Bioorganic & Medicinal Chemistry 19 (2011) 7186–7193

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Identification of novel CYP2A6 inhibitors by virtual screening

Minna K. Rahnasto *, Hannu A. Raunio, Carsten Wittekindt, Kaisa A. Salminen, Jukka Leppänen, Risto O. Juvonen, Antti Poso, Maija K. Lahtela-Kakkonen

School of Pharmacy, University of Eastern Finland, POB 1627, 70211 Kuopio, Finland

ARTICLE INFO

Article history:
Received 27 May 2011
Revised 1 September 2011
Accepted 28 September 2011
Available online 4 October 2011

Keywords:
P450
CYP2A6
CoMFA
3D-QSAR

ABSTRACT

The human CYP2A6 enzyme metabolises several xenobiotics including nicotine, the addictive component in tobacco. Reduced activity of CYP2A6, either for genetic reasons or by administering inhibitors of CYP2A6, reduces tobacco smoking. The aim was to design novel inhibitors of CYP2A6 using 3D-QSAR analysis combined with virtual screening. A 3D-QSAR model was utilised to identify the most important features of the inhibitors, and this knowledge was used to design inhibitors for CYP2A6. Chemical database screening yielded several potent inhibitor candidates such as alkylamine derivatives (compound no. **5**, IC_{50} = 0.1 μM) and 1-benzothiophene-3-carbaldehyde that can be used as lead compounds in the development of drugs for smoking reduction therapy.

1. Introduction

Tobacco smoking is associated with increased incidence of several types of cancer, respiratory and cardiovascular diseases, gastrointestinal diseases, and many other disorders. Nicotine is extensively metabolised to a number of intermediates through the cytochrome P450 (CYP) enzyme system in the liver. In humans, about 70–80% of nicotine is metabolised by the CYP2A6 enzyme and converted to cotinine.[1–4]

Genetic polymorphism of the CYP2A6 gene can lead to a complete lack or reduced activity of the CYP2A6 enzyme (http://cypalleles.ki.se/cyp2a6.htm). Such polymorphism causes slower nicotine metabolism in individuals carrying nonfunctional alleles, resulting in decreased nicotine plasma levels and changes in smoking behaviour.[5,6] Individuals with less than 50% CYP2A6 activity are about two times less likely to be smokers than people with normal CYP2A6 activity,[7,8] and even if they do smoke, their consumption of cigarettes per day is lower.[7,8]

The extent of smoking can be reduced by mimicking (phenocopying) the slow metabolism of nicotine by oral treatment with CYP2A6 inhibitors. Small pilot studies with some currently available CYP2A6 inhibitors, for example, methoxsalen, have indicated that these agents can reduce smoking.[9,10] However, none of these CYP2A6 inhibitors can be used clinically, because they are overly nonspecific and/or toxic.[11–14]

Most inhibitors bind to their target enzyme through reversible interactions. However, in some cases, enzymes are inactivated by mechanism-based inactivators because they form either irreversible or quasi-irreversible complexes with inhibitory molecules. In irreversible inactivation, inhibitors are catalytically transformed by CYP enzymes into reactive intermediates, which then inactivate the enzyme.[15,16] In quasi-irreversible inactivation, the initial product of the inhibitor coordinates to the Fe(II)-form of the CYP enzyme, resulting in the formation of a metabolite-intermediate complex (MIC).[17,18] Inactivation of the target enzyme by mechanism-based inactivators requires catalysed activation, which causes these inactivators to display slow binding kinetics.

This study is a part of a project with the long-term goal of developing novel, specific, and safe CYP2A6 inhibitors to be used in smoking reduction treatments. The main purpose of this study was to combine molecular modeling and in vitro enzymatic techniques to identify selective and potent CYP2A6 inhibitors. The first step was to construct improved predictive comparative molecular field analysis (CoMFA)[19,20] for CYP2A6 using its crystal structure. In the second step, data from these models were used to carry out virtual screening from a chemical database to search for novel molecules with the desired inhibitory characteristics. The third step involved a series of in vitro enzymatic analyses to understand the mechanism of inhibition of the novel CYP2A6 inhibitors.

2. Computational method

The crucial step in constructing a CoMFA model is the alignment of the training set molecules. The training set of CYP2A6 inhibitors consisted of 85 compounds published previously (see Supplementary data, section A).[12,13,21] Since several CYP2A6 crystal structures complexed with coumarin, methoxsalen and 3-heteroaromatic

* Corresponding author. Tel.: +358 40 355 3786; fax: +358 17 162424.
E-mail address: Minna.Rahnasto@uef.fi (M.K. Rahnasto).

0968-0896/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2011.09.054

pyridines have been determined,[11,12], we applied the virtual docking tool (see 2.2) to fit the inhibitor compounds into the active site of CYP2A6. The alignment of inhibitors obtained from molecular docking was used in the CoMFA study (see 2.3).

The crystal structure of CYP2A6 (pdb: 2FDW) was obtained from the PDB database.[11] The quality of the structure was assessed by the PROCHECK software[23] and was fixed and visually inspected using the Sybyl software (v. 8.1).[24] Solvent molecules were deleted from the structure and hydrogen atoms were added followed by visual inspection for the orientation of residues and tautomeric states of histidine residues in the active site. Finally, the structure was minimised using the Powell software[25] before molecular dynamics (MD) simulations (see 2.1). Before minimisation, the model was solvated using a cubic box including water and counterions (chloride ions). Energy minimisation of the structures was performed by an implicit solvent model with using CHARMM[26] force field. During the first minimisation step, the backbone and side chains were constrained while only hydrogens were able to move freely. During the next step, the backbone was constrained and the rest of the protein was free. The construction of the molecules, conformational analysis, minimisation of the molecules and CoMFA modelling were performed using the Sybyl 8.1 molecular with Tripos force field[27] modelling software. The steepest descent[28] and Powell methods were used to minimise the molecular structures of inhibitor with a termination gradient of 0.05 kcal/mol and maximum of 1000 iteration steps. Before performing CoMFA field calculations, the charges for each compound were calculated using the MMFF94 method.

2.1. Molecular dynamics

MD simulation was carried out using the GROMACS 3.3.3 program[29] to relax the crystal structure. In addition, parameters derived from the Morse potential were used for the Fe–N–C angle of the protoporphyrin ring as described by Rupp and co-workers.[30] These parameters had been previously determined for the octahedral coordination between the haeme iron, Cys-sulphur, and basic nitrogen of the porphyrin ring. During the MD simulation, the crystal structure was placed in a cubic water box at 300 K where periodic boundary conditions and the Particle Mesh Ewald (PME)[31] electrostatic conditions with Berendsen isotropic pressure coupling[32] were applied. First, the backbone and side chain atoms were constrained while the water molecules in the box and the hydrogens were able to move freely. In this simulation, the length of the time step was 0.002 ps when the simulation time was kept at 50 ps. Another simulation was performed so that the backbone atoms were constrained while the side chain atoms of amino acids were able to move freely. The simulation required 200 ps with a time step of 0.002 ps. The PROCHECK program[23] was used to test the stereochemical quality of the structure.

