Development of a Combined Protein and Pharmacophore Model for Cytochrome P450 2C9

Marcel J. de Groot,*[,†,‡] Alexander A. Alex,[†] and Barry C. Jones[‡]

Department of Molecular Informatics, Structure & Design and Department of Pharmacokinetics, Dynamics & Metabolism, Pfizer Global Research & Development, Sandwich Laboratories, Ramsgate Road, Sandwich, Kent CT13 9NJ, U.K.

Received October 29, 2001

A combined protein and pharmacophore model for cytochrome P450 2C9 (CYP2C9) has been derived using various computational chemistry techniques. A combination of pharmacophore modeling (using 31 metabolic pathways for 27 substrates), protein modeling (using the rabbit CYP2C5/3 crystal structure), and molecular orbital calculations was used to derive a model that incorporated steric, electronic, and chemical stability properties. The initial pharmacophore model (based on a subset of 17 metabolic pathways for 16 substrates) and the protein model used to construct the combined model were derived independently and showed a large degree of complementarity. The combined model is in agreement with experimental results concerning the substrates used to derive the model and with site-directed mutagenesis data available for CYP2C9. The model has been successfully used to predict the metabolism of substrates not used to construct the model, of which four examples are discussed in detail. The model has also been successful in explaining the differences in substrate specificity between CYP2C9 and CYP2C19.

Introduction

Cytochromes P450 (P450s) constitute a large superfamily of heme-containing enzymes capable of oxidizing and reducing a variety of substrates of both endogenous and exogenous origin. Theoretical models that can predict the possible involvement of P450s in the metabolism of drugs or drug candidates are important tools for use in drug discovery and development. An overview of recently derived models for cytochromes P450 can be found in several recent reviews.[1,2]

CYP2C9 was identified as one of the most important P450s involved in human drug metabolism,[3] thus highlighting the obvious advantage of predictive models for this enzyme. To determine the nature of the protein sites responsible for the interaction with the substrates in the pharmacophore models, a three-dimensional representation of the entire active site of CYP2C9 is required. Homology modeling has been used to develop models of P450s for which sequence information is available, but X-ray structures are sparse. These models have to be verified by either X-ray structures or supported by SDM (site-directed mutagenesis) experiments.[4] An increasing number of crystal structures of soluble bacterial P450s is becoming available, e.g., CYP101 ($P450_{cam}$),[5−10] CYP102 ($P450_{BM3}$),[11,12] CYP105 ($P450_{SCA2}$),[13] CYP107A ($P450_{eryF}$),[14,15] and CYP108 ($P450_{terp}$),[16] as well as the structures of the soluble eukaryotic CYP55 ($P450_{nor}$) from *Fusarium oxysporum*[17,18] and CYP119 from the Archaeon *Sulfolobus solfataricus*.[19] More recently, the first mammalian solubilized P450 structure (rabbit CYP2C5/3[20]) has been resolved.

In the literature, several homology models for eukaryotic P450s based on the crystal structure of CYP101, CYP102, or a combination of CYP101, CYP102, and CYP108 have been reported. Until the release of the CYP2C5/3 X-ray structure, CYP102 was considered to provide the most useful structural information for homology studies on eukaryotic P450s, since this well-characterized bacterial enzyme belongs to the class II P450s[11] to which many eukaryotic P450s belong.

A variety of three-dimensional representations of the active site of CYP2C9 have been constructed using various modeling techniques and reported in the literature. Pharmacophore models have been developed predicting either a hydrogen-bond donor atom[21] or an "anionic site" in the substrate.[22,23] A CoMFA (comparative molecular field analysis) model indicated that the active site contains a lipophilic binding site and at least one positively charged site (and possibly a second one).[24,25] Recently this model was refined and the interacting amino acids were proposed to be Phe^{114}, Arg^{105}, and possibly Asp^{297} (via a water molecule).[26] Ligand-based models derived using Catalyst[27] and PLS (partial least squares)—MS-WHIM (molecular surface-weighted holistic molecular)[28] 3D- and 4D-QSARs (quantitative structure−activity relationships) showed at least one hydrophobic site and one hydrogen-bond acceptor in the active site of CYP2C9.[29] Homology models have been constructed on the basis of the crystal structures of either CYP101[30] or CYP102.[31−34]

Recently, we have derived a combined model for CYP2D6[35] based on the combination of a homology model constructed using the crystal structures of CYP101, CYP102, and CYP108, a pharmacophore model, and MO (molecular orbital) calculations on the substrates, metabolic intermediates, and products.

In this work, we have constructed a three-dimensional model for CYP2C9 and its substrates, based on a

* To whom correspondence should be addressed. Phone: +44-1304-648746. Fax: +44-1304-658463. E-mail: degroot@sandwich.pfizer.com.
† Department of Molecular Informatics, Structure & Design.
‡ Department of Pharmacokinetics, Dynamics & Metabolism.

