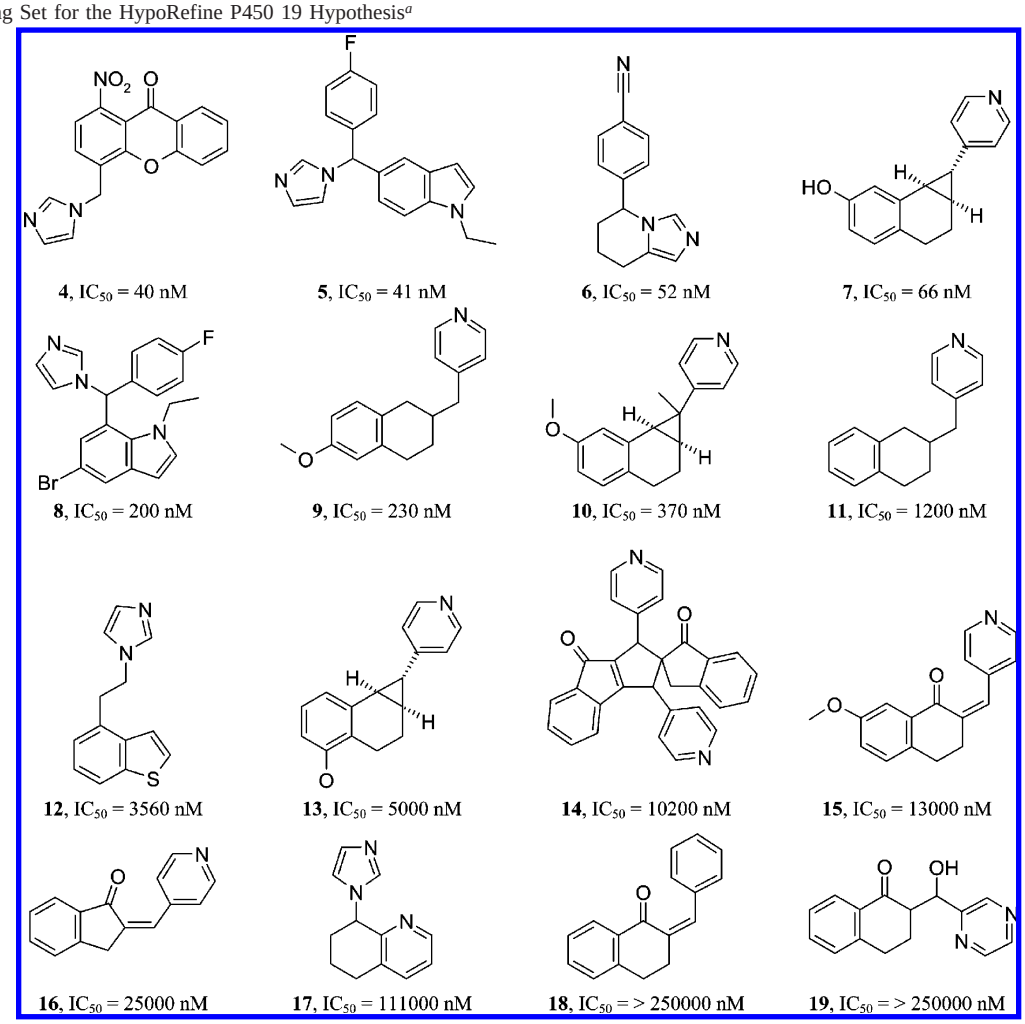
In our study, we used ilib diverse to generate a virtual 3D molecular structure library which should show (a) large diversity and (b) high druglikeness. To achieve these aims, we used the following procedure: In ilib diverse, the random generation mode and the predefined standard fragment set were used to generate two sets—each set consisting of 10 000 compounds. The compounds of the two sets were built from three and four fragments, respectively. We chose to use the default filter setting for the generation of our library and to do some advanced filtering for possibly reactive or labile groups afterward. Thus, the two sets of 10 000 compounds were generated within ilib diverse and exported as SD files. 3D structures were generated by the 3D structure generator Corina,[41] which is optionally bundled with ilib diverse. The two SD files were combined into one file, which was then filtered with the help of obgrep, a command line tool that is part of Open Babel (current version 1.100.2), a free, open-source, cross-platform program and library designed to interconvert between many file formats used in molecular

Chart 1. Nonsteroidal P450 19 Inhibitors Used for HipHop Pharmacophore Model Generation[a]



[a] **1** CAS148869-02-7;[58] **2** CAS157911-98-3;[59] **3** CAS185908-04-7.[60]

Chart 2. Training Set for the HypoRefine P450 19 Hypothesis[a]