2.2. Molecular docking

GOLD[33,34] is a docking program that evaluates solutions of conformers at the protein active site using scoring functions to estimate changes in free energy (ΔG) during binding. Kirton and coworkers[35] described the use of ligand-specific iron parameters (GOLD 3.2) for docking to haeme-containing proteins, and demonstrated improved performance. These parameters were used for docking new conformers of our previously published inhibitors.[12,13,21] These inhibitors were docked to the active site of CYP2A6 using GOLD (v. 3.2) with the metal coordination parameters of ChemScore.[35] The radius for docking was set to 20 Å around the porphyrin ring of haeme. The GOLD program produced up to 10 different docking poses for each inhibitor, which were then ranked based on CScore[36] (consensus score) ranking values and visual inspection. The best-ranked conformers were chosen for the CoMFA models. For some quinolines and lactones, H-bond constraint (Asn297) was applied because the inhibition studies showed that they were oriented in an incorrect way. The docking results were analysed using the protein–ligand interaction fingerprint (PLIF) analysis in the molecular orbital calculation engine (MOE) to define the interaction between the inhibitors and the protein backbone as well as side chains.[37]

2.3. CoMFA models

CoMFA models were developed based on the docking alignment of the inhibitors with CYP2A6. In the CoMFA models, both electrostatic and steric fields were calculated for each molecule using the Sybyl (v. 8.1) default settings with an sp^3 hybridised carbon with a charge of +1 as a probe atom and a grid spacing of 2 Å. The column filtering value was set at 2.0 kcal/mol. The models were validated using leave-one-out (LOO), leave-some-out (LSO) with 5 groups, and leave-half-out (LHO) with two groups cross-validation. All of these cross-validation procedures were repeated 20 times, and each time the means of the statistical values were calculated.

Also a progressive scrambling method was applied to validate the models. This validation technique addresses the overly optimistic cross-validation or response randomisation results for redundant data sets. In this approach, small random perturbations are introduced into a data set. This causes the nominal predictivity of unstable models to decline rapidly, whereas robust models are relatively stable. The instantaneous slope of the predictivity with respect to the degree of perturbation (dq^2/dr^2yy') is the statistical value obtained from the progressive scrambling test used to determine the quality of the model. It depicts the model sensitivity to perturbation at a critical threshold level of perturbation (here 0.85). This value has been reported to be a reliable indicator of model complexity, that is, it helps to avoid overfitting due to an excess of partial least-square (PLS) components. Specifically, the number of PLS components yielding a dq^2/dr^2Y' slope near unity should be optimal.[38]

In the final nonvalidated model, the number of components was chosen according to the statistical values of the cross-validated models and the scrambling test, such as the lowest S_{PRESS} or $SDEP_S$, and highest q^2 or $Q^{*2}{}_s$.

2.4. Virtual screening

Flexible 3D screening was performed using the Sybyl software (v. 8.1) with unity tool to screen the Maybridge database (www.maybridge.com) containing approximately 60,000 compounds. The database query was generated based on the CYP2A6 structure co-crystallised with the 3-heteroaromatic pyridine analogue of nicotine.[11] The query consisted of two hydrophobic sites (tolerance 1.0 Å) and one donor (tolerance 0.5 Å) atom near haeme or an acceptor atom close to Asn297 (tolerance 0.5 Å). Functional groups that can cause irreversible inhibition, such as aldehyde or thiophene, were included in the query. The Lipinski rule of five was applied with a maximum weight of 300 g/mol. The hit compounds were docked into the active site of CYP2A6 as described before (see 2.2). The molecules were ranked according to the predicted inhibitory potencies with CoMFA model and CScore (consensus score) ranking values. Some of the selected molecules also had a functional group that was placed at an optimal distance from the haeme moiety. The MOE similarity search tool was also used to find additional molecules similar to those identified in the first database screening.

3. Chemicals and biological material

4-(2-Thienyl) benzaldehyde, benzo[*b*]thiophene-2-carboxylic acid, 4-formylbenzonitrile, 5-bromo-1*H*-indole, 4-bromo-1*H*-indole, 1-benzothiophene-3-ylmethylamine hydrochloride, 1-benzothiophene-5-ylmethylamine, 2-(4-chlorophenyl)-3-oxopropane nitrile, 2-(4-methyl-2,3-dihydro-1,3-thiazol-2-ylidene)-3-oxobutanenitrile, 2-chloro-6-(2,2,2-trifluoroethoxy)benzonitrile, and 2-[(5-amino-1,3,4-thiadiazol-2-yl)thio]acetonitrile were purchased from Maybridge (www.maybridge.com). Furafylline, ticlopidine, methoxsalen, memantine, and α-naphthoflavone were purchased from Sigma–Aldrich (St. Louis, MO, USA). Their characterizations and purities, which were greater than 95%, were confirmed using ^{1}H NMR, mass spectrometry (MS), and high-performance liquid chromatography (HPLC) methods.

Human liver samples were obtained from patients undergoing surgery to remove hepatic tumours. The patients have been described previously.[21] The use of surplus tissue was approved by the Ethics Committee of the University of Kuopio. Liver samples were frozen in liquid nitrogen and stored at −70 °C. Only tumour-free tissue was used for the experiments. Baculovirus-insect cell-expressed human CYP2A6 was purchased from BD Biosciences Discovery Labware (Bedford, MA, USA).

4. In vitro assays

4.1. IC_{50} assay

A coumarin 7-hydroxylation activity assay with a 96-well plate method was used in the studies of CYP2A6-mediated metabolism. The method is based on the detection of fluorescence emitted by 7-hydroxycoumarin in alkaline conditions.[39] In each well, the reaction mixture contained 50 mM Tris–HCl buffer (pH 7.4), 5.0 mM $MgCl_2$, 10 μM coumarin, 20 μg microsomal proteins or 0.5 pmol of cDNA-expressed CYP2A6, and 0.3 mM NADPH. The reaction was initiated by addition of NADPH, incubated at 37 °C for 10 min, and terminated by adding 60 μl 10% TCA. Immediately before measurement, 140 μl 1.6 M glycine–NaOH buffer (pH 10.4) was added. The formed fluorescence was measured with a Victor^{2TM} plate counter (Perkin Elmer Life Sciences Wallac, Turku, Finland) at 355 nm excitation and 460 nm emission. Control incubations were carried out without the substrate, enzyme, or inhibitor. All inhibitors were dissolved in dimethylsulphoxide (DMSO), and a final concentration of 2% DMSO was used in the incubations to ensure complete solubility of the compounds. This concentration caused an about 20% inhibition in CYP2A6 catalytic activity; this effect was taken into account in the calculations. Each inhibitor was prescreened using inhibitor concentrations ranging from 0.1 to 100 μM. The actual IC_{50} values were determined using five to seven inhibitor concentrations. All IC_{50} values were determined in duplicate from liver microsomal preparations or 0.5 pmol cDNA-expressed human CYP2A6 enzyme. The IC_{50} was determined by nonlinear regression analysis using Sigma Plot.