10.1021/jm0110791 CCC: $22.00
Published on Web 04/06/2002

combination of (a) a homology model constructed using the crystal structures of CYP2C5 and CYP102, (b) a pharmacophore model, and (c) MO calculations on the substrates, metabolic intermediates, and products. This model will enable us to rationalize and predict sites of CYP2C9-mediated metabolism as described for CYP2D6 previously.[35,36]

As in the aforementioned CYP2D6 model, this combined methodology offers a better understanding of the CYP2C9 active site and its interactions with substrates than each of these approaches in isolation, since it takes into account the most reactive sites in the substrate as well as the conformational and stereochemical constraints imposed by the protein active site. The independent generation of the pharmacophore and protein models offers the opportunity to compare and contrast as well as "cross-validate" both approaches. Furthermore, when using the model for predictions of the metabolism catalyzed by CYP2C9 of compounds for which information on the actual site of oxidation is not available, the MO calculations indicate the electronically most likely sites of oxidation, which are used to guide inclusion of these compounds in the combined model.

Computational Methods

Construction of the Initial Pharmacophore Model for Substrates of CYP2C9. The initial model is based on 16 structurally diverse substrates (17 metabolic pathways) known to be metabolized by human CYP2C9 and on their sites of oxidation (see Figure 1 and Table 1). After construction of the substrates, a conformational search and energy minimization was performed[37] using Spartan.[38] The rigid substrate 58C80 (**1**) was selected as the template. Each substrate from the initial set of substrates (Figure 1a) was then overlaid onto **1**, matching the site of oxidation and potential hydrogen-bond donating/accepting groups or hydrophobic/aromatic moieties of the substrates using the "MULTIFIT" method in SYBYL.[39] Antipyrine (**3**) was added as a separate template for smaller molecules.

Construction of a Homology Model for CYP2C9. The recently published crystal structure of CYP2C5/3[20] was used as the template for the homology model. However, the B′-region in this crystal structure displays some very large temperature factors, and some amino acids possess unrealistic side chain orientations. Therefore, the B′-helical region of the crystal structure of CYP102 (2hpd.pdb, chain A[11]) was used as template for the CYP2C9 B′-region.[40] The consensus homology modeling program Modeler[41,42] (as implemented in Quanta97[43]) has been applied using the alignment shown in Figure 2.

A set of 20 models was produced, and the best model was selected on the basis of a Procheck[44] analysis of stereochemical quality and a visual inspection in Quanta97[43] to assess (a) the location of the iron relative to the heme moiety, (b) the orientation of the amino acids involved in hydrogen bonds with the heme propionate moieties, and (c) the location of amino acids that have been identified by site-directed mutagenesis.[45−51]

Molecular Orbital (MO) Calculations on CYP2C9 Substrates. As described previously[35] for each of the substrates (Figure 1), a conformational search followed by an energy minimization was performed[37] using the AM1 method[52] in Spartan.[38] Starting from the lowest energy conformation of the molecule, all likely radicals, hydroxylated products, and the radical cation were generated and their geometries optimized.[37] For each substrate, the chemically most likely products were compared with the experimentally observed metabolic products. The MO calculations are more important when predictions based on the model are made (Figure 1c) because in the absence of experimental information, the MO results for the various possible products are used to identify the most likely sites of oxidation.

After combination of the pharmacophore and homology models and minimization of the substrates in the presence of the protein (see below), a similar AM1 geometry optimization was performed on the geometry of the substrate resulting from this minimization in the protein (after which the fit with the active site was reassessed).

Combination of the Initial Pharmacophore and Homology Models for CYP2C9, Extension of the Pharmacophore Model, and Predictions Using the Combined Model. The derived CYP2C9 pharmacophore model (using substrates **1**−**16**, Figure 1a) was oriented in the active site of the CYP2C9 protein model[53] using SYBYL[39] without changing the orientations of the substrates relative to each other. The substrates were oriented with the site of oxidation approximately 3.0−3.5 Å above the iron atom of the heme, with the plane of the heme roughly perpendicular to the semiplanar region of the pharmacophore model, the acidic group/hydrogen-bond acceptor group oriented toward the side chain of Arg^{108}, and the aromatic parts of the molecules oriented such that an interaction with Phe^{476} was possible while avoiding steric clashes with the other amino acids lining the interior of the active site. The pharmacophore model fitted well in the protein model even though the two models were derived independently.

In the next step each individual substrate was energy-minimized within the active site using the SYBYL[39] force field. The entire protein and the substrate were allowed to minimize with one distance constraint and an additional oxygen atom bound to the iron (Figure 3)[54] to ensure that known heme−substrate interactions were included. After the optimization within the protein, the substrates were again optimized using AM1 to ensure that the docked conformations were in a local energy minimum, and then they were reinserted into the model. None of the docked conformations were more than 5 kcal/mol above the calculated global energy minimum (see Table 1).