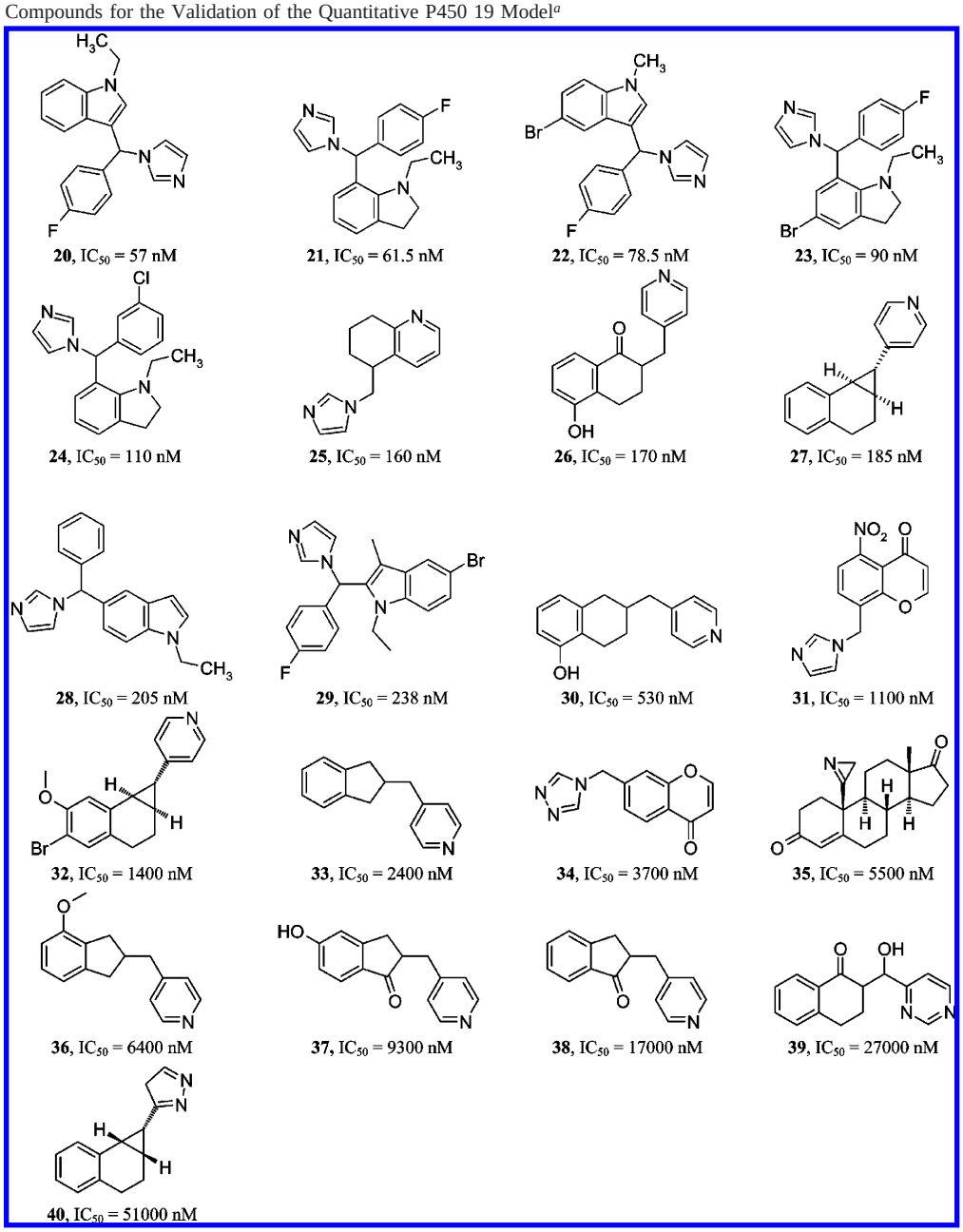
[a] **4** CAS331684-05-0;[67] **5** CAS222306-60-7;[68] **6** fadrozole;[67] **7** CAS165686-07-7;[69] **8** CAS405274-56-8;[70] **9** CAS154932-78-2;[71] **10** CAS165686-02-2;[69] **11** CAS154932-76-0;[71] **12** CAS273930-00-0;[72] **13** CAS165686-05-5;[69] **14** CAS749875-35-2;[73] **15** CAS135075-25-1;[74] **16** CAS154932-62-4;[71] **17** CAS273930-07-7;[72] **18** CAS57558-64-2;[75] **19** CAS173988-83-5.[75]

Table 3. Hitlist Derived from a Stockroom Database Search with the Hypothesis AroHipHopB.01+148869-02-7Shape Employing the BestFlexible Search Algorithm

hit compound	IC_{50} [nM]	BestFit value
CAS147223-54-9[62]	190	2.38
22	0.03	2.84
23	0.038	3.91
CAS204714-56-7[62]	180	3.31
CAS207288-27-5[63]	190	1.92
41	57	1.91
43	78.5	1.94
26	40	2.00
CAS288848-75-9[64]	650	3.51
25	40	0.22
42	61.5	0.80
letrozole[65]	11.5	2.96
vorozole[66]	2.6	2.97

modeling and computational chemistry.[42] Obgrep can read compounds from a variety of common chemical file formats and match them with user-defined SMARTS patterns. Being basically an extension of the SMILES language, SMARTS allows the user to specify flexible and efficient substructure-search specifications.[43] Before filtering, hydrogen atoms were removed from the input file with the "-d" option of Open Babel, because SMARTS distinguishes between implicit and explicit hydrogen atoms. Since obgrep can process only one SMARTS pattern at a time, we wrote a small Perl program which does the following: (i) take search patterns from a list, which contains each search pattern together with a short description on separate lines, and (ii) successively search the given SD file with each pattern from the list by running obgrep and write out the names of the filtered molecules as well as the reason for filtering to a new list. (iii) Using this list, compounds from the input SD file are divided between two files: one file contains the unfiltered compounds and the other one contains the compounds that were filtered, together with a new data field describing the reason each compound was filtered. The patterns that we used to avoid lability, reactivity, and toxic, mutagenic, teratogenic, or cancerogenic effects—both directly or via an unfavorable metabolism—were based on substructure patterns previously applied in the literature,[40,44,45] by chemical distributors for their in-house filtering,[46] or by other software publishers[47] as well as on our own chemical and toxicological knowledge. Only those patterns were used, which seemed possible to occur in our library based on the ilib diverse standard fragment set. A description of each applied filter as well as the number of compounds filtered by this pattern is given in Table 1. On a Pentium IV with 3.0 GHz and 1 GB RAM running Redhat Linux 9.0, matching 20 000 compounds against 38 patterns took 14 min. Of 20 000 compounds, 7215 (36%) were sorted out. With the Catalyst module catDB,

