

Competitive CYP2C9 Inhibitors: Enzyme Inhibition Studies, Protein Homology Modeling, and Three-Dimensional Quantitative Structure-Activity Relationship Analysis

LOVISA AFZELIUS, ISMAEL ZAMORA, MARIANNE RIDDERSTRÖM, TOMMY B. ANDERSSON, ANDERS KARLÉN, and COLLEN M. MASIMIREMBWA

Department of Drug Metabolism and Pharmacokinetics & Bioanalytical Chemistry, AstraZeneca R&D, Mölndal, Sweden (L.A., I.Z., M.R., T.B.A., C.M.M.); and Department of Organic Pharmaceutical Chemistry, Biomedical Center, Uppsala University, Uppsala, Sweden (L.A., A.K.)

Received October 3, 2000; accepted December 20, 2000

This paper is available online at <http://molpharm.aspetjournals.org>

ABSTRACT

This study describes the generation of a three-dimensional quantitative structure activity relationship (3D-QSAR) model for 29 structurally diverse, competitive CYP2C9 inhibitors defined experimentally from an initial data set of 73 compounds. In parallel, a homology model for CYP2C9 using the rabbit CYP2C5 coordinates was built. For molecules with a known interaction mode with CYP2C9, this homology model, in combination with the docking program GOLD, was used to select conformers to use in the 3D-QSAR analysis. The remaining molecules were docked, and the GRID interaction energies for all conformers proposed by GOLD were calculated. This was followed by a principal component analysis (PCA) of the GRID energies for all conformers of all compounds. Based on the

similarity in the PCA plot to the inhibitors with a known interaction mode, the conformer to be used in the 3D-QSAR analysis was selected. The compounds were randomly divided into two groups, the training data set ($n = 21$) to build the model and the external validation set ($n = 8$). The PLS (partial least-squares) analysis of the interaction energies against the K_i values generated a model with $r^2 = 0.947$ and a cross-validation of $q^2 = 0.730$. The model was able to predict the entire external data set within 0.5 log units of the experimental K_i values. The amino acids in the active site showed complementary features to the grid interaction energies in the 3D-QSAR model and were also in agreement with mutagenesis studies.

An understanding of the drug-metabolizing enzymes' role in the clearance of compounds and of drug-drug interactions caused by coadministered medications is a key issue in both the drug discovery process and in the use of therapeutics. Advances in analytical and in vitro metabolism technologies have made it possible to determine metabolic stability, identify metabolites, and specify enzymes involved in compound biotransformation. Genomic research will lead to an increased number of therapeutic targets, and combinatorial chemistry has increased the number and chemical diversity of compounds that can be tested against these targets. High-throughput screening approaches in metabolism research, therefore, are being developed to meet this challenge.

The use of computational techniques to predict metabolic properties has experienced a recent resurgence as attested by the increased number of publications (Ekins et al., 2000; Rao et al., 2000; Williams et al., 2000). This interest is based on the idea that such approaches will add chemical knowledge to the empirical data obtained with in vitro systems. The computer-based models could enable the prediction of sub-

strates or inhibitors of specific enzymes, ensuring that only those with desirable properties are synthesized. During lead optimization, predictions of the sites of oxidation and chemical features that cause metabolic instability or increased inhibitory potency would assist medicinal chemists in making necessary chemical modifications. Efforts to develop such predictive models have intensified on cytochromes P450 (P450s) because they are an important family of drug-metabolizing enzymes.

Current models range from descriptive structure activity relationship studies (Smith et al., 1997a,b; Lewis et al., 1999), to three-dimensional quantitative structure activity relationship (3D-QSAR) models. 3D-QSAR models for CYP2B6 (Ekins et al., 1999c) and CYP3A4 (Ekins et al., 1999d) substrates have been published. For P450 inhibitors, models have been proposed for CYP2D6 (Strobl et al., 1993; Ekins et al., 1999b), CYP1A2 (Moon et al., 2000), CYP3A4 (Ekins et al., 1999a) and CYP2C9 (Jones et al., 1996b; Ekins et al., 2000; Rao et al., 2000). Other models use electronic parameters to predict likely rates and sites of oxidation (Kor-

ABBREVIATIONS: P450, cytochrome P450; 3D, three-dimensional; QSAR, quantitative structure activity relationship; CoMFA, comparative molecular field analysis; MFC, 7-methoxy-4-trifluoromethylcoumarin; HPLC, high-performance liquid chromatography; PCA, principal component analysis; GOLPE, generating optimal linear partial least-squares analysis estimations; PLS, partial least-squares analysis.

zekwa et al., 1996), two-dimensional descriptors to predict likely substrates and sites of oxidation (Lewis et al., 1999), 3D descriptors to identify substrates and sites of oxidation [CYP2D6 (de Groot et al., 1999a,b), CYP2C9 (Mancy et al., 1995; Jones et al., 1996a)] and NMR studies for CYP2C9 substrates (Poli-Scaife et al., 1997). The challenges to model P450 substrates and inhibitors include the uncertainty in human P450 active site structure, the large chemical diversity of compounds that interact with each of the P450s, and the possible activation of some P450s because of homotropic and heterotropic cooperativity (Domanski et al. 1999; Hutzler and Tracy, 2000).

With no human P450 yet crystallized, homology modeling has been performed using coordinates of bacterial P450 crystal structures (CYP101, -102, -107, and -108) (Peterson and Graham, 1998). There is only up to 19% amino acid sequence identity of these bacterial P450s with the human P450s, which makes homology modeling difficult (Ridderström et al., 2000). Recently, a mammalian membrane-bound P450 (rabbit CYP2C5) has been crystallized (Williams et al., 2000), which improves our ability to model human P450s because the sequence identity between human CYP2C9 and rabbit CYP2C5 is 77%. Efforts to generate pharmacophore models without taking into account the active site geometry and chemistry can have a number of difficulties associated with the selection of the most likely active conformers and identifying the different modes ligands can bind in the active site.

Most receptor-ligand modeling in drug design is performed using compounds with a common core structure (Kubinyi, 1997a,b). P450s, on the other hand, have the capacity to interact with compounds with extremely diverse chemical features, some of which are specific for some P450 isoforms (Lewis et al. 1999) but not easily discernible for others. Substrates and inhibitors interact with P450s in a stereospecific manner, and most of the compounds are also very flexible (with at least three rotatable bonds), which increases the number of possible conformers from which one has to select the one likely to interact with the P450. These features make

it difficult to generate models based on alignment rules such as comparative molecular field analysis (CoMFA) as previous studies have done (Jones et al., 1996b; Rao et al., 2000). To overcome this problem, efforts are being made to use alignment-independent techniques with software like ALMOND, based on GRid-INdependent Descriptors (Pastor et al., 2000) and CATALYST (Molecular Simulations, San Diego, CA).

When identifying the site of metabolism, it must be noted that the metabolism of a compound by a single P450 can occur at many sites and can also be done by other P450s and at different rates. Identification of the site of metabolism, therefore, is noninformative about the importance of that route of biotransformation. Use of the apparent K_m value, which reflects affinity but not rate, to infer metabolic stability can also be misleading. The V_{max}/K_m ratio, which estimates clearance and significance of pathway, might be a better parameter to model against. With respect to inhibitors, their interaction can be through reversible inhibition (competitive, noncompetitive, uncompetitive, or mixed type) or irreversible, such as metabolite complexation or reaction with the heme iron (Murray and Reidy, 1990). The quality of the data used in modeling is also crucial because there is great variability in kinetic constants for the same compounds between laboratories and, correspondingly, when different sources of enzyme such as recombinant P450s, human liver microsomes, hepatocytes, and liver slices are used (Boobis et al., 1998; Iwata et al., 1998; Thummel and Wilkinson, 1998; Ekins et al., 1999d). Use of such variable literature data could have contributed to previous modeling efforts that have an error in predicting K_i or K_m values of 1 logarithmic unit (Ekins et al., 1999a,b). Such deviation causes significant difference in the prediction of the biologic effects of a test compound (e.g., a K_i value of 100 μM when the actual one is 10 μM) and are therefore of limited value.