A second set of 11 substrates (14 metabolic pathways, increasing the structural diversity of the set of substrates further, Figure 1b) was added to the pharmacophore after combination of the initial pharmacophore with the protein model, taking into account the influence of Arg^{108} and Phe^{476} (see below). To facilitate the addition of new molecules, a MOLCAD surface[55] was generated for the active-site region (Figure 4). This same method was used for the test set of compounds (Figure 1c).

Results and Discussion

General. Unlike CYP2D6, the interactions between protein and substrate seem to be less strictly defined. Whereas CYP2D6 relies on a very specific interaction between a basic nitrogen atom present in the substrate and Asp^{301} and on an aromatic interaction in all compounds that are hydroxylated or O-demethylated,[35] CYP2C9 metabolizes a more diverse range of compounds (e.g., aromatic compounds such as benzo[*a*]pyrene (**4**), acidic compounds such as diclofenac (**9**), basic compounds such as fluoxetine, and polar compounds such as phenytoin (**21**)). As a consequence, not all potential interaction sites within the protein are used by all substrates. The compounds generally possess a hydrophobic region (often an aromatic ring) 5.6−9.8 Å from the site of oxidation, while a variety of substrates contain a hydrogen-bond acceptor group 5−10 Å from the site of oxidation.

Homology Model. One of the key prerequisites for homology modeling is the accuracy of the sequence alignment. Although the sequence similarity between the P450 families is low,[56] the structural similarity of all significant secondary structural elements is conserved across the available bacterial structures. It is