Chart 3. Test Set Compounds for the Validation of the Quantitative P450 19 Model[a]



[a] **20** CAS222306-51-6;[68] **21** CAS405274-48-8;[70] **22** CAS222306-52-7;[68] **23** CAS405274-49-9;[70] **24** CAS405274-45-5;[70] **25** CAS273930-11-3;[72] **26** CAS135075-38-6;[74] **27** CAS158467-40-4;[69] **28** CAS222306-59-4;[68] **29** CAS222306-41-4;[68] **30** CAS154932-80-6;[71] **31** CAS331683-99-9;[67] **32** CAS165686-04-4;[69] **33** CAS154932-71-5;[71] **34** CAS331684-17-4;[67] **35** CAS174914-71-7;[76] **36** CAS154932-72-6;[71] **37** CAS154932-70-4;[71] **38** CAS4803-61-6;[71] **39** CAS173988-86-8;[75] **40** CAS173988-94-8.[75]

we generated a multiconformational 3D database of the remaining 12 785 compounds, applying the FAST algorithm and a maximum of 100 conformations per molecule. Finally, the database was checked for duplicates. This had to be done in such a late stage, because neither ilib diverse nor Open Babel support filtering for duplicates in their current releases. Basically, a good method to compare compounds to see whether they are identical is via so-called canonical SMILES, i.e., a method of conversion into SMILES strings that, regardless of the atom order and structure of the input molecules, always gives identical SMILES strings for identical compounds.[48] The major drawback of canonical SMILES is that—although its concept has been implemented in some computer programs—there are no general standard rules available yet, leaving it to programmers to generate their own canonical SMILES. The recently released IUPAC International Chemical Identifier (InChI),[49−51] providing a publicly available algorithm for structure canonicalization, might help overcome this problem. Catalyst also offers the possibility to export the contents of a database as a SMILES list, and although it uses some extensions to the SMILES language which are unreadable by many other programs, this code is—to the best of our knowledge—canonical. We thus generated a SMILES list for all compounds in the database, sorted it alphabetically, so that identical SMILES strings appeared in consecutive lines, and then filtered this list with a small Perl script by comparing each pair of consecutive lines for identical SMILES strings. With this method, we could find 10 duplicate structures. This small number (0.08% of 12 785 compounds) proves the power of ilib diverse to

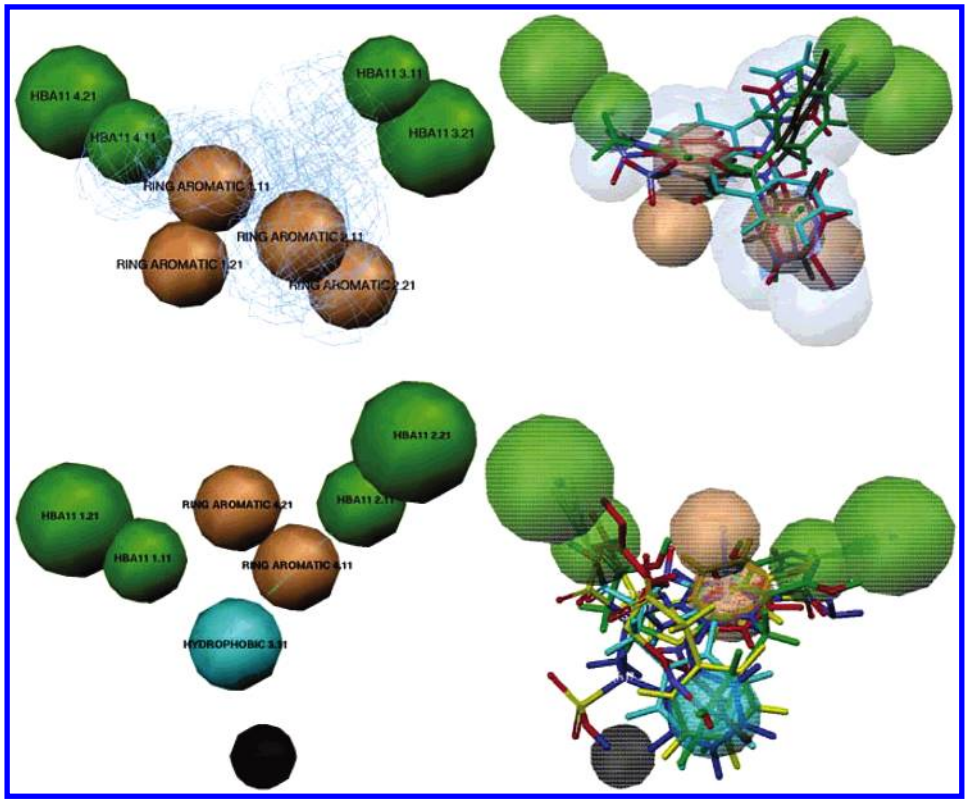


Figure 2. The HipHop hypothesis AroHH-B.01+148869-02-7Shape (left) consists of two RA features (brown), two HBA+F+Cl features (green), and the shape of **1** merged into the hypothesis (blue). **1** (red), **2** (dark gray), vorozole (green), and **4** (blue) are mapped into AroHH-B.01+148869-02-7Shape. The HypoRefine hypothesis AroHG-D.8 is composed of one RA feature, one H feature (cyan), two HBA+F+Cl features, and one XVOL (black). The NCI database search hits 288 730 (red), 92 674 (green), 6320 (dark blue), 401 336 (yellow), and 245 220 (cyan) are overlaid in AroHG-D.8.

generate large collections of diverse compounds, even with a small number of fragments and with the default settings where some groups and fragments are deselected. The 10 duplicate compounds were removed from our database, resulting in a virtual 3D database of 12 775 unique, diverse, and druglike compounds.