CYP2C9 is one of the most important drug-metabolizing P450s. It is responsible for the metabolism of up to 15% of currently used therapeutics. Among them are low therapeutic index drugs such as warfarin, for which drug-drug interactions can involve enzyme inhibition (Miners and Birkett, 1998). CYP2C9 also shares high sequence identity (77%) with the recently crystallized mammalian CYP2C5 and therefore a greater chance to generate a reliable protein homology model. In this study, we attempt to generate a 3D-QSAR model for CYP2C9 inhibitors covering a large chemical space, taking into account important parameters such as the mechanism of inhibition and the stereochemistry of the inhibitors.

Experimental Procedures

Materials

Chemicals. *CYP2C9* substrates and their metabolites: Diclofenac and 7-methoxy-4-trifluoromethylcoumarin (MFC) were purchased from SIGMA (St. Louis, MO) and 4'-OH-diclofenac, 7-OH-4-trifluoromethylcoumarin, and 7-OH-warfarin were purchased from Gen-Test (Woburn, MA). NADPH was bought from Sigma Aldrich Research (Stockholm, Sweden). Recombinant human CYP2C9 expressed in yeast was produced in-house (AstraZeneca R & D, Mölndal, Sweden). *CYP2C9* inhibitors: Sulfaphenazole was purchased from Ultrafine (Manchester, UK). Dicoumarol, kaempferol, phenytoin, phenylbutazone, progesterone, pyrimethamine, quercetin, quinine, and thiazendazole were purchased from Sigma. (3*R*,5*S*)-Fluvastatin sodium and (3*S*,5*R*)-fluvastatin sodium [racemate from Sandoz (Summit, NJ), enantiomers separated at Astra-

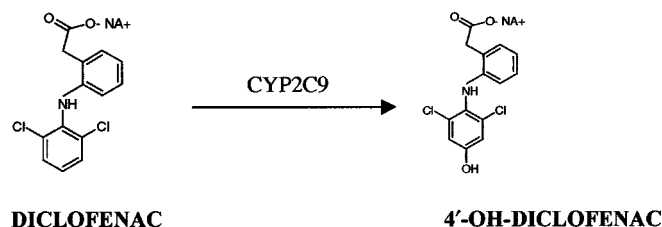
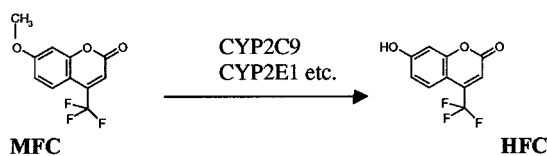


Fig. 1. Marker substrates for CYP2C9 activity, MFC (7-methoxy-4-trifluoromethyl-coumarin) demethylation and diclofenac 4'-hydroxylation

Zeneca R & D]; fluvoxamine was from Chemtronica (Ballwin, MO); (*R*)-warfarin and (*S*)-warfarin were from Ubichem (Eastleigh, UK); zafirlukast was from AstraZeneca R & D; (*R*)-omeprazole, (*R*)-pantoprazole, (*S*)-pantoprazole, (*R*)-rabeprazole, (*S*)-rabeprazole, omeprazole sulfone, spiro (fluoren-9,4-imidazolidine)-2',5'-dione (SFID), A-44338, and D-62126 were from AstraZeneca R & D; and (–)-pyranocoumarin, (+)-pyranocoumarin, (+)-miconazole, and (–) miconazole were from Sigma (enantiomers separated at AstraZeneca R & D, Sweden)

Equipment, Software, and Databases. MFC dealkylation was assessed by a fluorometric method (Bapiro et al., 2001) and diclofenac hydroxylation by a HPLC method (Masimirembwa et al., 1999). Enzyme kinetic analysis was done using GraFit 4.0.12 (Erithacus Software Limited, Middlesex, UK) and SIMFIT 5.3 (Bardsley et al., 1995). Protein homology and QSAR modeling were done on Silicon Graphics Octane and O₂ workstations, respectively (Silicon Graphics Inc., Mountain View, CA). Insight II 98.0 (Molecular Simulations Inc.) was used in the protein homology modeling. The software utilized in the computational analysis were: GOLD 1.1 (Dr Gareth Jones, University of Sheffield, U.K.), GRID (Molecular Discovery Ltd., University of Oxford, U.K.), GOLPE VOLSURF (MIA, Perugia, Italy), SYBYL 6.5.3, and CONCORD (Tripos Associates Inc., St. Louis, MO). Chemical structures were imported from ISIS-BASE database or drawn in ISIS-Draw (MDL information Systems Inc., San Leandro, CA).

Methods

Overview. The present work was initiated by performing *inhibition studies* to establish a data set of structurally diverse competitive CYP2C9 inhibitors. Simultaneously, a *protein homology model* was built based on the coordinates of the crystallized rabbit CYP2C5 structure. *Mutagenesis experiments* were performed guided by the homology model to enable validation of the final enzyme model. The homology model was then used for *active site docking* of compounds with known orientation toward the heme and the selected conformers were used as templates in a similarity analysis. The remaining compounds were then docked into the homology model and the resulting conformers were evaluated in the similarity analysis (based on grid interaction energies) to choose conformers similar to the templates. This approach enabled a *conformer selection* that constrained the conformational space to the active site of the homology model. The grid interaction fields for the selected conformers were statistically analyzed together with the experimental data to *derive a 3D QSAR model*. Data not incorporated in the model (external test set) was then used for external prediction to validate the model.

Inhibition Studies. Drug-like compounds known from the literature or suspected to interact with CYP2C9, as substrates or inhibitors, were compiled. The initial screen of 73 compounds as potential inhibitors of CYP2C9 was performed using MFC dealkylation (Fig. 1), which is a fast screening assay with good correlation with conventional HPLC assays. (Bapiro et al., 2001). The inhibition assays were done with the substrate, MFC, at K_m concentration (50 μ M). The inhibitors were serially diluted from 200 to 0.009 μ M in the 96-microtiter well plate assay design with sulfaphenazole used as the positive control inhibitor. This experimental set-up allowed for the use of the relationship $K_i = IC_{50} / 2$ to estimate inhibitory potency assuming competitive inhibition at $[S] = K_m$. Of the 73 compounds, 40 compounds with a defined stereochemistry showed inhibitory effect and were investigated further.

The mechanism of inhibition, for the 40 compounds, was determined from saturation curve measurements of CYP2C9 catalyzed diclofenac 4'-hydroxylation (Fig. 1). Diclofenac was used for the mechanistic studies instead of MFC because it is more specific for CYP2C9 and therefore a better probe for specific enzyme active-site interactions. Data were analyzed using GraFit 4.0.12 and SIMFIT 5.3. Of the 40 compounds, 29 inhibited CYP2C9 competitively with the remaining 11 displaying either uncompetitive or unclear mechanisms of inhibition. Of the 29 compounds shown to be competitive

inhibitors, 21 compounds, the training data set (Table 1), were randomly selected manually for deriving the 3D-QSAR model, and the remaining eight compounds, the external test set (Table 2), were used to validate the model.