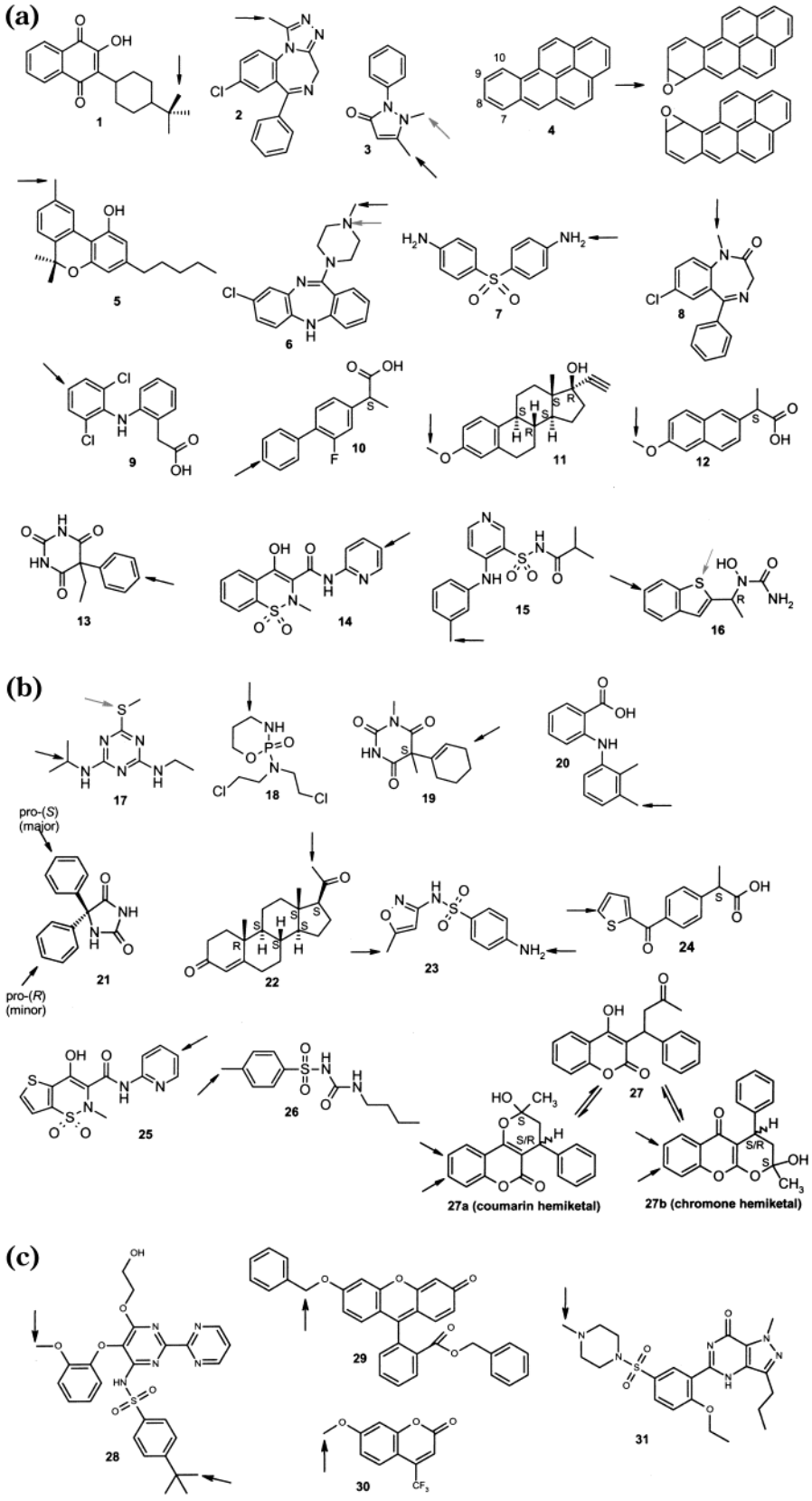

Figure 1. Structures and sites of metabolism. (a) Substrates used for initial pharmacophore are the following: 58C80[71] (**1**), alprazolam[72] (**2**), antipyrine[73] (**3**), benzo[*a*]pyrene[74] (**4**), cannabinol[75] (**5**), clozapine[76] (**6**), dapsone[77] (**7**), diazepam[78] (**8**), diclofenac[3] (**9**), (*S*)-flurbiprofen[79] (**10**), mestranol[80] (**11**), (*S*)-naproxen[81] (**12**), phenobarbital[82] (**13**), piroxicam[3] (**14**), torsemide[83] (**15**), and zileuton[84] (**16**). (b) Substrates used to extend the pharmacophore after combination with the protein model are the following: ametryne[66,85] (**17**), cyclophosphamide[86] (**18**), hexobarbital[87] (**19**), mefenemic acid[3] (**20**), phenytoin (leading to (*S*)-hydroxy product and (*R*)-hydroxy product)[3,34,88,97] (**21**), progesterone[89] (**22**), sulfamethoxazole[77] (**23**), suprofen[3] (**24**), tenoxicam[3] (**25**), tolbutamide[3,90,91] (**26**), and warfarin (in closed coumarin hemiketal form (**27a**) and closed chromone hemiketal form (**27b**)[25]). (c) Test substrates are the following: bosentan[92,93] (**28**), DBF[94] (dibenzylfluorescein, **29**), MFC[94] (7-methoxy-4(trifluoromethyl)coumarin, **30**), and sildenafil[95,96] (**31**). Gray arrows indicate known sites of metabolism (N-oxidation and S-oxidation) not used in the model.

Table 1. Substrates and Metabolic Pathways Used To Construct and Test the Model

	pathway	ΔE (AM1) (kcal/mol)	distance oxidation site to dummy[a] (Å)	ref
1	t-butyl hydroxylation	0.0	1.86	71
2	methyl hydroxylation	0.0	1.69	72
3	methyl hydroxylation	0.0	1.63	73
4	7,8-epoxide formation	+0.07	1.85	74
	9,10-epoxide formation	0.0	1.96	
5	benzylic hydroxylation	+0.09[b]	1.67	75
6	N-demethylation	+3.72	1.75	76
7	N-oxidation	0.0	1.91	77
8	N-demethylation	0.0	1.75	78
9	aromatic hydroxylation	0.0	2.00	3
(*S*)-**10**	aromatic hydroxylation	+0.18	1.82	79
11	O-demethylation	+0.01	1.73	80
(*S*)-**12**	O-demethylation	+0.28	1.75	81
13	aromatic hydroxylation	0.0	1.89	82
14	aromatic hydroxylation	0.0	1.82	3
15	benzylic hydroxylation	0.0[c]	1.83	83
(*R*)-**16**	aromatic hydroxylation	0.0	1.81	84
17	*i*-propyl hydroxylation	+0.01	2.25	85
18	ring opening	0.0	1.76	86
(*S*)-**19**	ring hydroxylation	+0.01	1.97	87
20	benzylic hydroxylation	0.0	1.73	3
21[d]	aromatic hydroxylation (pro-*S*)	+1.85	0.0	3, 34, 88
	aromatic hydroxylation (pro-*R*)	+1.83	0.0	
22	methyl hydroxylation	0.0	1.79	89
23	N-oxidation	0.0	1.76	77
	methyl hydroxylation	0.0	1.81	
(*S*)-**24**	ring hydroxylation	0.0	2.00	3
25	aromatic hydroxylation	0.0	1.81	3
26	benzylic hydroxylation	+0.01	1.81	3, 90, 91
(*S*,*S*)-**27a**[e]	aromatic hydroxylation (6-OH)	0.0	1.96	3, 25
	aromatic hydroxylation (7-OH)	0.0	1.93	
(*S*,*R*)-**27a**[e]	aromatic hydroxylation (6-OH)	+1.19	1.93	3, 25
	aromatic hydroxylation (7-OH)	+1.19	1.70	
(*S*,*S*)-**27b**[e]	aromatic hydroxylation (6-OH)	+0.02	1.78	3, 25
	aromatic hydroxylation (7-OH)	+0.01	1.89	
(*S*,*R*)−**27b**[e]	aromatic hydroxylation (6-OH)	+1.05	1.86	3, 25
	aromatic hydroxylation (7-OH)	0.0	1.95	
28	*t*-butyl hydroxylation	+1.44	1.83	92, 93
	O-demethylation	+0.32	2.09	
29	O-debenzylation	+0.56	1.84	94
30	O-demethylation	0.0	1.73	94
31	N-demethylation	+0.01	1.80	95, 96