Biological Assays for P450 17 and P 450 19 Inhibition. As source of the enzymes the following microsomal preparations were used: for P450 19, human placenta[52,53] and for P450 17, *E. coli* expressing human P450 17.[54] The P450 19 assay was performed as described using the ^{3}H2O-method: [1β-^{3}H]androstenedione/androstenedione (0.5 μm)[55] was used as substrate. The P450 17 assay was performed with nonlabeled progesterone, and an HPLC-procedure was employed for the separation of the substrate and androstenedione using UV-detection.[54,55]

RESULTS AND DISCUSSION

In the literature, one can find several classifications on P450 19 inhibitors, dependent on market launch[56] or mode of interaction with the enzyme.[57] For modeling purposes, interaction properties play crucial roles, so the second classification was considered. Three chemical classes of P450 19 inhibitors are distinguished (Table 2): competitive steroidal inhibitors, competitive nonsteroidal inhibitors, and irreversible inactivators.

Today, type I P450 19 inhibitors play only a minor role in the search for new breast cancer therapies. For this class, no model was created.

Table 4. Validation of AroHG-D.8 with the Test Set[a]

compound	activity [nM]	estimated activity	error[b]
20	57	520	−9.2
21	61.5	110	1.8
22	78.5	250	3.1
23	90	99	1.1
24	110	310	2.8
25	160	9300	58
26	170	960	5.7
27	185	9300	50
28	205	1700	8.5
29	238	250	1.1
30	530	740	1.4
31	1100	350	−3.1
32	1400	9400	6.7
33	2400	9200	3.8
34	3700	400	−9.1
35	5500	9300	1.7
36	6400	2800	−2.3
37	9300	1400	−6.5
38	17000	11000	−1.6
39	27000	9200	−2.9
40	51000	13000	−3.9

[a] Activities are given in nM. Nineteen out of 21 (90%) P450 19 inhibitors were predicted correctly within one log unit of activity. [b] The error factor is computed as the ratio of the measured activity to the activity estimated by the hypothesis or the inverse if estimated is greater than measured.

Inactivators such as exemestane take an important position among breast cancer drugs; however, there are comparably few data available on P450 19 inactivators. Therefore, the generation and validation of a pharmacophore model would have proved difficult, so we disregarded this class as well.

Table 5. Screening of the Hitlists against Antitarget Pharmacophore Models as Well as the Rule of Five

filter criteria	no. of filtered compounds
hERG block	90
potent PXR activation	0
rule of five − violation	226

Some type II inhibitors are already used in the clinic (e.g. anastrozole). Regularly, these new classes of nonsteroidal P450 19 inhibitors are reported in the literature. Here we introduce both a qualitative (HipHop) and a quantitative (HypoRefine) model for nonsteroidal P450 19 inhibitors.

Common Feature Pharmacophore Model for P450 19 Inhibitors. Data on P450 19 inhibitors and their selectivity toward other enzymes involved in sex hormone biosynthesis are abundant. Reviews on nonsteroidal compounds appear regularly; the most recent was published by Recanatini et al.[57] Starting from a set of 17 highly active P450 19 inhibitors (Chart 1), a training set consisting of the three most potent compounds (**1**−**3**) was submitted to the hypothesis generation process.

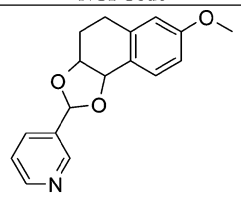
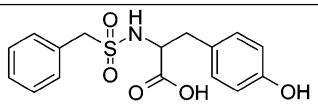
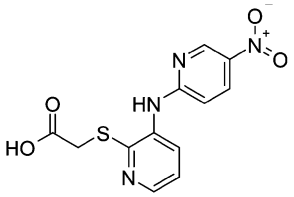
Most active P450 19 inhibitors show aromatic structures; therefore, the Har and RA features were considered for pharmacophore generation in addition to the H and the HBA+F+Cl features. The hypothesis generation process returned 10 hypotheses of which the hypothesis that identified most P450 19 inhibitors as actives and which was ranked best by Catalyst was chosen for further work. The hypothesis—AroHipHopB.01—consisted of two RA features and two HBA+F+Cl features. When performing a database search of our in-house library of 57 P450 19 inhibitors of different activities (the stockroom database, consisting of 24 P450 19 inhibitors with IC_{50} values < 200 nM, 9 inhibitors with IC_{50} values from 200 to 1.000 nM, 20 inhibitors with IC_{50} values from 1000 to 111.000 nM, and three inactive compounds), 20 compounds (35.1%) were identified correctly as active, 16 of which show an IC_{50} value < 200 nM. Screening of the World Drug Index Database (WDI)[61] which consists of 63 307 compounds that are currently under clinical investigation or marketed as drugs returned 14.799 hits (23.4%). Of the 31 nonsteroidal, competitive aromatase inhibitors included in the WDI database 16 (52%) were correctly identified by our model.

To increase model selectivity, **1** was fitted into model AroHipHopB.01, converted into a shape query, and merged with the preliminary features to the more sophisticated hypothesis AroHipHopB.01+148869-02-7Shape. A search of the stockroom database returned 13 hits (22.8%), all of them with an IC_{50} value in the nanomolar range (Table 3) and including 12 compounds with an IC_{50} value of < 200 nM. A WDI search returned 1.729 hits (2.7%). Of the 31 aromatase inhibitors included in the WDI database mentioned above, 12 (39%) were correctly identified by our model. A search in the ilib diverse virtual database returned 3.0% hits while correctly identifying all 10 active compounds. Due to the good enrichment of P450 19 inhibitors in the hitlist, the hypothesis AroHipHopB.01+148869-02-7Shape was employed for database mining.