The 29 compounds used in this study exhibit great chemical diversity ranging from rigid structures (steroids) to very flexible structures (more than three rotatable bonds) and molecular mass ranging from 200 to 575 kDa. The compounds belong to diverse chemical classes such as acids, bases, and heterocyclic amines and compound classes including flavonoids, steroids, and coumarins. The diversity of our training data set was further substantiated by comparing their 2D-descriptors in VOLSURF (using the H₂O, DRY, and O probes) with literature data sets of substrates (Miners and Birkett, 1998), competitive inhibitors (Jones et al., 1996b) and compounds believed to be substrates, inhibitors, or mentioned in the context of CYP2C9. In the principal component analysis (PCA) plot (Fig. 2), compounds used in this study (squares) are well scattered over the chemical space.

Protein Homology Modeling. The crystal structure of CYP2C5 was used as template in the modeling of CYP2C9 (Williams et al., 2000). The sequence identity between CYP2C9 and CYP2C5 is 77% (similarity, 83%) which makes CYP2C5 a good template for the modeling of CYP2C9. Table 3 shows the amino acid alignment of the 2C5_3LVN and CYP2C9 because the alignment is done to the mutated protein (see Williams et al., 2000).

The crystal structure of CYP2C5 lacks the first 29 N-terminal amino acids and has five mutations compared with the wild-type sequence. The corresponding 29 N-terminal amino acids in CYP2C9 were therefore also removed before amino acid sequence alignment of the two P450s. The modeling was carried out with the Insight II Homology module. The alignment between CYP2C9 and CYP2C5 was adjusted manually. The CYP2C9 amino acid sequence contains three additional amino acids (Lys-275, Gln-278, and Pro-279). These additional amino acids are located in the loop between helices H and I in the CYP2C9 model. The models were analyzed using Prostat in the Insight II Homology module and the SYBYL ProTable module.

Mutagenesis Experiments. The mutants were generated as described by Ridderström et al. (2000). The primers for site-directed mutagenesis used in the mutagenic PCR were (forward primers): Leu362Ala, 5'-ATACATTGACCTT**G**CCCCACCAGCCTGCC-3' and for Leu362Ile, 5'-GATACATTGACCTT**A**TCCCCACCAGCCTG-3'. The nucleotides that introduced changes in the cDNA sequence are bold and underlined. The reverse primers were complementary to the forward primers.

The CYP2C9 mutant variants were expressed in *Saccaromyces cerevisiae* INVSc1-HR as described previously (Ridderström et al., 2000). The assays for (*S*)-warfarin hydroxylation, diclofenac hydroxylation, and the HPLC-UV analysis were carried out as described previously (Jung et al., 1998; Masimirembwa et al., 1999).

Active Site Docking and Conformer Selection. Because the crystal structure of CYP2C5 on which we base our CYP2C9 homology model was crystallized with water hexacoordinated to the heme, we carried out the docking of template molecules with the water still resident in the active site. Because 'nonligand' or 'ligand-type' compounds displace water when they interact with P450s (Klaassen, 1996), we are therefore modeling for the stage just before the displacement of water as the sixth heme ligand. The distances of docked compounds' site of interaction to the heme are therefore likely to be different from when the water is displaced, a process that is probably accompanied by conformational changes. The possibility of such conformational changes is to some extent supported by the fact that we could not dock tienilic acid or diclofenac into the active site of our CYP2C9 homology model active site using the distances of these molecules to the heme as measured by NMR (Poli-Scaife et al., 1997).

The structures were imported from ISIS Base and converted into 3D structures using the CONCORD program in a Silicon Graphics O₂ UNIX workstation. Each structure was energy minimized using the Tripos force field with the Gesteiger-Marssili atom charges and

TABLE 1

Competitive inhibitors of CYP2C9 used in 3D-QSAR model

The experimental K_i values, K_i values obtained for each compound in the leave-one-out internal validation process, and the standard deviation of these predictions are indicated.

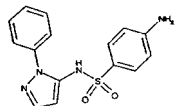
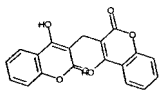
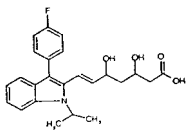
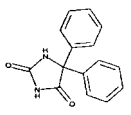
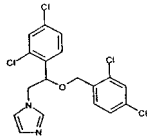
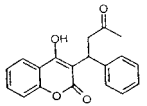
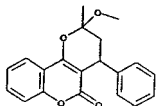
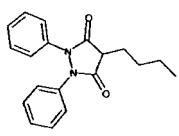
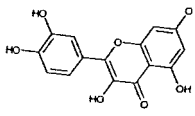
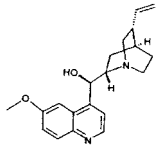
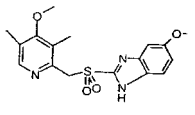
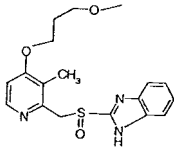
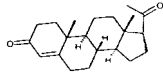
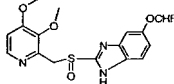
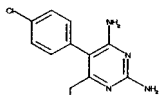
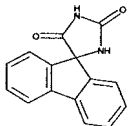
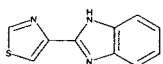
Inhibitors	K_i – exp	K_i – int	S.D. (log units)	
	μM			
Sulfaphenazole	0.5	1.0	0.21	
Dicoumarol	1.9	3.47	0.18	
(3 <i>R</i> ,5 <i>S</i>)-fluvastatin sodium (3 <i>S</i> ,5 <i>R</i>)-fluvastatin sodium	2.2 3.3	1.49 1.47	0.11 0.24	
Phenytoin	6.0	6.1	0.005	
(<i>S</i>)-Miconazole	6.0	4.97	0.06	
(<i>R</i>)-Warfarin (<i>S</i>)-Warfarin	13.6 20	15.5 18	0.04 0.03	
(3 <i>R</i> ,13 <i>S</i>)-Pyranocoumarin	16	17	0.02	
Phenylbutazone	19.3	15	0.08	
Quercetin	27	29	0.02	
Quinine	32	44	0.1	
Omeprazole sulfone	35	50	0.11	

TABLE 1
Continued

Inhibitors	K_i - exp	K_i - int	S.D. (log units)	
(<i>R</i>)-Rabeprazole (<i>S</i>)-Rabeprazole	36 37	27 60	0.09 0.15	
Progesterone	5.5	5.6	0.003	
(<i>S</i>)-Pantoprazole (<i>R</i>)-Pantoprazole	64 145	73 114	0.04 0.07	
Pyrimethamine	51.5	46	0.04	
SFID	125	106	0.05	
Thiabendazole	245	176	0.1	

in vacuo conditions. To analyze the ligand-receptor interaction, several docking experiments were performed using the program GOLD. The program was set to propose the 10 best solutions, and early termination was allowed if the first three energetically favorable ones were within 1.5 Å root mean square distance of each other. No ligand bumps were used in the fitness function settings and the default conditions were used. No constraints were used except in the case of sulfaphenazole, in which a constraint was set to guide the pyrazole nitrogen toward the iron. The template molecules were docked into an active site volume of 10 Å.

Substrates and nonsubstrate inhibitors of CYP2C9 whose site of oxidation and/or mode of interaction with the heme was known from published works were used as templates in the selection of conformers for the compounds whose interaction with active site were not known. The positions of hydroxylation at CYP2C9 have been established for (*S*)-warfarin, phenytoin, and progesterone (Miners and Birkett, 1998; He et al., 1999; Payne et al., 1999). Of the 10 best solutions found by GOLD for each of these inhibitory substrates, the conformers with the hydroxylation site in close proximity to the heme were chosen (Fig. 2). The distances between the water and the 7-position of (*S*)-warfarin, progesterone, and phenytoin were 2.79 Å, 3.62 Å, and 3.02 Å, respectively. For the ligand-type inhibitor sulfaphenazole, solutions with the ligating nitrogen closest to the heme were chosen. In this study, we chose the nitrogen in the pyrazole ring at a distance of 3.98 Å based on a better alignment with the other template molecules. There are, however, two studies giving different suggestions as to which nitrogen of sulfaphenazole ligates with the heme. Mancy et al. (1995) postulate that it is the terminal anilinic nitrogen whereas Rao et al. (2000) proposed the pyrazole ring nitrogen.