4.2. Time-dependent inhibition assay

Time-dependent inhibition was investigated by two experimental procedures. The initial assay used different inhibitor concentrations and preincubation times to generate IC_{50} values. Although this assay was carried out as the IC_{50} assay, the reaction was initiated by the addition of coumarin and was preincubated up to 30–60 min. Control reactions were carried out either without coumarin, without the inhibitor, or by adding NADPH just after preincubation. For determination of irreversible inhibition, various concentrations of the inhibitors were preincubated up to 4–60 min. The reaction mixture consisted of 50 mM Tris–HCl buffer (pH 7.4), 5 mM $MgCl_2$, 200–800 μg microsomal protein or 5–20 pmol cDNA-expressed human CYP2A6 enzyme, 0.3 mM NADPH, and the inhibitor. The reaction mixture was diluted 10–40 times with 50 mM Tris–HCl buffer (pH 7.4), 5 mM $MgCl_2$, 10 μM coumarin, and 0.3 mM NADPH, and the incubation was continued for 10 min at 37 °C. The reaction was terminated and measured as described for the IC_{50} assay.

Measurements detecting slow binding inhibition were carried out at 37 °C in a reaction mixture containing 50 mM Tris–HCl buffer (pH 7.4), 5 mM $MgCl_2$, 12 μM NADPH, 10 μg microsomal protein, and 2.5 or 5 μM coumarin. The reaction was initiated by adding the enzyme, and the reaction curves were monitored every 2 min for 60 min using the Victor2 fluorescence plate reader at excitation and emission wavelengths of 355 and 460 nm, respectively. The progress curves of the reactions were fitted to the Morrison equation,[40] in which the curves with the inhibitors present are characterised by an initial burst phase followed by a slower steady-state rate. The product concentration is given by

$$[p] = v_s t + \frac{v_0 v_s}{k_{obs}}(1 - e^{-k_{obs}t}) \quad (1)$$

where v_0 and v_s are the initial and final steady state reaction rates, respectively, t is the time, and k_{obs} is the apparent first-order rate constant.

Two basic reaction mechanisms have been proposed to explain slow enzyme inhibition. In the one-step mechanism, either the formation of the EI complex is a slow process (low association constant, k_{on}), or both the association (k_{on}) and dissociation (k_{off}) processes are slow. In the two-step mechanism, rapid formation of the EI complex (fast association and dissociation constants k_1 and k_2, respectively) is followed by slow isomerization to a more stable E·I* complex (slow association and dissociation constants k_3 and k_4, respectively). The apparent first-order rate constant (k_{obs}) is calculated from the progress curves. The association and dissociation constants of a slow-binding inhibitor can be determined by plotting the apparent first-order constant k_{obs} against inhibitor concentrations. In the one-step mechanism, k_{obs} is linearly dependent on the inhibitor concentration, whereas in the two-step mechanism the dependence is hyperbolic. The association rate constants were solved by fitting data to the following equations.

$$k_{obs} = k_{off} + k'_{on}[I] \quad (2)$$

$$K_{on} = k'_{on}(1 + S/K_m) \quad (3)$$

where k_{off} is the dissociation constant and k'_{on} the association constant. k'_{on} is the slope from the plot of k_{obs} against the inhibitor concentration. k_{on} was corrected for substrate competition using Eq. 2; k_{off} values were determined from the y-axis intercepts.[41]

5. Results

5.1. Docking

The active site of CYP2A6 is surrounded mostly by lipophilic amino acids that limit the size of inhibitors[22] and contains an amino acid residue (Asn297) that is capable of acting as an H-bond donor. The docked CYP2A6 inhibitors consisted of 85 compounds published previously (see Supplementary data, section A).[12,13,21] According to the PLIF analysis, the training set could be divided into four categories. In the first category, potent inhibitors with a carbonyl oxygen acceptor atom (e.g., coumarin) form an H-bond with the NH_2 group of Asn297. In addition, Asn297 acts as an H-bond donor to the carbonyl oxygen of lactone derivatives. In the second category, the nitrogen atom of an amine group of alkylamine derivatives coordinates towards the haeme, indicating potent

inhibition. The third category consists of weak inhibitors with acceptor atoms, such as lactone derivatives with a short alkyl chain (<3 carbons), which are usually too small to fill the active site of CYP2A6 (260 Å). The fourth category consists of some potent inhibitors without donor/acceptor atoms, such as halogen-substituted naphthalene derivatives, which interact effectively with the hydrophobic amino acids in the active site.

5.2. Model statistics

Most of the values obtained from the CoMFA models were statistically significant (Table 1). However, the LSO method was more stringent than the LOO method. Therefore, LSO cross-validation with 5 groups was used to validate the model.

A progressive scrambling test slope (dq^2/dr_{yy}^2) of more than 1.2 can lead to overfitting. Therefore, the number of components in the final model was chosen so that the slopes were <1.2 (Table 2). The statistical values $Q^{*2}{}_s$ and SDEP*s were comparable to the corresponding q^2 and S_{PRESS} values of cross-validation in all models.

5.3. CoMFA fields

The CYP2A6 CoMFA contour map contained electrostatic fields near Asn297 consisting of electronegative (red) and electropositive favoured (blue) regions (Fig. 1). Electrostatic interaction regions near Asn297 were associated with the presence of H-bonds between the inhibitor and protein. Another electronegative-favoured region was located near the nitrogen atom of alkyl amine derivatives close to the haeme iron, indicating coordination of nitrogen towards the haeme. This was confirmed by measuring a type II binding spectrum. The steric fields (green) suggested that inhibitory potency should increase if the inhibitor is occupied by a hydrophobic group. At the active site cavity, three phenylalanines form a cluster that can undergo aromatic π–π stacking interactions with aromatic-type inhibitors. The steric fields in the CoMFA map were located near this cluster, where hydrophobic interactions are thought to be important. The sterically favoured and disfavoured regions provided indications about the optimal volume and shape of a potent inhibitor.

5.4. Virtual screening

The CYP2A6 crystal structures complexed with 3-heteroaromatic pyridine derivatives showed that these compounds are tightly coordinated towards haeme[11] and lead to potent enzyme inhibition. One of these crystal structures (PDB: 2FDW) and the CoMFA model were used to create the features in the virtual screening queries. According to the model, a partial negative charge near the haeme and an H-bond atom near Asn297 are favourable for potency. Thus, either a donor atom (see Supplementary data, section B) above the haeme or an acceptor atom close to Asn297 was included in the queries. The optimal size and position of the hydrophobic sites in the queries were constructed on the grounds of the sterically favoured regions of the models. Given that one goal of the screening procedure was to identify mechanism-based enzyme inactivators, different functional groups potentially mediating irreversible binding were included in the queries. The database search query returned hundreds of molecules, many of which contained the desired functional groups. The selection of compounds for further analysis was based on either high predicted inhibitory potency, the scoring values in docking, or visual inspection for the desired properties. Molecules that were predicted to have an IC_{50} value of <100 μM or contained functional groups located at an optimal distance from the haeme were selected. Nine of the hit molecules were tested for inhibitory potency (Fig. 2). Another search was carried out using the MOE similarity search tool to find additional molecules with similar desirable characteristics.