Quantitative Pharmacophore Model for P450 19 Inhibitors. A HypoRefine pharmacophore model was derived from a training set of 16 compounds with in vitro activities from 40 nM to >250 000 nM (Chart 2). All these data had been determined in the same laboratory using equal assays (see Materials and Methods).

Regarding the relatively narrow spectrum of activity values represented in this training set (3.5 log units), the uncertainty values of all compounds except for the two least active ones were set to 2. The activity value of the two least active compounds was set to 300 000, and their uncertainty value was set to 3. The hypothesis generation settings demanded one or two HBA+F+Cl features, a maximum of four H

Chart 4. Inhibition of P450 19 and P450 17 by Compounds Retrieved from the NCI Database Search and Processed through the Filter

Compound Structure NCI Code	Inhibition on P450 19[a)]	Inhibition on P450 17[b)]	Retrieved by Model for
 378898	68.2 % inhibition at 36 μM	no inhibition at 2.5 μM	P450 19
 401336	no inhibition at 36 μM	no inhibition at 2.5 μM	P450 19
 288730	15 % inhibition at 36 μM	no inhibition at 2.5 μM	P450 19

[a] Substrate: 1β ^{3}H- androstendione/androstendione 500 nM. [b]Substrate: progesterone 25 μM enzyme: bacterial membranes containing recombinantly expressed human P450 17 or human placenta microsomes containing P450 19 reference compounds: BW19 IC_{50}: 0.15 μM, ketoconazole IC_{50}: 4.5 μM (P450 17); aminoglutethimide IC_{50}: 29.75 μM (P450 19).

features, at most two Har features, at most two RA features, and at most 20 XVOLs.

According to the cost analysis of the 10 returned hypotheses, each of them had a delta cost (the difference between null cost and total cost of the respective hypothesis) of 40−80 bits and was therefore relevant. The evaluation of a test set (Chart 3 and Table 4) consisting of 21 compounds revealed that the hypothesis AroHG-D.08 was most suitable to predict the IC_{50} of new compounds within one log unit of error. All test set activities have been determined in the same laboratory as the training set compounds, using comparable biological assays.

Figure 2 shows both the HipHop and the HypoRefine hypotheses for P450 19 inhibitors.

Selection of Test Compounds From the National Cancer Institute (NCI) Database. The hypothesis AroHipHopB.01+148869-02-7Shape was employed to search the database of the NCI (comprising 123 219 compounds) for new P450 19 inhibitors. By employing the FastFlexible Search algorithm, 938 hits (0.76%) were retrieved by the model. These hits were then estimated in their activity by the HypoRefine hypothesis AroHG-D.8 and ranked according to their predicted potency. Subsequently, the hits underwent additional filtering to exclude compounds with unfavorable properties: The hits were screened against a combined pharmacophore model to predict hERG potassium channel block[77] as well as a model for highly active pregnane X receptor (PXR) activators.[78] Positive hits from these searches were excluded from the hitlists. Additionally, hits that did not satisfy the criteria for favorable absorption and permeation properties—the famous "rule of five":[34] molecular weight ≤ 500, ≤ 5 hydrogen bond donors, ≤ 10 hydrogen bond acceptors, cLogP ≤ 5—were also disregarded. The number of filtered compounds are listed in Table 5.

Unfavorable chemical classes such as esters and glycosides that are unlikely to be stable in human plasma were eliminated as well. Finally, we ran an experience-based toxicity prediction program (Osiris Property Explorer from www.organic-chemistry-org/prog/peo/index.html) over the remaining top-scored hits. From the final hitlists we selected five compounds for biological testing. The test results are shown in Chart 4.

Two of the three hits derived from the P450 19 model showed selective activity in the micromolar range.

CONCLUSION

In this work we present a common feature as well as a quantitative pharmacophore model for P450 19 inhibitors. The HipHop model for P450 19 inhibitors consists of two RA features in the center flanked by a HBA group each at both ends. The quantitative (HypoRefine) model was able to predict 90% of all tested compounds within an error of one activity log unit. Using the HipHop model as search queries and employing a sophisticated filtering procedure to increase the druglikeness of the test compounds, we discovered two new P450 19 inhibitors. However, the number of compounds tested was very small. Further biological testing of hits would be necessary to definitely determine the hit rate of this procedure and subsequently further optimize it.

ACKNOWLEDGMENT

We thank the National Cancer Institute for providing the test compounds free of charge, Dr. Rémy D. Hoffman, Accelrys SARL Paris, for conducting the searches in the WDI, and Inte:Ligand GmbH for funding and providing the ilib diverse software.