Derivation of a 3D-QSAR Model. The grid interaction field (Goodford, 1985) for all 155 GOLD solutions for all compounds docked into the active site were calculated with the DRY (for hydrophobic interactions) and the OH (for polar interactions) probe in a

grid box extending 4 Å beyond all molecules. The grid interaction energies were calculated as defined in the program and imported into GOLPE [Generating Optimal Linear Partial Least-Squares Analysis (PLS) Estimations] (Massimo et al., 1993), a chemometric toolbox for 3D-QSAR enabling statistical analysis such as PCA and PLS analysis (Wold et al., 1983, 1987) for data sets with thousands of variables. The x-matrix was pretreated by deleting the points with a standard deviation lower than 0.02 kcal/mol and with an energy of interaction lower than 0.02 kcal/mol. A PCA analysis was performed for all docked conformers and a clustering of one or few conformers of each compound close to the selected template ones for (*S*)-warfarin, phenytoin, progesterone, and sulfaphenazole was identified in the score plot by a similarity analysis. The distances in the four principal component score plot of each GOLD conformer solution toward the selected template molecules were calculated and used as the similarity index. The one with the shortest distance to the templates was chosen as the "active conformer" for each compound and was used for further analysis.

The grid interaction fields were recalculated using the DRY probe and the OH probe in a box that extended 4 Å beyond all ligands. The results were imported into GOLPE together with the K_i values, obtaining a two-block matrix, one block for the DRY and the other for the OH probe. A pretreatment was performed in the independent variables by scaling them using a block unscaled weights, so each block (probe) had the same variance in the initial model; also, any variable with a standard deviation lower than 0.02 kcal/mol and all (positive and negative) interactions with an absolute value lower than 0.02 kcal/mol were removed, together with the n -variables that were in 2 or 3 levels. After these pretreatments, 7600 active variables were obtained and the data set was divided into the training set (21 compounds) and the validation set (8 compounds). To predict the K_i value for the validation set, it is important that both the training and the validation sets have the same variable pretreatment.

The PLS modeling with a variable selection process was then performed in three steps: First, D-optimal designs (Baroni et al. 1993) were applied to the PLS loading space. This is a statistical method that enables the selection of variables that are the most spread and the least complementary to each other to increase the signal-to-noise ratio. This process was run using the two first components in the PLS model in an iterative manner and reducing 50% of the variables in each iteration. Two D-optimal design iterations were used, yielding 1900 variables. Second, a smart region definition selection (Pastor et al., 2000) was made with a critical distance set to 2.0 grid units (2 Å) and the collapsing distance to 25.0 grid units (25 Å). The smart region definition is a mathematical technique that groups variables that are correlated to each other and close in the space. The third step was a fractional factorial design variable selection. The variables were selected to increase the predictive power of the model and the technique used fractional factorial design matrix and Student *t* test cut-off values (Massimo et al., 1993). The resulting model had one principal component and a correlation coef-

ficient (r^2) of 0.947 and a predictive power (q^2) of 0.730. The internal cross-validation of the model was done using the leave one out approach, in which one object at a time is excluded from the model and predicted, the procedure is repeated until all objects have been held out once. A more rigorous measurement of internal predictive power was also performed by leaving out 20% and 50% of the compounds, producing q^2 values of 0.672 and 0.497, respectively. The external prediction capabilities of the model were checked by the eight-compound validation set, giving a maximum error of 0.36 log units.

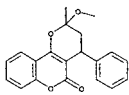
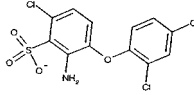
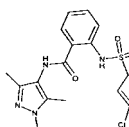
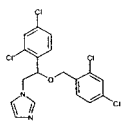
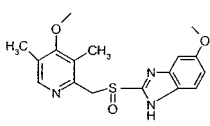
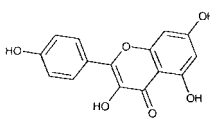
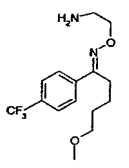
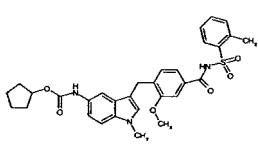
Results and Discussion

This study describes the generation of a 3D-QSAR model for CYP2C9 inhibitors using a combination of QSAR and protein homology modeling. The model captures the complexity of enzyme-inhibitor interactions by taking into account

TABLE 2

Competitive inhibitors of CYP2C9 used to validate the 3D-QSAR model

The experimental K_i values, K_i values obtained for each compound in the external prediction validation process, and the standard deviation of these predictions are indicated.

Validation Data Set	K_i - exp	K_i - ext	S.D. (log units)	
	μM			
(3 <i>R</i> ,13 <i>S</i>)-Pyranocoumarin	16	28	0.17	
A-44338	20	16	0.07	
D-62126	17	11	0.13	
(<i>R</i>)-Miconazole	6.0	7.46	0.07	
(<i>R</i>)-Omeprazole	45	47	0.01	
Kaempferol	6.0	18	0.34	
Fluvoxamine	8.5	21	0.28	
Zafirlukast	2.5	8	0.35	

the mechanism of inhibition and the issue of active conformer selection. The good quality ($r^2 = 0.947$, $q^2 = 0.730$) and capacity to predict external data sets within 0.5 logarithmic units of the experimental value attest to the validity of our 3D-QSAR modeling approach.

CYP2C9 Inhibitor Data Set. Starting with a set of 73 compounds and ending up with 29 compounds suitable for use in the modeling work means that to get more compounds in the data set, one needs to screen a large number of compounds. Of the 40 compounds that inhibited CYP2C9, 29 were competitive inhibitors and 11 were either uncompetitive or of unclear mechanism of inhibition. The model developed for competitive inhibitors presented here can only be used for inhibitors with defined types of inhibition. This is valuable information in both modeling work and application of models to predict biological effect. The use of literature inhibition data of unspecified mechanism or K_i values approximated from IC_{50} determinations should be practiced with caution.

This modeling work describes a procedure to find an align-

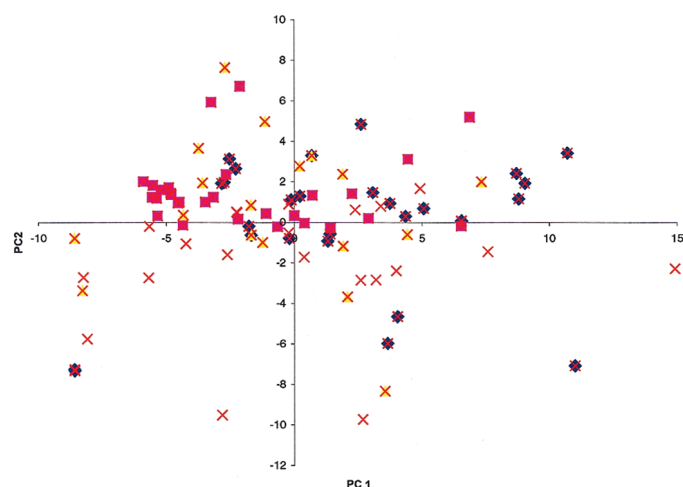


Fig. 2. Distribution of model compounds in chemical space explained in two components for VOLSURF 2D-descriptors calculated using H_2O , DRY, and O probes. ◆, Model compounds; ■, Jones et al. (1996b); △, Miners and Birkett (1998); ×, D-propoxyphene, amiodarone, clozapine, diethylstilbestrol, flecainide, fluconazole, ketoconazole, lovastatin, mestranol, norfluoxetine, paroxetine, pravastatin, simvastatin, and trimethoprim, and ◆, ■, and △.