[a] After optimization in the presence of protein. Distance was constrained at 1.75 Å during calculation.[54] [b] 5.97 kcal/mol before modifying clashing side chain orientation. [c] 4.97 kcal/mol before modifying clashing side chain orientation. [d] Metabolism of (*S*)-product dominates over (*R*)-product 43:1,[34,97] but both major and minor pathways fit within model, as can be expected from a *binding* model. [e] The open form of (*S*)-warfarin can form four closed-form enantiomers. Each of these forms has been taken into consideration. Some forms can directly interact with Phe[114] (e.g., conformations leading to (*S*,*R*)-**27b** 6-hydroxylation and (*S*,*S*)-**27a** 7-hydroxylation), while other enantiomers are interacting with Arg[108], which is influenced by the orientation of the Phe[114] side chain.

believed that this structural homology is also conserved in the membrane-bound mammalian P450s.[57,58] The rabbit CYP2C5/3 crystal structure[20] showed that this belief is realistic. However, a few regions, most notably the region between the B and C helices, are diverse in terms of both structure and sequence across the bacterial P450 crystal structures and are not well defined in the rabbit CYP2C5/3 crystal structure,[20] which indicates that modeling of such regions is challenging. Therefore, the B′-loop region of CYP102[11] was used as a template for the CYP2C9 model instead of the rabbit CYP2C5/3 structure[20] (see Figure 2). Quality checks were performed on the resulting model using criteria as described in Computational Methods.

Creation of Initial Pharmacophore and Protein Models and Combination of These Models. A set of 16 substrates (Figure 1a) was used to construct a pharmacophore model using the site of oxidation, a hydrogen-bond acceptor site, and an aromatic region to superimpose the molecules. 58C80 (**1**) and antipyrine (**3**) were used as template molecules. Insertion of the initial pharmacophore model (Figure 1a) into the protein model did not result in steric clashes between substrates and protein. Because the two models were derived independently, this was interpreted as a validation of both the pharmacophore and the protein model.

The combination of the pharmacophore model with the homology model indicated two amino acids important for binding the substrates. When present, the acidic or hydrogen-bond acceptor group forms an interaction with Arg[108], while the aromatic moieties within the substrates form an aromatic stacking interaction with Phe[476]. Both Arg[108] and Phe[476] have some conformational freedom, which allows interactions with hydrogen-bond acceptor groups and hydrophobic regions at various distances from the site of oxidation within the substrates. Not all substrates interact with both of these amino acids. This is in line with the previously proposed multiple binding sites within certain P450s (including CYP2C9) resulting in simultaneous binding of two (different) substrates within the CYP2C9 active site.[25,59,60]

```
CYP2C9      SSGRGKLPPGPTPLPVIGNILQIGIKDISKSLTNLSKVYGPVFTLYFGLK
CYP2C5/3    -------PPGPTPFPIIGNILQIDAKDISKSLTKFSECYGPVFTVYLGMK
                    |       |       |       |       |       |
                   23      30      40      50      60      70

CYP2C9      PIVVLHGYEAVKEALIDLGEEFSGRGI-FPLAERANRGFGIVFSNGKKWK
CYP102      --------------------FDKNLSQALKFVRDFAGDGLFTS------
CYP2C5/3    PTVVLHGYEAVKEALVDLGEEF--------------------FSNAKTWK
                    |       |       |       |       |       |
                   73      80      90      100     110     120

CYP2C9      EIRRFSLMTLRNFGMGKRSIEDRVQEEARCLVEELRKTKASPCDPTFILG
CYP2C5/3    EMRRFSLMTLRNFGMGKRSIEDRIQEEARCLVEELRKTNASPCDPTFILG
                    |       |       |       |       |       |
                  122     130     140     150     160     170

CYP2C9      CAPCNVICSIIFHKRFDYKDQQFLNLMEKLNENIKILSSPWIQICNNFSP
CYP2C5/3    CAPCNVICSVIFHNRFDYKDEEFLKLMESLHENVELLGTPWLQVYNNFPA
                    |       |       |       |       |       |
                  172     180     190     200     210     220

CYP2C9      IIDYFPGTHNKLLKNVAFMKSYILEKVKEHQESMDMNNPQDFIDCFLMKM
CYP2C5/3    LLDYFPGIHKTLLKNADYIKNFIMEKVKEHQKLLDVNNPRDFIDCFLIKM
                    |       |       |       |       |       |
                   222     230     240     250     260     270

CYP2C9      EKEKHNQPSEFTIESLENTAVDLFGAGTETTSTTLRYALLLLLKHPEVTA
CYP2C5/3    EQENNL---EFTLESLVIAVSDLFGAGTETTSTTLRYSLLLLLKHPEVAA
                    |       |       |       |       |       |
                  272     280     290     300     310     320

CYP2C9      KVQEEIERVIGRNRSPCMQDRSHMPYTDAVVHEVQRYIDLLPTSLPHAVT
CYP2C5/3    RVQEEIERVIGRHRSPCMQDRSRMPYTDAVIHEIQRFIDLLPTNLPHAVT
                    |       |       |       |       |       |
                  322     330     340     350     360     370

CYP2C9      CDIKFRNYLIPKGTTILISLTSVLHDNKEFPNPEMFDPHHFLDEGGNFKK
CYP2C5/3    RDVRFRNYFIPKGTDIITSLTSVLHDEKAFPNPKVFDPGHFLDESGNFKK
                    |       |       |       |       |       |
                  372     380     390     400     410     420

CYP2C9      SKYFMPFSAGKRICVGEALAGMELFLFLTSILQNFNLKSLVDPKNLDTTP
CYP2C5/3    SDYFMPFSAGKRMCVGEGLARMELFLFLTSILQNFKLQSLVEPKDLDITA
                    |       |       |       |       |       |
                  422     430     440     450     460     470

CYP2C9      VVNGFASVPPFYQLCFIPV--
CYP2C5/3    VVNGFVSVPPSYQLCFIPIHH
                    |       |       |
                  472     480     490
```