REFERENCES AND NOTES

(1) Neumann, F.; Schenck, B.; Schleusener, H.; Schweikert, H. U. Endocrinology. *General and specific aspects of pharmacology and toxicology*; Spektrum Verlag: München, 1992.
(2) Hartmann, R. W.; Ehmer, P. B.; Haidar, S.; Hector, M.; Jose, J. et al. Inhibition of CYP 17, a New Strategy for the Treatment of Prostate Cancer. *Arch. Pharm. (Weinheim)* **2002**, *335*, 119−128.
(3) Forbes, J. F. The control of breast cancer: the role of tamoxifen. *Semin. Oncol.* **1997**, *24*, S1−5-S1−19.
(4) Miller, W. R. Aromatase inhibitors: mechanism of action and role in the treatment of breast cancer. *Semin. Oncol.* **2003**, *30*, 3−11.
(5) Beatson, G. T. On the Treatment of Inoperable Cases of Carcinoma of the Mamma. Suggestion for a New Method of Treatment, with Illustrative Cases. *Lancet* **1896**, *2*, 104−107.
(6) Pasqualini, J. R. The selective estrogen enzyme modulators in breast cancer: a review. *Biochim. Biophys. Acta* **2004**, *1654*, 123−143.
(7) Nagy, P. I.; Tokarski, J.; Hopfinger, A. J. Molecular shape and QSAR analyses of a family of substituted dichlorodiphenyl aromatase inhibitors. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1190−1197.
(8) Auvray, P.; Moslemi, S.; Sourdaine, P.; Galopin, S.; Seralini, G.-E. et al. Evidence for new nonsteroidal human aromatase inhibitors and comparison with equine aromatase inhibition for an understanding of the mammalian active site. *Eur. J. Med. Chem.* **1998**, *33*, 451−462.
(9) Cavalli, A.; Greco, G.; Novellino, E.; Recanatini, M. Linking CoMFA and protein homology models of enzyme-inhibitor interactions: an application to nonsteroidal aromatase inhibitors. *Bioorg. Med. Chem.* **2000**, *8*, 2771−2780.
(10) Cavalli, A.; Recanatini, M. Looking for Selectivity among Cytochrome P450s Inhibitors. *J. Med. Chem.* **2002**, *45*, 251−254.
(11) Auchus, R. J.; Miller, W. L. Molecular modeling of human P450c17 (17a-hydroxylase/17,20-lyase): insights into reaction mechanisms and effects of mutations. *Mol. Endocrinol.* **1999**, *13*, 1169−1182.
(12) Clement, O. O.; Mehl, A. T. HipHop: Pharmacophores Based on Multiple Common-Feature Alignments. *Pharmacophore Perception, Development, and Use in Drug Design*; International University Line: La Jolla, CA, 2000; pp 69−84.
(13) Catalyst Catalyst software package, Tutorial Release 4.9, 2003.
(14) Li, H.; Sutter, J.; Hoffmann, R. HypoGen: An Automated System for Generating 3D Predictive Pharmacophore Models. *Pharmacophore Perception, Development, and Use in Drug Design*; International University Line: La Jolla, CA, 2000; pp 172−189.
(15) Catalyst / HypoRefine; Accelrys, 2004. www.accelrys.com
(16) Kurogi, Y.; Guner, O. F. Pharmacophore modeling and three-dimensional database searching for drug design using catalyst. *Curr. Med. Chem.* **2001**, *8*, 1035−1055.
(17) Sutter, J.; Güner, O.; Hoffmann, R.; Li, H.; Waldman, M. Effect of Variable Weights and Tolerances on Predictive Model Generation. *Pharmacophore Perception, Development, and Use in Drug Design*; International University Line: La Jolla, CA, 2000; pp 499−511.
(18) Dunitz, J. D. Organic fluorine: Odd man out. *ChemBioChem* **2004**, *5*, 614−621.
(19) Boehm, H.-J.; Banner, D.; Bendels, S.; Kansy, M.; Kuhn, B.et al. Fluorine in medicinal chemistry. *ChemBioChem* **2004**, *5*, 637−643.
(20) Ahmed, S. The use of the novel substrate-heme complex approach in the derivation of a representation of the active site of the enzyme complex 17a-hydroxylase and 17,20-lyase. *Biochem. Biophys. Res. Commun.* **2004**, *316*, 595−598.
(21) Wolber, G.; Langer, T. Comb'Gen: a novel software package for the rapid generation of virtual combinatorial libraries. *Rational Approaches to Drug Design, Proceedings of the European Symposium on Quantitative Structure−Activity Relationships, 13th, Duesseldorf, Germany, Aug. 27−Sept. 1, 2000*, 2001; pp 390−399.
(22) *ilib:diverse*; Inte:Ligand GmbH. www.inteligand.com
(23) Krovat, E. M.; Fruehwirth, K. H.; Langer, T. Pharmacophore Identification, in Silico Screening, and Virtual Library Design for Inhibitors of the Human Factor Xa. *J. Chem. Inf. Comput. Sci.* **2005**, *45*, 146−159.
(24) Steindl, T.; Langer, T. Docking Versus Pharmacophore Model Generation: A Comparison of High-Throughput Virtual Screening Strategies for the Search of Human Rhinovirus Coat Protein Inhibitors. *QSAR Comb. Sci.* **2005**, *24*, 470−479.