Table 1
Statistical values of CYP2A6 CoMFA models

Model no.	Cv[a]	q^{2b}	N[c]	S_{PRESS}[d]
1	LOO	0.42	1	0.89
2	LOO	0.47	2	0.85
3	LOO	0.52	3	0.81
4	LOO	0.58	4	0.77
5	LOO	0.62	5	0.74
6	LSO	0.41	1	0.89
7	LSO	0.48	2	0.86
8	LSO	0.52	3	0.82
9	LSO	0.58	4	0.77
10	LSO	0.62	5	0.75
11	LHO	0.50	1	0.96
12	LHO	0.50	2	0.87
13	LHO	0.58	3	0.80
14	LHO	0.61	4	0.75
15	LHO	0.66	5	0.75

[a] Cv = cross-validation method.
[b] q^2 = cross-validated correlation coefficient.
[c] N = number of components.
[d] S_{PRESS} = standard deviation for the prediction error; Final model (no. 10) r^2 (correlation coefficient) = 0.93; S (standard error of estimate) = 0.32.

Table 2
Progressive scrambling statistics for the CYP2A6 CoMFA model

Model	N[a]	$Q^{*2}{}_s$[b]	$SDEP^*{}_s$[c]	dq^2/dr_{yy}^{2}[d]
CYP2A6	1	0.33	0.94	0.55
	2	0.32	0.96	1.04
	3	0.32	0.86	1.22
	4	0.38	0.93	1.18
	5	0.43	0.89	1.14

[a] N, number of PLS components.
[b] $Q^{*2}{}_s$, predictivity at the critical threshold level of perturbation s (s = maximum value of $Q^{*2}{}_s$; s = 0.85).
[c] SDEPs*, standard error of prediction at the critical threshold level of perturbation s.
[d] dq^2/dr_{yy}^2, sensitivity to perturbation at s.

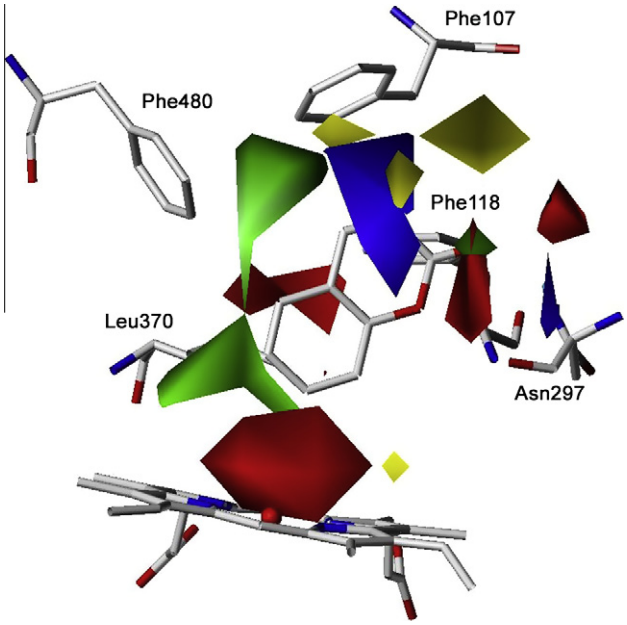

Figure 1. CoMFA fields of the model based on CYP2A6 crystal structure with coumarin as the template. CoMFA electrostatic fields: blue, negative-charge disfavoured area; red, negative-charge favoured area. CoMFA steric field; green, bulk favoured area; yellow, bulk disfavoured area.

Two molecules with the greatest similarities to the most potent inhibitors (**3**–**4**) of the first search were selected for further testing.

5.5. Prediction by the models

Overfitting is a common problem in quantitative structure–activity relationship (QSAR) approaches related to overoptimistic q^2values, that is, a high q^2 value does not necessarily indicate significant predictivity of the model.[42] Therefore, it is recommended that external tests are used to evaluate the reliability of the models.[42,43] We used external test sets that consisted of both well-known CYP inhibitors and the hit molecules of the database screening. The models predicted inhibitory potency (pIC_{50}) values quite accurately, although not all of the molecules in the test set were similar to those in the training set. Many test molecules contained functional groups not present in the training molecules. A few outlier molecules were also found where the residual for prediction was >1 (Fig. 3). These outlier molecules were structurally different from the training set molecules; this is a possible reason for the poor predictivity.

5.6. Inhibition characteristics

The hit molecules were tested for their ability to inhibit CYP2A6-mediated coumarin 7-hydroxylase activity in human liver microsomes and in cDNA-expressed CYP2A6 enzyme (Table 3, Fig. 4). The most potent inhibitors identified using the MOE similarity search tool were molecule **5** (IC_{50} = 0.1 μM; no preincubation) and molecule **6** (IC_{50} = 0.05 μM; after preincubation). The alkylamine derivatives **3** and **4** were also quite potent. The mechanism of CYP2A6 inhibition was examined in greater detail for selected compounds.

The inhibitory potency of compounds **3** and **4** was enhanced after a 30-min preincubation period. Alkylamine derivatives are known to form an MIC with CYP protein.[44] However, more detailed analysis with methods detecting time-dependent inhibition using cDNA-expressed CYP2A6 did not reveal quasi-irreversible inhibition for compounds **3** and **4**. For compound **6**, the inhibition was dependent on NADPH, preincubation time, and inhibitor concentration (Fig. 5A). The K_{inact} (the maximal rate of enzyme inactivation at saturated inhibitor concentration) value was 0.36 min^{-1}, whereas the K_I (the concentration required for half-maximal enzyme inactivation) was 29 μM for compound **6** (Fig. 5B).

5.7. Inhibition kinetics of the most potent inhibitors

The inhibitor binding rates were evaluated. The progress curves of the enzyme catalytic reaction did not display a simple linear product-versus-time relationship for inhibitors **3**–**6** when the reaction was initiated by enzyme. The enzyme reaction progress curves for compounds **3**, **5**, and **6** are shown in Figures 6–8. The curves showed a gradually decreasing rate of reaction, indicating slow binding of the inhibitors. Given that k_{obs} was linearly dependent on the inhibitor concentration for compounds **3**–**5**, slow formation of the EI complex indicates a one-step mechanism (Fig. 6 for compound **3**, Fig. 7 for compound **5**). These compounds had low dissociation and association constants. In contrast, the dependence of k_{obs} on the inhibitor concentration was hyperbolic for compound **6** (Fig. 8), suggesting that binding occurs according to a two-step mechanism.