TABLE 3

Amino acid sequence alignment for 2C5_3LVN and CYP2C9

CYP2C9	30	PPGPTPLPVI	GNILQIGIKD	ISKSLTNLSK	VYGPVFTLYF	GLKPIVVLHG
2C5_3LVN	30	PPGPTPPPII	GNILQIDAKD	ISKSLTKFSE	CYGPVFTVYL	GMKPTVVLHG
CYP2C9	80	YEAVKEALID	LGEEFSGRGI	FPLAERANRG	FGIVFVSNKK	WKEIRRFSLM
2C5_3LVN	80	YEAVKEALVD	LGEEFAGRGS	VPILEKVSKE	LGIAFSNAKT	WKEMRRFSLM
CYP2C9	130	TLRNFVGMGKR	SIEDRVQEEA	RCLVEELRKT	KASPCDPTFI	LGCAPCNVIC
2C5_3LVN	130	TLRNFVGMGKR	SIEDRIQEEA	RCLVEELRKT	NASPCDPTFI	LGCAPCNVIC
CYP2C9	180	SIIFHKRFDY	KDQFLNLME	KLNENIKILS	SPWIQICNNF	SPIIDYFPPT
2C5_3LVN	180	SVIFHNRFDY	KDEEFLKLM	SLHENVELLG	TPWLQVYNNF	PALLDYFPGI
CYP2C9	230	HNKLLKNVAF	MKSYLEKVK	EHQESMDMNN	PQDFIDCFIL	KMEKEKHNQP
2C5_3LVN	230	HKTLKNADY	IKNFIMEKVK	EHQKLLDVNN	PRDFIDCFIL	KMEQE--
CYP2C9	280	SEFTIESLEN	TAVDLFGAGT	ETTSTTLRYA	LLLLLKHPEV	TAKVQEEIER
2C5_3LVN	277	LEFTLESLEVI	AVSDFGAGT	ETTSTTLRYS	LLLLLKHPEV	AARVQEEIER
CYP2C9	330	VIGNRSPCM	QDRSHMPYTD	AVVHEVQRYI	DLLPTSLPHA	VTCDIKFRNY
2C5_3LVN	327	VIGRHRSPCM	QDRSRMPYTD	AVIHEIQRFI	DLLPTNLPHA	VTRDVFRNY
CYP2C9	380	LIPKGTTLILI	SLTSVLHDK	EFPNPEMFDP	HHFLDEGGNF	KKSKYFMPFS
2C5_3LVN	377	FIPKGTDLIIT	SLTSVLHDEK	AFPNPKVFDP	GHFLDESGNF	KKSDYFMPFS
CYP2C9	430	AGKRICVGEA	LGMELFLFL	TSILQNFNLK	SLVDPKNLDT	TPVNGFASV
2C5_3LVN	427	AGKRMVGEA	LARMELFLFL	TSILQNFKLQ	SLVEPKDLDI	TAVVNGFVSV
CYP2C9	480	PPFYQLCFIP	V			
2C5_3LVN	477	PPSYQLCFIP	IHH--			

ment although the compounds might have different binding modes in the CYP2C9 active site despite an identical mechanism of inhibition (competitive inhibition). In this study (as well as that of Rao et al., 2000) the stereochemistry of the molecules was taken into account by only including nonchiral compounds and enantiomers of chiral compounds.

The inhibitors included in our model include very potent inhibitors ($K_i < 1 \mu M$; $n = 1$), potent inhibitors ($K_i = 1-10 \mu M$; $n = 6$), medium potency inhibitors ($K_i = 10-20 \mu M$ ($n = 4$), and poor inhibitors $K_i > 20 \mu M$ ($n = 10$) when classified according to their potential to cause clinically significant inhibitory effects. It would have been ideal to have at least 10 compounds in each category, which also exhibit wide intra-category spread of values. To obtain such a data set, a very large number of compounds need to be screened. This will probably be possible in the near future using high-throughput screening procedures. Data used in this study was generated using uniform conditions to avoid the sometimes more than 10-fold variations in K_i values for the same compounds generated in different laboratories or using different sources of enzyme (Thummel and Wilkinson, 1998).

CYP2C9 Homology Model. A CYP2C9 homology model was constructed using coordinates of the recently crystallized mammalian CYP2C5 (Williams et al., 2000). The root mean square distance in the polypeptide backbone of the CYP2C9 model was 0.18 Å in relation to the CYP2C5 crystal structure. SYBYL ProTable was used to calculate the φ - ψ angles for the Ramachandran plot. Most of the φ - ψ angles (83%) were located in the core region of the plot. The MatchMaker average energy score for the model was -0.09 kT (for the CYP2C5 template, -0.13 kT). Comparison of the MatchMaker energy graphs for the model and the template did not reveal any poorly modeled regions. Analysis of the dihedral angles ω showed that two angles were larger than 5 S.D. from the reference value, where one of these ω angles was correspondingly high in the CYP2C5 template. Because CYP2C9 and CYP2C5 share 77% amino acid identity and 83% amino acid similarity, this model is a substantial improvement compared with the model of CYP2C9 based on the bacterial CYP102 (Ridderström et al., 2000).

Mutagenesis Data. The rationale for making Leu362Ile and Leu362Ala mutants was 2-fold: in the protein homology model, leucine points into the active site, and in the same

position, the isoform CYP2C19, which has different substrate specificity, has an isoleucine. These mutations resulted in increased K_m and reduced V_{max} values of CYP2C9 toward diclofenac (Table 5). The alanine mutant had a K_m value eight times higher than that of the wild-type enzyme. The isoleucine and alanine mutations also resulted in reduced (*S*)-warfarin 7-hydroxylation. The Leu362Ile and Leu362Ala mutations had 90 and 20%, respectively, of the wild-type activity. The effects of the mutations on CYP2C9 function are in line with the QSAR model of CYP2C9-inhibitor interactions where hydrophobic interactions at that site contribute to more potent inhibitory effects (see section below).

3D-QSAR Model. The 3D-QSAR model should have a bearing on appropriate amino acids in the active site of the CYP2C9 model to make biochemical sense. The examination of the PLS coefficients of the grid interaction fields, proposed for the model, in comparison with the active site of the homology model showed credible interaction patterns. The amino acids directed into the active site have representatives from all substrate recognition sites SRS 1 to 6 except for SRS 3 (Gotoh, 1992; Williams et al., 2000). Of the 17 amino acids present in the active site, two were polar (Thr-301 and Thr-304), two were acidic (Asp-293 and Glu-300), and the rest were hydrophobic (Table 4). The hydrophobic grid interac-

TABLE 4

Amino acids present in the active site of the CYP2C9 homology model are listed under active site.

The residues involved in interactions with the pharmacophore model are listed under model. Mutagenesis experiments lists mutations of amino acids and the effect of the mutations upon activity, indicated by arrows.