(25) Sadowski, J.; Kubinyi, H. A scoring scheme for discriminating between drugs and nondrugs. *J. Med. Chem.* **1998**, *41*, 3325−3329.
(26) Byvatov, E.; Fechner, U.; Sadowski, J.; Schneider, G. Comparison of Support Vector Machine and Artificial Neural Network Systems for Drug/Nondrug Classification. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1882−1889.
(27) Mueller, K.-R.; Raetsch, G.; Sonnenburg, S.; Mika, S.; Grimm, M. et al. Classifying 'Drug-likeness' with Kernel-Based Learning Methods. *J. Chem. Inf. Model.* **2005**, *45*, 249−253.
(28) Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. A Knowledge-Based Approach in Designing Combinatorial or Medicinal Chemistry Libraries for Drug Discovery. 1. A Qualitative and Quantitative Characterization of Known Drug Databases. *J. Comb. Chem.* **1999**, *1*, 55−68.
(29) Lee, M.-L.; Schneider, G. Scaffold architecture and pharmacophoric properties of natural products and trade drugs: application in the design of natural product-based combinatorial libraries. *J. Comb. Chem.* **2001**, *3*, 284−289.
(30) Mozziconacci, J. C.; Arnoult, E.; Baurin, N.; Marot, C.; Morin-Allory, L. Preparation of a Molecular Database from a Set of 2 Million Compounds for Virtual Screening Applications: Gathering, Structural Analysis and Filtering. www.univ-orleans.fr/SCIENCES/ICOA/eposter/eccc9/ECCC9i.htm.
(31) Oprea, T. I.; Davis, A. M.; Teague, S. J.; Leeson, P. D. Is There a Difference between Leads and Drugs? A Historical Perspective. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1308−1315.
(32) Walters, W. P.; Murcko, M. A. Library filtering systems and prediction of drug-like properties. *Virtual Screening for Bioactive Molecules*; Wiley: New York, 2000; pp 15−32.
(33) Fichert, T.; Yazdanian, M.; Proudfoot, J. R. A structure-Permeability study of small drug-like molecules. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 719−722.
(34) 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 Deliv. Rev.* **1997**, *23*, 3−25.
(35) Palm, K.; Luthman, K.; Ungell, A.-L.; Strandlund, G.; Artursson, P. Correlation of Drug Absorption with Molecular Surface Properties. *J. Pharm. Sci.* **1996**, *85*, 32−39.
(36) Palm, K.; Stenberg, P.; Luthman, K.; Artursson, P. Polar molecular surface properties predict the intestinal absorption of drugs in humans. *Pharm. Res.* **1997**, *14*, 568−571.
(37) Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W. et al. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **2002**, *45*, 2615−2623.
(38) Darvas, F.; Keseru, G.; Papp, A.; Dorman, G.; Urge, L. et al. In silico and ex silico ADME approaches for drug discovery. *Curr. Top. Med. Chem.* **2002**, *2*, 1287−1304.
(39) Van de Waterbeemd, H.; Camenisch, G.; Folkers, G.; Chretien, J. R.; Raevsky, O. A. Estimation of blood-brain barrier crossing of drugs using molecular size and shape, and H-bonding descriptors. *J. Drug. Targeting* **1998**, *6*, 151−165.
(40) Oprea, T. I. Property distribution of drug-related chemical databases. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 251−264.
(41) *Corina Version 3.0*; 3.0 ed.
(42) *Open Babel version 1.100.2*.
(43) Systems, D. C. I. SMARTS tutorial by Daylight Chemical Information Systems Inc.
(44) Hann, M.; Hudson, B.; Lewell, X.; Lifely, R.; Miller, L. et al. Strategic Pooling of Compounds for High-Throughput Screening. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 897−902.
(45) Sawatari, K.; Nakanishi, Y.; Matsushima, T. Relationships between chemical structures and mutagenicity: a preliminary survey for a database of mutagenicity test results of new work place chemicals. *Ind. Health* **2001**, *39*, 341−345.
(46) Maybridge Exclusion criteria of the Maybridge Screening Collection database.
(47) MOE Manual, 2004.
(48) Weininger, D.; Weininger, A.; Weininger, J. L. SMILES. 2. Algorithm for generation of unique SMILES notation. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 97−101.
(49) Prasanna, M. D.; Vondrasek, J.; Wlodawer, A.; Bhat, T. N. Application of InChI to curate, index, and query 3-D structures. *Proteins: Struct., Funct., Bioinformat.* **2005**, *60*, 1−4.
(50) Coles, S. J.; Day, N. E.; Murray-Rust, P.; Rzepa, H. S.; Zhang, Y. Enhancement of the chemical semantic web through the use of InChI identifiers. *Org. Biomol. Chem.* **2005**, *3*, 1832−1834.
(51) IUPAC The IUPAC International Chemical Identifier (InChI) - official homepage.
(52) Thompson, E. A., Jr.; Siiteri, P. K. Utilization of oxygen and reduced nicotinamide adenine dinucleotide phosphate by human placental microsomes during aromatization of androstenedione. *J. Biol. Chem.* **1974**, *249*, 5364−5372.
(53) Hartmann, R. W.; Batzl, C. Aromatase inhibitors. Synthesis and evaluation of mammary tumor inhibiting activity of 3-alkylated 3-(4-aminophenyl)piperidine-2,6-diones. *J. Med. Chem.* **1986**, *29*, 1362−1369.
(54) Hutschenreuter, T. U.; Ehmer, P. B.; Hartmann, R. W. Synthesis of Hydroxy Derivatives of Highly Potent Non-steroidal CYP 17 Inhibitors as Potential Metabolites and Evaluation of their Activity by a Non Cellular Assay using Recombinant Human Enzyme. *J. Enzyme Inhib. Med. Chem.* **2004**, *19*, 17−32.
(55) Wachall, B. G.; Hector, M.; Zhuang, Y.; Hartmann, R. W. Imidazole substituted biphenyls: A new class of highly potent and in vivo active inhibitors of P450 17 as potential therapeutics for treatment of prostate cancer. *Bioorg. Med. Chem.* **1999**, *7*, 1913−1924.
(56) Sainsbury, R. Aromatase inhibition in the treatment of advanced breast cancer: is there a relationship between potency and clinical efficacy? *Br. J. Cancer* **2004**, *90*, 1733−1739.
(57) Recanatini, M.; Cavalli, A.; Valenti, P. Nonsteroidal aromatase inhibitors: recent advances. *Med. Res. Rev.* **2002**, *22*, 282−304.
(58) Okada, M.; Kawaminami, E.; Yoden, T.; Kudo, M.; Isomura, Y. Preparation of triazolyl-substituted tertiary amines as aromatase inhibitors. In *PCT Int. Appl.*; Yamanouchi Pharmaceutical Co., Ltd.: Japan, Wo, 1993; pp 113 pp.
(59) Okada, M.; Yoden, T.; Kawaminami, E.; Shimada, Y.; Kudoh, M.et al. Studies on aromatase inhibitors IV. Synthesis and biological evaluation of N, N-Disubstituted-5-aminopyrimidine derivatives. *Chem. Pharm. Bull.* **1997**, *45*, 1293−1299.
(60) Cozzi, P.; Briatico, G.; Giudici, D.; Rossi, A.; Di Salle, E. Imidazol-1-yl and 1,2,4-triazol-1-yl derivatives of 3-cyano-4-phenyl-1,4-dihydropyridines and analogs as potent aromatase inhibitors. *Med. Chem. Res.* **1996**, *6*, 611−617.
(61) *Derwent World Drug Index*; 2003 ed.; Derwent Publications Ltd.: London.
(62) Karjalainen, A.; Kalapudas, A.; Sodervall, M.; Pelkonen, O.; Lammintausta, R. Synthesis of new potent and selective aromatase inhibitors based on long-chained diarylalkylimidazole and diarylalkyltriazole molecule skeletons. *Eur. J. Pharm. Sci.* **2000**, *11*, 109−131.
(63) Vinh, T. K.; Ahmadi, M.; Delgado, P. O. L.; Perez, S. F.; Walters, H. M. et al. 1-[(benzofuran-2-yl)phenylmethyl]-triazoles and -tetrazoles - potent competitive inhibitors of aromatase. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2105−2108.
(64) Sonnet, P.; Dallemagne, P.; Guillon, J.; Enguehard, C.; Stiebing, S. et al. New aromatase inhibitors. Synthesis and biological activity of aryl-substituted pyrrolizine and indolizine derivatives. *Bioorg. Med. Chem.* **2000**, *8*, 945−955.
(65) Bhatnagar, A. S.; Haeusler, A.; Schieweck, K.; Lang, M.; Bowman, R. Highly selective inhibition of estrogen biosynthesis by CGS 20267, a new non-steroidal aromatase inhibitor. *J. Steroid Biochem. Mol. Biol.* **1990**, *37*, 1021−1027.
(66) Vanden Bossche, H.; Willemsens, G.; Roels, I.; Bellens, D.; Moereels, H. et al. R 76713 and enantiomers: selective, nonsteroidal inhibitors of the cytochrome P450-dependent estrogen synthesis. *Biochem. Pharmacol.* **1990**, *40*, 1707−1708.
(67) Recanatini, M.; Bisi, A.; Cavalli, A.; Belluti, F.; Gobbi, S. et al. A New class of nonsteroidal aromatase inhibitors: design and synthesis of chromone and xanthone derivatives and inhibition of the P450 enzymes aromatase and 17a-hydroxylase/C17,20-lyase. *J. Med. Chem.* **2001**, *44*, 672−680.
(68) Le Borgne, M.; Marchand, P.; Delevoye-Seiller, B.; Robert, J.-M.; Le Baut, G. et al. New selective nonsteroidal aromatase inhibitors: synthesis and inhibitory activity of 2-, 3-, or 5-(a-azolylbenzyl)-1H-indoles. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 333−336.
(69) Hartmann, R. W.; Bayer, H.; Gruen, G.; Sergejew, T.; Bartz, U. et al. Pyridyl-Substituted Tetrahydrocyclopropa[a]naphthalenes: Highly Active and Selective Inhibitors of P450 arom. *J. Med. Chem.* **1995**, *38*, 2103−2111.
(70) Marchand, P.; Le Borgne, M.; Palzer, M.; Le Baut, G.; Hartmann, R. W. Preparation and pharmacological profile of 7-(a-Azolylbenzyl)-1H-indoles and indolines as new aromatase inhibitors. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1553−1555.
(71) Hartmann, R. W.; Bayer, H.; Gruen, G. Aromatase Inhibitors. Syntheses and Structure-Activity Studies of Novel Pyridyl-Substituted Indanones, Indans, and Tetralins. *J. Med. Chem.* **1994**, *37*, 1275−1281.
(72) Jacobs, C.; Frotscher, M.; Dannhardt, G.; Hartmann, R. W. 1-imidazolyl(alkyl)-substituted di- and tetrahydroquinolines and analogs: syntheses and evaluation of dual inhibitors of thromboxane A2 synthase and aromatase. *J. Med. Chem.* **2000**, *43*, 1841−1851.