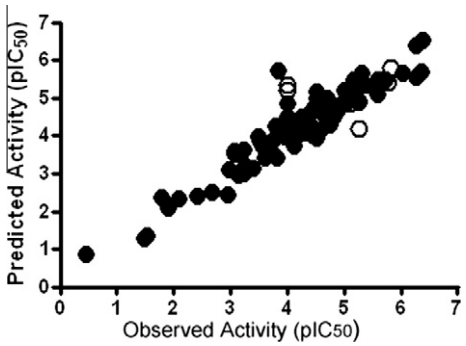

Figure 3. Plots of observed versus predicted activities of the test (○) and training (●) sets for cyp2a6 enzyme.

Table 3
IC_{50} values of inhibitors

No.	IC_{50} value[a] (μM)	K_{on}[b]	K_{off}[c]
1	>100		
2	>100		
3	1.5/0.2[b]	0.17	0.003
4	1.7/0.6[b]	1.02	0.08
5	0.1	0.13	0.002
6	1.6/0.05[b]		
7	6.4		
8	>100		
9	>100		
10	>100		
11	>100		

[a] IC_{50} values represent the mean from 2 to 4 independent experiments performed in duplicate.
[b] μM^{-1} min^{-1}.
[c] min^{-1}.

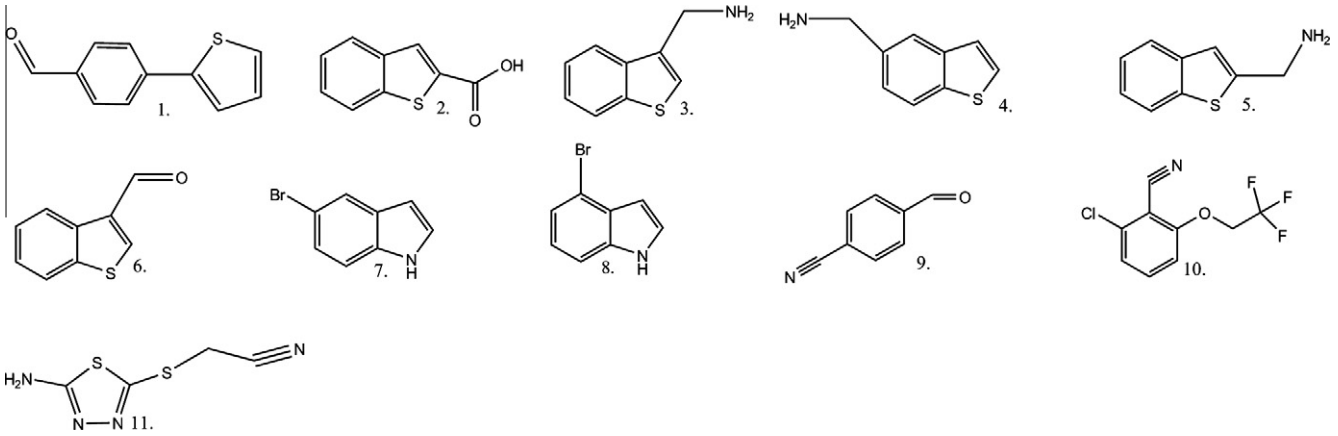

Figure 2. Structures of hit molecules from virtual screening.

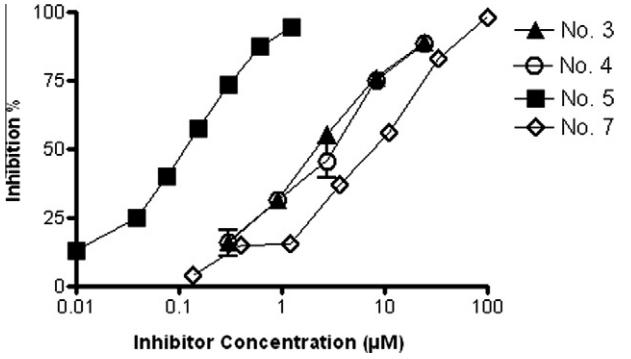

Figure 4. Inhibition of CYP2A6 activity by compounds **3**–**5** and **7**.

6. Discussion

In this study we employed three-dimensional quantitative structure–activity relationship (3D-QSAR) modeling, virtual screening and biological testing to identify novel scaffolds for CYP2A6 inhibitors. Three alkylamine derivatives (**3**–**5**) and 1-benzothiophene-3-carbaldehyde (**6**) were the most potent inhibitors of CYP2A6. The results show that crystal structure-based CoMFA combined with virtual screening can be used successfully to design novel inhibitors for the CYP2A6 enzyme. Our previous studies have also shown that 3D-QSAR methods are valuable in evaluating the structural features that are important for inhibitory potency against other CYP enzymes.[12,45,46]

A CoMFA model was used to identify specific binding features at the CYP2A6 active site. The CoMFA models created in this study improved upon our previous models.[12,13] First, the training sets contained more molecules than the previously constructed models. Our previous CYP2A6 models did not contain potent inhibitors with acceptor/donor atoms, which turned out to be an important feature for CYP selectivity. Second, the use of ligand-specific iron parameters in the docking analysis yielded a more realistic positioning of lactone.

Third, the use of the CYP2A6 crystal structures together with docking was an improvement over the previous 3D-QSAR models, because a more precise definition of the bioactive conformation of molecules was achieved. According to the current docking results, Asn297 in the CYP2A6 binding pocket formed H-bonding with CYP2A6 inhibitors containing an acceptor atom. Val117 and Leu370 in the B′-C loop and K-K′-loop regions interacted with most of the inhibitors, which is consistent with previous mutagenesis and docking studies.[47,48] Our earlier study showed that the optimal size of CYP2A6 inhibitors is limited and the presence of overly large substituents on the lactone moiety will decrease inhibitory potency.[49] This was verified in the docking analysis, where lactone derivatives with a substitution of a 3–4 carbon alkyl chain had the optimal conformation, interacting with Val117 and Leu370. The present crystal structure-based CoMFA model also showed that a hydrogen bond with Asn297 and van der Waals interactions between the aromatic cluster of phenylalanines and aromatic ring of inhibitors are important determinants of inhibitory potency. These observations are in agreement with other recent CYP2A6 QSAR studies (partially based on our previously published datasets) showing the dependence of binding affinity of inhibitors on multiple factors such as electronic, hydrophobic and spatial properties of the molecules.[50–53]

Although 3D-QSAR approaches are very useful in drug discovery, some limitations are apparent. Overfitting is a common problem of QSAR with the PLS method leading to a decrease of the prediction accuracy of the external dataset.[42,54] In addition, the chosen alignment procedure can affect dramatically the 3D-QSAR models. If the 3D structure of the target is not available, the active conformation of a molecule is often difficult to define, leading to