Figure 2. Sequence alignment between CYP2C5, CYP102 (B′-region only), and CYP2C9 used to construct the protein model.[40] Identical regions are shown with gray background. Amino acids that are different in the active site between CYP2C9 and CYP2C19 are shown in bold.

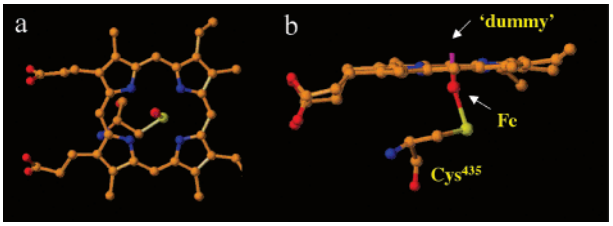

Figure 3. Active-site geometry used for the calculations, showing the dummy oxygen atom used.

Minimization of each individual substrate contributing to this initial pharmacophore within the protein[54] induced only minor changes to the overall pharmacophore model (see distances in Table 1). When these substrates were subsequently submitted to an AM1 geometry optimization (starting from the orientation resulting from the geometry optimization in the presence of the protein), there were again only very minor changes in substrate geometry. This indicates that the conformations in the protein are low-energy conformations and not a result of protein-imposed steric restrictions. It was still possible to incorporate all substrates within the active-site region of the homology model, and the energy differences between these AM1-optimized geometries and the AM1 energies for the global minimum were generally less than 5 kcal/mol (Table 1). An example of a docked substrate (**1**) within the protein is shown in Figure 5.

The pharmacophore was extended with a second set of known CYP2C9 substrates (Figure 1b), taking into account the environment of the protein as described above. All substrates could be incorporated into the pharmacophore in low-energy conformations, which after optimization within the protein and subsequent AM1 minimization (without protein present) were no more than 5 kcal/mol higher in energy than the calcu-

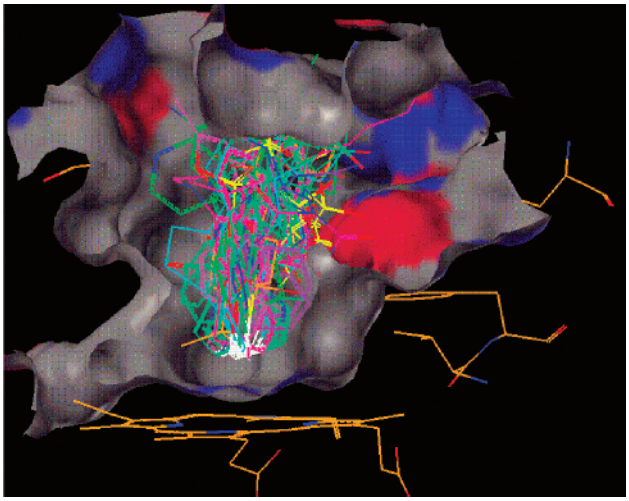

Figure 4. MOLCAD surface[55] generated within the active site of the protein model, showing the pharmacophore model contained within the protein model, localized over the heme moiety. Surface is color-coded by hydrogen-bonding potential (blue, hydrogen-bond acceptor in protein; red, hydrogen-bond donor in protein).