(73) Gupta, R.; Jindal, D. P.; Jit, B.; Narang, G.; Palusczak, A. et al. Synthesis and evaluation of a dimer of 2-(4-pyridylmethyl)-1-indanone as a novel nonsteroidal aromatase inhibitor. *Arch. Pharm. (Weinheim)* **2004**, *337*, 398−401.
(74) Bayer, H.; Batzl, C.; Hartman, R. W.; Mannschreck, A. New aromatase inhibitors. Synthesis and biological activity of pyridyl-substituted tetralone derivatives. *J. Med. Chem.* **1991**, *34*, 2685−2691.
(75) Waechter, G. A.; Hartmann, R. W.; Sergejew, T.; Gruen, G. L.; Ledergerber, D. Tetrahydronaphthalenes: Influence of Heterocyclic Substituents on Inhibition of Steroidogenic Enzymes P450 arom and P450 17. *J. Med. Chem.* **1996**, *39*, 834−841.
(76) Njar, V. C. O.; Duerkop, J.; Hartmann, R. W. Synthesis of 10b-(1′-azirinyl)estr-4-en-3,17-dione as an aromatase inhibitor. *Steroids* **1996**, *61*, 138−143.
(77) Niederreiter, D. In *Why Drugs Fail - A Study on Side Effects in New Chemical Entities*; Department of Pharmaceutical Chemistry, University of Innsbruck: Innsbruck, 2004.
(78) Schuster, D.; Langer, T. The Identification of Ligand Features Essential for PXR Activation by Pharmacophore Modeling. *J. Chem. Inf. Model.* **2005**, *45*, 431−439.

CI050237K