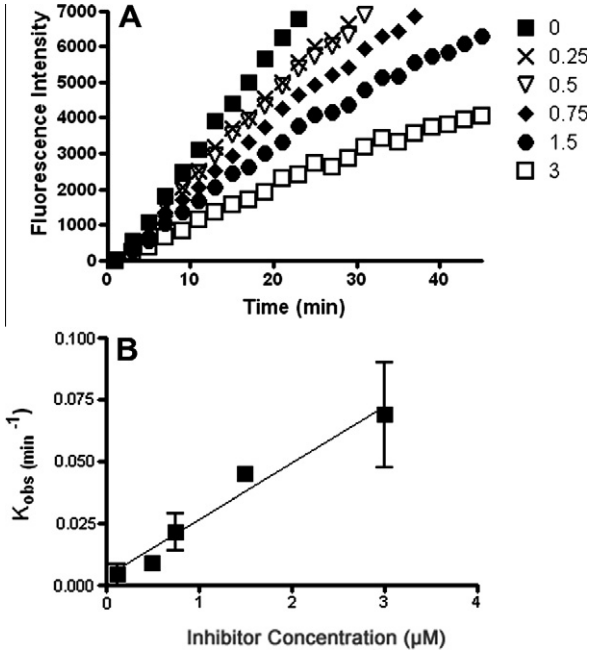

Figure 6. (A) Progress curves for hydroxylation of coumarin (2.5 μM) in the presence of various concentrations (μM) of compound **3**; (B) K_{obs} as a function of inhibitor concentration for compound **3**.

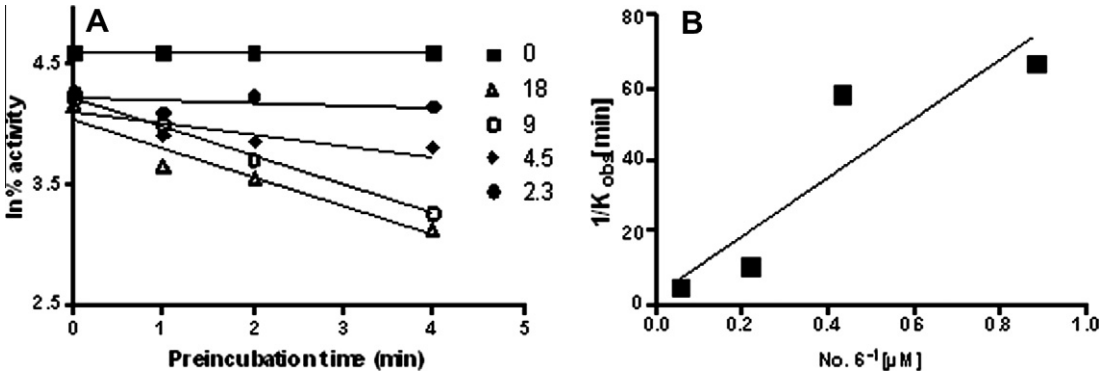

Figure 5. (A) Time- and concentration-dependent inactivation of CYP2A6 enzyme activity by compound **6**; (B) Double-reciprocal plot of the relationship between K_{obs} and inhibitor concentration.

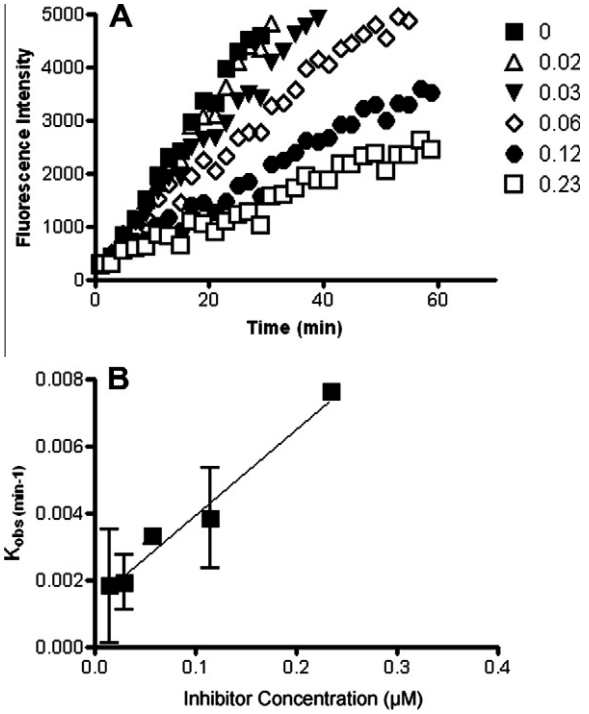

Figure 7. (A) Progress curves for hydroxylation of coumarin (5 μM) in the presence of various concentrations (μM) of compound **5**; (B) K_{obs} as a function of inhibitor concentration for compound **5**.

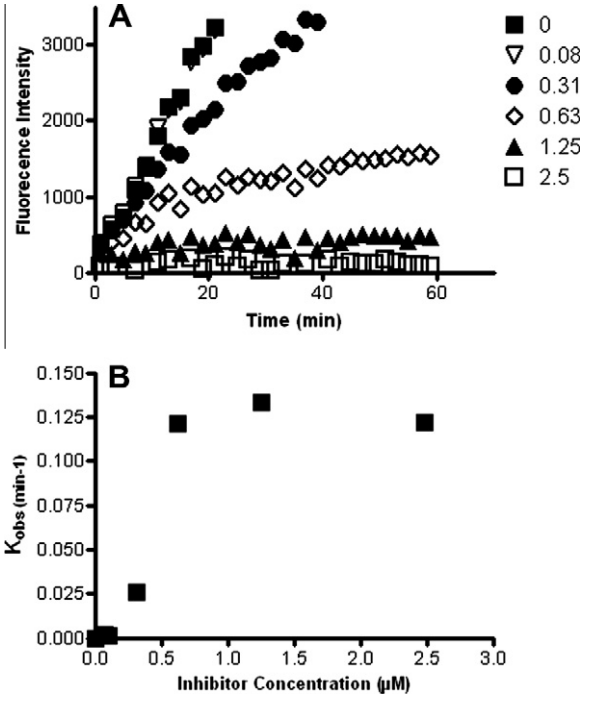

Figure 8. (A) Progress curves for hydroxylation of coumarin (5 μM) in the presence of various concentrations (μM) of compound **6**; (B) K_{obs} as a function of inhibitor concentration for compound **6**.

wrong conclusions, since descriptors of the model are dependent on the conformation of the compounds.[54,55]

A database search was conducted using CYP2A6 co-crystallised with 3-heteroaromatic pyridine analogues to define the features in the query. The purpose of the search was to identify both potent and reversible mechanism-based CYP2A6 inhibitors. There are many organic functional groups known to be capable of binding irreversibly to CYPs and destroying the enzyme.[17] Compounds can covalently bind to the enzyme active site, leading to a long-lasting inactivation of the enzyme. Electrophilic reactive metabolites formed in the reaction can cause irreversible mechanism-based inactivation by modifying nucleophilic amino acid residues in the apoprotein or by alkylating porphyrin nitrogen prosthetic groups. The destruction of the prosthetic haeme group leads to haeme-derived fragments that covalently modify the apoprotein.[15,16] The fields in the CoMFA models were used to define the position of these irreversible functional groups such as thiophene, nitriles, benzaldehydes, and alkyl amines near the haeme.