Region	Amino Acids			
	Active site	Model	Mutagenesis Experiments	
B-C loop	Leu-102	Leu-102	Leu-102 Ala ^a	Activity ↓
SRS-1	Val-113	Val-113	Val-113 Leu ^b	Activity ↓
	Phe-114	Phe-114	Phe-114 Leu ^b	Activity ↓
F-helix	Leu-201	Leu-201		
SRS-2	Ile-205	Ile-205		
	Gln-214			
I-helix	Asp-293	Asp-293		
SRS-4	Gly-296	Gly-296		
	Ala-297	Ala-297		
	Glu-300	Glu-300		
	Thr-301	Thr-301		
	Thr-304	Thr-304		
SRS-5	Leu-361	Leu-361		
	Leu-362	Leu-362	Leu-362 Ala, Ile ^c	Activity ↓
	Leu-366	Leu-366		
B2	Phe-476	Phe-476		
SRS-6	Ala-477	Ala-477		

^a Ridderström et al. (2000)

^b Haining et al. (1998)

^c Table 4

tions of the PLS coefficients (PC1) are shown in Fig. 4. In the **SRS-1** region (B-C loop), Phe-114 shows a considerable contribution consistent with mutagenesis data (Haining et al., 1999) and could provide π - π stacking interactions with ligands. It contributes to a hydrophobic pocket for the most potent inhibitor, sulfaphenazole, together with Leu-366 (**SRS-5**) and Phe-476 (**SRS-6**). Leu-201 and Ile-205 (**SRS-2**) in the F-helix constitute another hydrophobic pocket for interaction with sulfaphenazole. The backbone of Ile-205 that is in close proximity probably forms a hydrogen bond to the anilino nitrogen of sulfaphenazole. (*S*)-Warfarin, which has a higher K_i value, does not show this interaction. The amino acid Ala-297 in region **SRS-4** forms a third hydrophobic pocket together with Val-113 (**SRS-1**) and Leu-366. A direct hydrophobic interaction is suggested from Leu-362 (**SRS-5**) toward both (*S*)-Warfarin and sulfaphenazole and is strongly supported by the mutagenesis data (Table 4). The Ala-477 in the **SRS-6** region (β 2) also shows very strong hydrophobic interactions and could be considered a mutagenesis target for SAR analysis. Leu-102 builds a fourth hydrophobic pocket together with Val-292 in the I helix. The polar grid interactions determined by the OH probe (Fig. 5) showed directionally dependent favorable interactions with Asp-293 in SRS-4 and for the more potent inhibitors also Asp-204 in the F-helix (SRS-2). An unexpected interaction between a methyl group and the OH probe was observed. The amino acid associated with this interaction was shown to be threonine (Thr 301), which could explain the contradictory finding, in that the interaction could be with the oxygen of threonine due to amino acid conformational flexibility.

For the template molecules, 'nonligand' and 'ligand' type inhibitors, we get very good alignments (Fig. 3) that are also consistent with mutagenesis data discussed above. We therefore think that our working hypothesis of modeling for interactions just before water displacement, which is consistent with using the static model of the available crystal structure with water as the sixth ligand, captures inhibitor-enzyme interactions that determine inhibitory potency. The fact that we also get a robust model also reassures us that our hypothesis is reasonable. Understanding the dynamic changes that occur when water is displaced, associated with NMR data for substrates (Poli-Scaife et al., 1997) and the ligation of the nitrogen with the heme (Mancy et al., 1995), will only be possible when there are crystal structures of the CYP2C9 with and without inhibitors.

Despite the structural diversity of the data set (Fig. 2), the grid energy interactions for the conformers of each molecule are sufficiently described in one component, giving a good

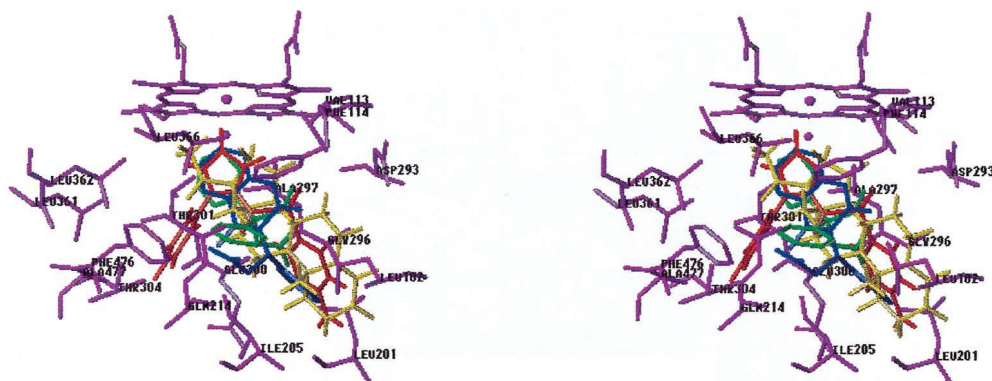


Fig. 3. Alignment of (*S*)-warfarin (blue), phenytoin (green), progesterone (yellow), and sulfaphenazole (red) into the active site of CYP2C9.

model with $r^2 = 0.948$ (Fig. 6) and an internal validation of $q^2 = 0.730$. Internal data was predicted within an S.D. of 0.24 logarithm units and external data set within 0.36 logarithm units from the experimental value (Table 1 and 2). The alignment approach for this model allows for variability in the importance of each pharmacophoric feature for different compounds as long as energy interactions are within the set that describes the molecules used in the model. This overcomes the problem of finding an alignment dependent model with fixed pharmacophoric features of fixed importance that might not be realistic when dealing with P450s. Although the predictions are very satisfactory, predictions outside the range defined by the K_i values in the model are uncertain. This is particularly so with respect to very potent inhibitors and poor inhibitors due to scarce information in our model.

Jones et al. (1996b) and Rao et al. (2000) have published CYP2C9 inhibitor models based on CoMFA. These models were based on a data set of coumarins and sulfaphenazole-like compounds and the researchers were able to find alignment rules and generate models with good predictive power. Compounds in our study are well spread over the chemical space, whereas those used by Jones et al. are confined to two quadrants of the first two components of the PCA space (Fig. 2). We could not find any alignment rules for our training data set using DISCO. Compared with the CoMFA models, our 3D-QSAR study provides a way to tackle the alignment problems associated with chemically diverse data sets. Although Rao et al. (2000) combined the pharmacophore and protein modeling, in our study we had the advantage of having a closely related P450 crystal structure to model the

active site of CYP2C9. Table 6 shows the results we obtained using our model to predict the K_i values of the molecules in Jones et al. (1996b). Compounds with an experimental K_i value below $1.0 \mu\text{M}$ were not included because our model has an endpoint at 0.5. We were able to predict 19 of 24 compounds within 0.5 log units and none of the other five were above 1.0 log units (Table 6). It should be noted that these are literature data and that the worst predicted compound (number 10) has a K_i value of $1.0 \mu\text{M}$, which is near the endpoint of our model where the information is scarce. Our results are comparable with those obtained by Jones et al. (1996b). Their mean S.D. error of *cross validation* was between 0.21 and 0.28 for the four different models for comparable compounds, whereas our mean S.D. error of *prediction* is 0.34. This further validates the robustness of our model and also the good quality of data generated under uniform conditions in Jones et al. (1996b) study.

Ekins et al. (1999a,b,c,d) used CATALYST to generate a model for CYP2C9 inhibitors. The CATALYST is a nonalignment dependent program for generating 3D pharmacophore models. The model generated did not have very good predictive power (predictive to 1.0 log unit of actual values). We believe that for models to be of any value, they should predict to within at least 0.5 log units of the actual value, otherwise one obtains values that are associated with significantly different biologic effects. Although CATALYST probably represents a powerful modeling tool to handle chemically diverse data sets, the use of highly variable literature data might have contributed to its failure to produce robust models. We think, therefore, that our approach represents a significant

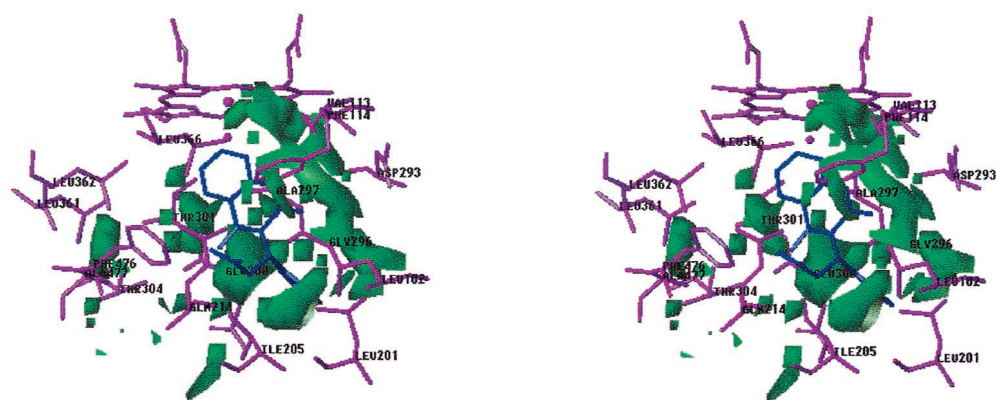


Fig. 4. PLS coefficients for the DRY probe at -0.0006 level (green). Because the DRY interactions had only negative energy of interaction, a negative coefficient region is interpreted as positive contribution to the activity.