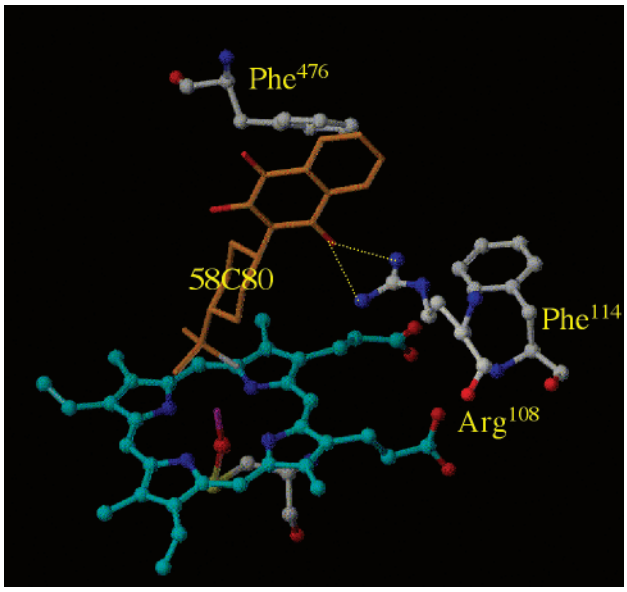

Figure 5. 58C80 (**1**) within the active site of the CYP2C9 protein model. Only a few amino acids are shown. A hydrogen bond is formed between Arg[108] and the carbonyl moiety of **1**, while an aromatic interaction between Phe[476] and the aromatic ring of **1** stabilizes this binding conformation. Phe[114] is shown for reference.

lated global minimum structures, and movement of the substrates within the active site was small (Table 1). The same amino acids were involved in binding the substrates. When present, the acidic or hydrogen-bond acceptor groups form interactions with Arg[108], while the aromatic moieties of the substrates form an aromatic stacking interaction with Phe[476]. Although the original set of substrates (Figure 1a) did not contain large compounds that reached the Val[113]−Phe[114] region, the second set of substrates (Figure 1b) contained warfarin either in the ring-closed coumarin hemiketal form (**27a**) or in the ring-closed chromone hemiketal form (**27b**).[25] Both these forms can interact directly with Val[113] and/or Phe[114] (amino acids shown by SDM to be involved in binding of warfarin[49]). For example, Phe[114] can stabilize binding orientations resulting in (*S*,*R*)-**27b** 6-hydroxylation and (*S*,*S*)-**27a** 7-hydroxylation. However, Phe[114] is partially shielded from interactions with other substrates within the CYP2C9 active site by the side chain of Arg[108]. Therefore, our model suggests that the role of Phe[114] might not, as previously[49] suggested, be to directly interact with substrates, but it could influence the orientation of side chains of other amino acids (in this case Arg[108]) responsible for interactions with substrates. This concept was recently suggested by Tsao et al. to explain (*S*)-mephenytoin 4′-hydroxylation.[51]

Predictions. In the case of predictions based on the model (Figure 1c), the MO results for the various possible products were used to guide the modeling, thereby increasing the importance of these MO calculations. Although the sites of metabolism for these compounds were available in the literature, this information was not taken into account at the start of the experiment; therefore, all potential orientations leading to metabolism at various sites within the molecule were docked. In most cases, only one or two sites of metabolism could be accommodated within the active site in a low-energy conformation (see Table 1). All experimentally observed metabolites were predicted correctly by the model. MO calculations alone were not expected to reliably predict the metabolic site of oxidation, since oxidation next to either a nitrogen or oxygen atom is generally favored in the absence of a catalyst or protein. As was the case for CYP2D6,[35,36] the site of metabolism in CYP2C9 is mainly dependent on steric factors instead of the electronic factors calculated using MO techniques.