Also a similarity search can be performed to identify inhibitors. However, similarity searches are based on the concept that chemically closely related analogues should be related in their mode of action, as well as in their relative potencies.[56] Therefore the identified hit structures are close analogues with the template structures. 3D-QSAR modelling combined with virtual screening yields structural diversity in hit molecules. In this study the purpose of virtual screening was to identify new scaffolds by combining different functional groups in CYP2A6 inhibitors. Therefore the active compounds in this study consisted of structural elements that were known to be important for inhibitory activity. However, the active compounds identified here consisted of a flat benzothiophene and indole type cores, constituting more compact structures than previously disclosed inhibitors. None of these inhibitors are suitable for routine clinical use as such and thus there is a need to design novel derivatives with more suitable toxicity profiles.

Several types of CYP2A6 inhibitors have been characterised to date. Some 3-heteroaromatic pyridine derivatives have been characterised as type II inhibitors, indicating that an amino group of the side chain coordinates to the CYP haeme iron. Some of these agents have been identified to be MBIs.[11] Alkylamine derivatives can tightly coordinate towards haeme, explaining their slow inhibition kinetics. Alkyl and aromatic amines can form quasi-irreversible nitroso-iron complexes that do not actually destroy the enzyme although it becomes catalytically inert. The term quasi-irreversible inhibition indicates that catalytically active CYPs can be regenerated in vitro but not in vivo.[57] Alkyl amine derivatives require oxidation to *N*-hydroxylamine (N-oxidation) and nitrosoalkane derivatives to form a quasi-irreversible alkylamine MIC.[18,44,58] However, it has been reported that N-oxidation and MIC formation for primary amines is slow and less extensive than for secondary amines.[59] Compound **6** may be a mechanism-based inactivator or become a more stable metabolite by a slow isomerization process. This inhibitor contained thiophene and aldehyde functional groups, each capable of inhibiting CYP enzymes in an irreversible manner.[21,60] The inactivation of CYP enzymes by aldehydes appears to involve the formation of free radicals that can form haeme adducts.[61] The general mechanism of inactivation by thiophene derivatives appears to involve the formation of an epoxide intermediate, which can attack the haeme or one of the nitrogen atoms in the protein side chain.[17] In both cases, the result is the formation of a covalent bond between the reactive metabolite and the CYP enzyme, which is consequently inactivated.

The alkylamine derivatives (**3**–**5**), 1-benzothiophene-3-carbaldehyde (**6**), and 5-bromo-1*H*-indole (**7**) were more potent inhibitors than nicotine. The scaffolds of these structures are more rigid than that of nicotine. Nicotine is a tertiary amine with one free electron pair whereas the alkyl amine derivatives in the current study have two free electron pairs causing increased alkalinity. It has been speculated that the reactive intermediate of nicotine, nicotine-Δ1′(5′)-iminium ion, is responsible for enzyme inactivation[62,63] indicating a different mechanism of irreversible inhibition in comparison with compound **6**.

7. Conclusion

The aim of the present study was to design novel CYP2A6 inhibitors to be evaluated as potential lead compounds in development of drugs for tobacco smoking reduction therapy. Different molecular modelling methods were used to determine typical features of CYP2A6 inhibitors and obtain more detailed information about the important interactions between inhibitors and the CYP2A6 protein. 3D-QSAR methods combined with docking tools and virtual screening proved to be a valuable approach in identifying inhibitor candidates. Several novel scaffolds of inhibitors such as benzothiophene and indole derivatives were identified for use as leads in the design of CYP2A6 inhibitor drugs. In conclusion, the combination of CoMFA and docking analysis together with virtual screening was a good approach to design CYP2A6 inhibitors. Several types of potent and novel CYP2A6 inhibitors, to be further assessed in the future, were identified in the study.

Acknowledgments

We thank Ms. Hannele Jaatinen for excellent technical help, Dr. Jarkko Venäläinen for his help in analyses, Dr. Tiina Kääriäinen for her important support and Dr. Ewen MacDonald for help in preparing the manuscript. Funding from the Academy of Finland, Substance Abuse and Addiction Program, and the Yrjö Jahnsson Foundation are gratefully acknowledged. We thank CSC-Scientific Computing Ltd. (Espoo, Finland) for the software and hardware resources.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2011.09.054.