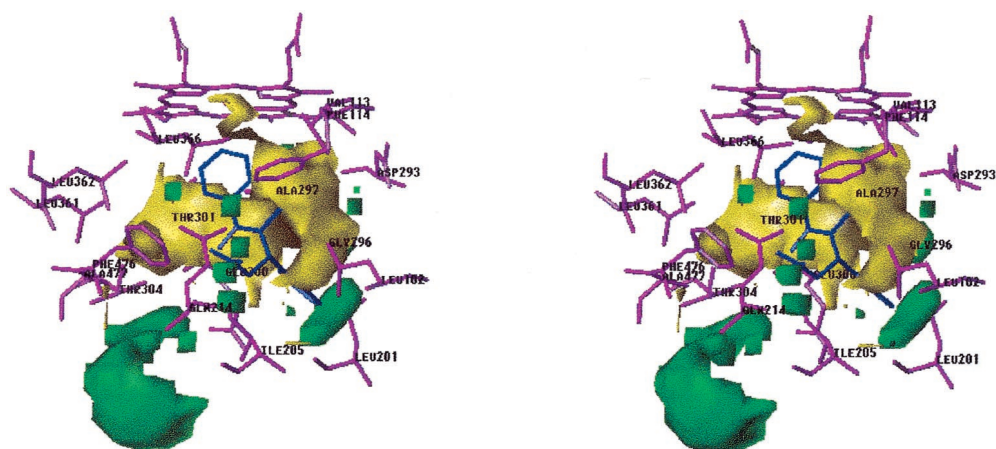


Fig. 5. PLS coefficients for the OH probe at -0.0004 level (green) and $+0.0012$ level (yellow). Because the grid interaction field had positive and negative energies, this PLS coefficient field can not be directly interpreted as regions with positive or negative contribution to the activity.

TABLE 5

Kinetic parameters for the formation of 4'-hydroxydiclofenac measured with CYP2C9 wild-type and mutants

P450	K_M	V_{max}
	μM	$nmol/min/nmol$
CYP2C9 wild-type	3.6 ± 0.4	41
Leu362Ala	30 ± 7	13
Leu362Ile	4.0 ± 0.4	31

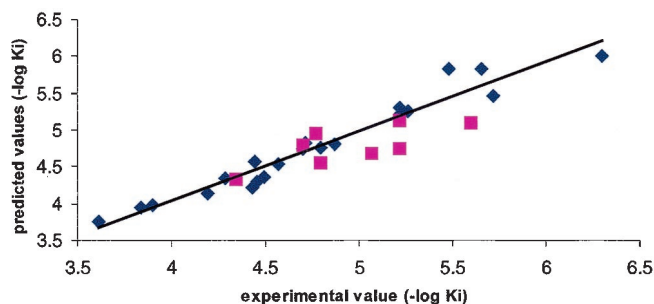


Fig. 6. Correlation between model and predicted compounds in logarithmic units $r^2 = 0.947$ and $q^2 = 0.730$. \blacklozenge , training set; \blacksquare , external test set.

TABLE 6

Competitive inhibitors of CYP2C9 used in 3D-QSAR model of Jones et al (1996)

The experimental K_i values, K_i values obtained from external prediction in our model, and the standard deviation of these predictions are indicated.

Compound Jones et al (1996)	Exp. value	External prediction (our model)	S.D. (log units)
		μM	
1	7	12	0.16
2	4	19	0.49
3	20	19	0.04
4	12	21	0.25
5	10	7	0.03
6	7	10	0.26
11	1	9	0.66
12	14	19	0.10
13	37	19	0.21
14	10	9	0.02
15	11	11	0.01
16	29	20	0.12
17	21	8	0.31
18	13	11	0.06
19	17	10	0.18
20	14	20	0.11
21	8	12	0.12
22	10	14	0.11
23	70	15	0.47
24	11	60	0.52
25	45	15	0.34
26	15	8	0.20
27	50	4	0.79
28	2	7	0.39

contribution toward 3D-QSAR modeling efforts in the P450 field and CYP2C9 in particular.

Acknowledgments

We thank our colleagues Hanna Nelander, for separation of chiral compounds, and Marie Ahlström, for technical assistance.

References

Bardsley WG, Bukhari NAJ, Ferguson MWJ, Cachaza JA and Burguillo FJ (1995) Evaluation of model discrimination, parameter estimation and goodness of fit in nonlinear regression problems by test statistics distributions. *Comput Chem* **19**: 75–84.

Bapiro TE, Egnell AC, Hasler JA and Masimirembwa CM (2001) Application of higher throughput screening (HTS) inhibition assays to evaluate the interaction of antiparasitic drugs with cytochrome P450s. *Drug Metab Dispos* **29**:1–7

Baroni M, Clementi S, Cruciani G, Kettaneh-Wold N and Wold S (1993) D-optimal designs in QSAR. *Quant Struct Act Relat* **12**:225–231.

Boobis AR, McKillop D, Robinson DT, Adams DA and McCormick DJ (1998) Inter-laboratory comparison of the assessment of P450 activities in human hepatic microsomal samples. *Xenobiotica* **28**:493–506

de Groot MJ, Ackland M, Horne V, Alexander A and Barry J (1999a) A novel approach to predicting P450 mediated drug metabolism. CYP2D6 catalyzed N-dealkylation reactions and qualitative metabolite predictions using a combined protein and pharmacophore model for CYP2D6. *J Med Chem* **42**:4062–4070.

de Groot MJ, Ackland M, Horne V, Alexander A and Barry J (1999b) Novel approach to predicting P450 mediated drug metabolism: Development of combined protein and pharmacophore model for CYP2D6. *J Med Chem* **42**:1515–1524.

Domanski TL, He YA, Harlow GR and Halpert J (1999) Dual role of human cytochrome P450 3A4 residue Phe-304 in substrate specificity and cooperativity. *J Pharmacol Exp Ther* **293**:585–591

Ekins S, Bravi G, Binkley S, Gillespie JS, Ring B, Wikel J and Wrighton S (2000) Three- and four-dimensional-quantitative structure activity relationship (3D/4D-QSAR) analyses of CYP2C9 inhibitors. *Drug Metab Dispos* **28**:994–1002.

Ekins S, Bravi G, Binkley S, Gillespie JS, Ring BJ, Wikel JH and Wrighton SA (1999a) Three- and four-dimensional quantitative structure activity relationship analyses of cytochrome P-450 3A4 inhibitors. *J Pharmacol Exp Ther* **290**:429–438.

Ekins S, Bravi G, Binkley S, Gillespie JS, Ring BJ, Wikel JH and Wrighton SA (1999b) Three and four dimensional-quantitative structure activity relationship (3D/4D-QSAR) analyses of CYP2D6 inhibitors. *Pharmacogenetics* **9**:477–489.

Ekins S, Bravi G, Ring BJ, Gillespie TA, Gillespie JS, Vandenbranden M, Wrighton SA and Wikel JH (1999c) Three-dimensional quantitative structure activity relationship analyses of substrates for CYP2B6. *J Pharmacol Exp Ther* **288**:21–29.