References and notes

1. Berkman, C. E.; Park, S. B.; Wrighton, S. A.; Cashman, J. R. *Biochem. Pharmacol.* **1995**, *50*, 565.
2. Nakajima, M.; Yamamoto, T.; Nunoya, K.; Yokoi, T.; Nagashima, K.; Inoue, K.; Funae, Y.; Shimada, N.; Kamataki, T.; Kuroiwa, Y. *Drug Metab. Dispos.* **1996**, *24*, 1212.
3. Messina, E. S.; Tyndale, R. F.; Sellers, E. M. *J. Pharmacol. Exp. Ther.* **1997**, *282*, 1608.
4. Yamazaki, H.; Inoue, K.; Hashimoto, M.; Shimada, T. *Arch. Toxicol.* **1999**, *73*, 65.
5. Benowitz, N. L.; Swan, G. E.; Jacob, P., 3rd; Lessov-Schlaggar, C. N.; Tyndale, R. F. *Clin. Pharmacol. Ther.* **2006**, *80*, 457.
6. Ho, M. K.; Mwenifumbo, J. C.; Al Koudsi, N.; Okuyemi, K. S.; Ahluwalia, J. S.; Benowitz, N. L.; Tyndale, R. F. *Clin. Pharmacol. Ther.* **2009**, *85*, 635.
7. Malaiyandi, V.; Lerman, C.; Benowitz, N. L.; Jepson, C.; Patterson, F.; Tyndale, R. F. *Mol. Psychiatry* **2006**, *11*, 400.
8. Schoedel, K. A.; Hoffmann, E. B.; Rao, Y.; Sellers, E. M.; Tyndale, R. F. *Pharmacogenetics* **2004**, *14*, 615.
9. Sellers, E. M.; Kaplan, H. L.; Tyndale, R. F. *Clin. Pharmacol. Ther.* **2000**, *68*, 35.
10. Tyndale, R. F.; Sellers, E. M. *Drug Metab. Dispos.* **2001**, *29*, 548.
11. Yano, J. K.; Denton, T. T.; Cerny, M. A.; Zhang, X.; Johnson, E. F.; Cashman, J. R. *J. Med. Chem.* **2006**, *49*, 6987.
12. Rahnasto, M.; Raunio, H.; Poso, A.; Wittekindt, C.; Juvonen, R. O. *J. Med. Chem.* **2005**, *48*, 440.
13. Rahnasto, M.; Wittekindt, C.; Juvonen, R. O.; Turpeinen, M.; Petsalo, A.; Pelkonen, O.; Poso, A.; Stahl, G.; Holtje, H. D.; Raunio, H. *Pharmacogenomics J.* **2008**, *8*, 328.
14. Zhang, W.; Kilicarslan, T.; Tyndale, R. F.; Sellers, E. M. *Drug Metab. Dispos.* **2001**, *29*, 897.
15. Kent, U. M.; Juschyshyn, M. I.; Hollenberg, P. F. *Curr. Drug Metab.* **2001**, *2*, 215.
16. Kalgutkar, A. S.; Obach, R. S.; Maurer, T. S. *Curr. Drug Metab.* **2007**, *8*, 407.
17. Fontana, E.; Dansette, P. M.; Poli, S. M. *Curr. Drug Metab.* **2005**, *6*, 413.
18. Polasek, T. M.; Elliot, D. J.; Lewis, B. C.; Miners, J. O. *J. Pharmacol. Exp. Ther.* **2004**, *311*, 996.
19. Cramer, R. D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
20. Cramer, R. D., 3rd; Patterson, D. E.; Bunce, J. D. *Prog. Clin. Biol. Res.* **1989**, *291*, 161.
21. Rahnasto, M.; Raunio, H.; Poso, A.; Juvonen, R. O. *Xenobiotica* **2003**, *33*, 529.
22. Yano, J. K.; Hsu, M. H.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. *Nat. Struct. Mol. Biol.* **2005**, *12*, 822.
23. Laskowski, R. A.; Macarthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283.
24. Tripos Associates. In http://www.tripos.com/: St. Louis, Missouri, USA.
25. Fletcher, A. *Comput. J.* **1970**, *13*, 317.
26. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
27. Clark, M.; Cramer, R. D.; Vanopdenbosch, N. *J. Comput. Chem.* **1989**, *10*, 982.
28. Shanno, D. *Math. Comput.* **1970**, *24*, 647.
29. Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model* **2001**, *7*, 306.
30. Rupp, B.; Raub, S.; Marian, C.; Holtje, H. D. *J. Comput. Aided Mol. Des.* **2005**, *19*, 149.
31. Essman, U.; Perela, L.; Berkowitz, M.; Darden, T.; Lee, H.; Pederson, L. A. *J. Chem. Phys.* **1995**, *38*, 8577.
32. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
33. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *Abstr. Papers Am. Chem. Soc.* **1997**, *214*, 154.
34. Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. *Proteins-Struct. Funct. Genet.* **2003**, *52*, 609.
35. Kirton, S. B.; Murray, C. W.; Verdonk, M. L.; Taylor, R. D. *Proteins* **2005**, *58*, 836.
36. Tripos Associates. In http://www.tripos.com/sciTech/inSilicoDisc/virtualScreening/cscore.html#references: St. Louis, Missouri, USA.
37. Chemical Computing Group. In http://www.chemcomp.com/journal/ligintdia.htm: Montreal, QC, Canada.
38. Clark, R. D.; Fox, P. C. *J. Comput. Aided Mol. Des.* **2004**, *17*, 1.
39. Aitio, A. *Anal. Biochem.* **1978**, *85*, 488.
40. Morrison, J. F.; Walsh, C. T. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1988**, *61*, 201.
41. Morrison, J. F. *Trends. Biochem. Sci.* **1982**, *7*, 102.
42. Golbraikh, A.; Tropsha, A. *J. Mol. Graph. Model* **2002**, *20*, 269.
43. Golbraikh, A.; Tropsha, A. *J. Comput. Aided Mol. Des.* **2002**, *16*, 357.
44. Polasek, T. M.; Miners, J. O. *Br. J. Clin. Pharmacol.* **2008**, *65*, 87.
45. Korhonen, L. E.; Rahnasto, M.; Mahonen, N. J.; Wittekindt, C.; Poso, A.; Juvonen, R. O.; Raunio, H. *J. Med. Chem.* **2005**, *48*, 3808.
46. Korhonen, L. E.; Turpeinen, M.; Rahnasto, M.; Wittekindt, C.; Poso, A.; Pelkonen, O.; Raunio, H.; Juvonen, R. O. *Br. J. Pharmacol.* **2007**, *150*, 932.
47. He, X. Y.; Shen, J.; Hu, W. Y.; Ding, X.; Lu, A. Y.; Hong, J. Y. *Arch. Biochem. Biophys.* **2004**, *427*, 143.
48. DeVore, N. M.; Smith, B. D.; Urban, M. J.; Scott, E. E. *Drug Metab. Dispos.* **2008**, *36*, 2582.
49. Juvonen, R. O.; Gynther, J.; Pasanen, M.; Alhava, E.; Poso, A. *Xenobiotica* **2000**, *30*, 81.
50. Roy, K.; Roy, P. P. *Eur. J. Med. Chem.* **2009**, *44*, 1941.
51. Leong, M. K.; Chen, Y. M.; Chen, H. B.; Chen, P. H. *Pharm. Res.* **2009**, *26*, 987.
52. Van Damme, S.; Bultinck, P. *J. Comput. Chem.* **2009**, *30*, 1749.
53. Wang, Y. H.; Li, Y.; Wang, B. *Int. J. Mol. Sci.* **2007**, *8*, 166.
54. Kubinyi, H. *Drug Discovery Today* **1997**, *2*, 538.
55. Golbraikh, A.; Bernard, P.; Chretien, J. R. *Eur. J. Med. Chem.* **2000**, *35*, 123.
56. Kubinyi, H. *Perspect. Drug Discov. Des.* **1998**, *9–11*, 225.
57. Polasek, T. M.; Miners, J. O. *Expert. Opin. Drug Metab. Toxicol.* **2007**, *3*, 321.
58. Polasek, T. M.; Elliot, D. J.; Somogyi, A. A.; Gillam, E. M.; Lewis, B. C.; Miners, J. O. *Br. J. Clin. Pharmacol.* **2006**, *61*, 570.
59. Murray, M. *Clin. Exp. Pharmacol. Physiol.* **1997**, *24*, 465.
60. Raner, G. M.; Chiang, E. W.; Vaz, A. D.; Coon, M. J. *Biochemistry* **1997**, *36*, 4895.
61. Correia, M. A.; Ortiz de Montellano, P. In Ortiz de Montellano, P., Ed.; Cytochrome P450 Structure, Mechanism, and Biochemistry; Kluwer AcademicPlenum Publisher: New York, 2005; p 247.
62. Denton, T. T.; Zhang, X.; Cashman, J. R. *Biochem. Pharmacol.* **2004**, *67*, 751.
63. Murphy, S. E.; von Weymarn, L. B.; Brown, K. M. *J. Pharmacol. Exp. Ther.* **2006**, *316*, 295.