Ekins S, Bravi G, Wikel JH and Wrighton SA (1999d) Three-dimensional-quantitative structure activity relationship analysis of cytochrome P-450 3A4 substrates. *J Pharmacol Exp Ther* **291**:424–433.

Goodford PJ (1985) A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J Med Chem* **28**:849–857.

Gotoh O (1992) Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *J Biol Chem* **267**:83–90.

Haining RL, Jones JP, Henne KR, Fisher MB, Koop DR, Trager WF and Rettie AE (1999) Enzymatic determinants of the substrate specificity of CYP2C9: Role of B'-C loop residues in providing the π - π stacking anchor site for warfarin binding. *Biochemistry* **38**:3285–3292.

He M, Korzekwa KR, Jones JP, Rettie AE and Trager WF (1999) Structural forms of phenprocoumon and warfarin that are metabolized at the active site of CYP2C9. *Arch Biochem Biophys* **372**:16–28.

Hutzler JM and Tracy TS (2000) Activation of CYP2C9 by dapsone and analogs. ISSX meeting; Indianapolis, IN; 2000 Oct 24–28.

Iwata H, Fujita K, Kushida H, Suzuki A, Konno Y, Nakamura K, Fujino A and Kamataki T (1998) High catalytic activity of human cytochrome P450 co-expressed with human NADPH- cytochrome P450 reductase in *Escherichia coli*. *Biochem Pharmacol* **55**:1315–1325.

Jones BC, Hawksworth G, Horne VA, Newlands A, Morsman J, Tute MS and Smith DA (1996a) Putative active site template model for cytochrome P4502C9 (tolbutamide hydroxylase). *Drug Metab Dispos* **24**:260–266.

Jones JP, He M, Trager WF and Rettie AE (1996b) Three-dimensional quantitative structure-activity relationship for inhibitors of cytochrome P4502C9. *Drug Metab Dispos* **24**:1–6.

Jung F, Griffin KJ, Song W, Richardson TH, Yang M and Johnson EF (1998) Identification of amino acid substitutions that confer a high affinity for sulfa-phenazole binding and a high catalytic efficiency for warfarin metabolism to P450 2C19. *Biochemistry* **37**:16270–16279.

Klaassen CD (1996) *Casarett and Doull's Toxicology: The Basic Science of Poisons*, 5th ed., pp. 141, Culinary and Hospitality Industry Publications Services, Weimar, TX.

Korzekwa KR, Grogan J, DeVito S and Jones JP (1996) Electronic models for cytochrome P450 oxidations. *Adv Expl Med Biol* **387**:361–369.

Kubinyi H (1997a) QSAR and 3D QSAR in drug design. Part 1: Methodology. *Drug Discov Today* **2**:457–467.

Kubinyi H (1997b) QSAR and 3D QSAR in drug design Part 2: Applications and problems. *Drug Discov Today* **2**:538–546.

Lewis DF, Dickins M, Eddershaw PJ, Tarbit MH and Goldfarb PS (1999) Cytochrome P450 substrate specificities, substrate structural templates and enzyme active site geometries. *Drug Metabol Drug Interact* **15**:1–49.

Mancy A, Broto P, Dijols S, Dansette PM and Mansuy D (1995) The substrate binding site of human liver cytochrome P450 2C9: An approach using designed tienilic acid derivatives and molecular modeling. *Biochemistry* **34**:10365–10375.

Masimirembwa CM, Otter C, Berg M, Jönsson M, Leidvik B, Jonsson E, Johansson T, Bäckman A, Edlund A and Andersson TB (1999) Heterologous expression and kinetic characterization of human cytochromes P-450: Validation of a pharmaceutical tool for drug metabolism research. *Drug Metab Dispos* **27**:1117–1122.

Massimo B, Cruciani G, Costantino G, Riganelli D, Valigi R and Clementi S (1993) Generating optimal linear PLS estimations (GOLPE): An advanced chemometric tool for handling 3D-QSAR problems. *Quant Struct-Activity Relat* **12**:9–20.

Miners JO and Birkett DJ (1998) Cytochrome P4502C9: An enzyme of major importance in human drug metabolism. *Br J Clin Pharmacol* **45**:525–538.

Moon T, CHI MH, Kim DH, Yoon CN, Choi YS (2000) Quantitative structure-activity relationships (QSAR) study of flavonoid derivatives for inhibition of cytochrome P450 1A2. *Quant Struct-Activity Relat* **19**:257–263.

Murray M and Reidy GF (1990) Selectivity in the inhibition of mammalian cytochromes P-450 by chemical agents. *Pharmacol Rev* **42**:85–101.

Pastor, Cruciani G, McLay I, Pickett S and Clementi S (2000) GRIND-INdependent Descriptors (GRIND): A novel class of alignment-independent three-dimensional molecular descriptors. *J Med Chem* **17**:3233–3243.

- Payne VA, Chang YT and Loew GH (1999) Homology modeling and substrate binding study of human CYP2C9 enzyme. *Proteins* **37**:176–190.
- Peterson JA and Graham SE (1998) A close family resemblance: The importance of structure in understanding cytochromes P450. *Structure* **6**:1079–1084.
- Poli-Scaife S, Attias R, Dansette PM and Mansuy D (1997) The substrate binding site of human liver cytochrome P450 2C9: An NMR study. *Biochemistry* **36**:12672–12682.
- Rao S, Aoyama R, Schrag M, Trager WF, Rettie A and Jones PJ (2000) A refined 3-dimensional QSAR of cytochrome P450 2C9: Computational predictions of drug interactions. *J Med Chem* **43**:2789–2796.
- Ridderström M, Masimirembwa C, Trump-Kallmeyer S, Ahlefeldt M, Otter C and Andersson TB (2000) Arginines 97 and 108 in CYP2C9 are important determinants of the catalytic function. *Biochem Biophys Res Commun* **270**:983–987.
- Smith DA, Ackland MJ and Jones BC (1997a) Properties of cytochrome P450 isoenzymes and their substrates. Part 1: Active site characteristics. *Drug Discov Today* **2**:406–414.
- Smith DA, Ackland MJ and Jones BC (1997b) Properties of cytochrome P450 isoenzymes and their substrates. Part 2: Properties of cytochrome P450 substrates. *Drug Discov Today* **2**:479–485.
- Strobl GR, von Kruedener S, Stockigt J, Guengerich FP and Wolff T (1993) Development of a pharmacophore for inhibition of human liver cytochrome P-450 2D6: Molecular modeling and inhibition studies. *J Med Chem* **36**:1136–1145.
- Thummel KE and Wilkinson GR (1998) In vitro and in vivo drug interactions involving human CYP3A. *Annu Rev Pharmacol Toxicol* **38**:389–430.
- Williams PA, Cosme J, Sridhar V, Johnson EF and McRee DE (2000) Mammalian microsomal cytochrome P450 monooxygenase: Structural adaptations for membrane binding and functional diversity. *Mol Cell* **5**:121–131.
- Wold S, Albano C, Dunn WJ, Edlund U, Esbensen K, Geladi P, Helberg S, Johansson E, Lindberg W and Sjöström M (1983) Multivariate data analysis in chemistry, In *Chemometrics Mathematics and Statistics in Chemistry* (Kowalsky BR ed) pp. 17–96, Reidel, Dordrecht, Holland.
- Wold S, Esbensen K and Geladi P (1987) Principal component analysis. *Chem Intell Lab Syst* **2**:37–52.

Send reprint requests to: Collen M. Masimirembwa, Department of DMPK & Bioanalytical Chemistry, AstraZeneca R & D Mölndal, S-431 83 Mölndal, Sweden. E-mail: collen.masimirembwa@astrazeneca.com