Mutants and Site-Directed Mutagenesis in the Literature. The allelic variant that has Cys[144] instead of Arg[144] results in impaired metabolism of (*S*)-warfarin by CYP2C9.[45] This amino acid is on the surface of the protein, and it has been suggested that it disrupts a reductase contact point.[32] The allelic variant Leu[359] (instead of Ile[359]) responsible for a lower V_{max} for the metabolism of (*S*)-warfarin and tolbutamide[61] is also not in direct contact with the active-site region. The influence of various amino acids on the metabolism of CYP2C9 substrates has previously been examined using SDM. Initially, efforts were directed at mutating specific amino acids in CYP2C19 in order to obtain CYP2C9 metabolic efficiency for various substrates or the other way around. CYP2C9 mutations of Ile[99], Pro[220], and Thr[221] to the equivalent residues in CYP2C19 (His[99], Ser[220], and Pro[221]) conferred omeprazole activity onto CYP2C9.[46] The amino acid mutations that confer high affinity to CYP2C19 for sulfaphenazole binding and the ability to metabolize warfarin were found to be Asn[286] to Ser[286] and Ile[289] to Asn[289] (the CYP2C9 equivalents).[47] The same mutations conferred diclofenac and ibuprofen activity to CYP2C19.[48] These residues are not part of the active site but are part of the I-helix interacting with the B′-region and are probably involved in the stabilization of these regions relative to each other in a way similar to the hydrophobic interactions that stabilize the CYP101 fold.[62]

Mutagenesis experiments guided by homology modeling and CoMFA analysis indicate that Phe[114] but not Phe[110] is consistent with the observed metabolic variation and is together with Val[113] involved in hydrophobic interactions with substrates.[49] However, as described

above, this might not be caused by a direct interaction with warfarin, but by influencing the orientation of Arg[108], although a direct interaction can be formed with certain substrates.

Mutagenesis in the region of Arg[105] indicates that this region is not an important anionic binding site.[50] Mutation of Arg[97] has an influence on diclofenac metabolism, but this amino acid is not predicted to be in the active site by either homology modeling using CYP102 or the CYP2C5/3 crystal structure itself.[20,50] Mutation of Arg[108] to alanine reduces the formation of 4′-hydroxydiclofenac 100-fold compared with CYP2C9 wild type and is another potential anionic binding site, although in a recent homology model based on the crystal structure of bacterial CYP102, Arg[108] is not in the active site.[50] However, in the currently described model for CYP2C9, Arg[108] is an essential part of the CYP2C9 active site. Recent SDM experiments designed to confer (*S*)-mephenytoin activity onto CYP2C9 suggested that this activity is caused by amino acids that influence the packing of structural elements or influence the substrate access channel rather than altering the active site directly (His[99] in the B′-region, Pro[220] and Thr[221] in the F−G-loop, and Asn[286], Ala[292], and Leu[295] in the I-helix, none of which are in direct contact with (*S*)-mephenytoin).[51] A similar role has now been proposed for Phe[114] (see above).

Literature Models for CYP2C9. A variety of models for CYP2C9 have been described in the literature, including protein models, pharmacophore models, and CoMFA/QSAR based models.

Protein Models. The first protein models for CYP2C9[30,31] were based on the crystal structures of either CYP101 or CYP102. A similar homology model (based on CYP102) incorporated all known SDM results and antibody recognition information but was quite different from the previous and later CYP2C9 models and suggested Lys[77] as an amino acid involved in substrate recognition (located in the substrate access channel) and Ile[99], Phe[100], and Ser[365] as being involved in substrate binding.[32] A CYP2C9 homology model[34] based on CYP101, CYP102, CYP107A, and CYP108 suggested that steric restrictions were more important than electrostatic interactions and highlighted Arg[105] and Arg[97] as key cationic residues with minor binding contributions from Pro[101], Leu[102], Asn[202], Gly[296], Asp[393], and Phe[476].

Another homology model of CYP2C9, based on CYP101, CYP102, and CYP107A, was constructed for use with a CoMFA model (see below) and suggested that Phe[110] and Phe[114] might be involved in aromatic binding interactions and that Arg[105]and Asp[293] might be responsible for electrostatic interactions.[26]

A homology model based on the bacterial CYP102 crystal structure (used to explain the SDM results for Phe[97], Phe[105], and Phe[108]; see above)[50] was recently followed by the first homology model for CYP2C9 based on the crystal structure of CYP2C5/3.[63] The latter model suggested a role for Val[113], Phe[114], Leu[201], Asp[204], Ile[205], Asp[293], Ala[297], Leu[362], Leu[366], and Phe[476] in interacting with various inhibitors.[63] A 3D-QSAR model was inserted into the latter CYP2C5/3 based homology model (see below). Another set of homology models for CYP2C9 and CYP2C19, based on CYP2C5/3, was constructed to examine and explain the amino acid residues responsible for conferring (*S*)-mephenytoin 4′-hydroxylation activity onto CYP2C9.[51] This suggested that the selective metabolism of (*S*)-mephenytoin over (*R*)-mephenytoin was caused by amino acids that influence the packing of structural elements or influence the substrate access channel instead of altering the active site directly.[51]

Because of the high-sequence homology between CYP2C9 and CYP2C5/3, models based on CYP2C5/3 will be more reliable than models based on bacterial crystal structures. The three protein models based on the CYP2C5/3 crystal structure (the currently described model and the models by Afzelius et al.[63] and Tsao et al.[51]) differ in the alignment used to construct the protein models and the software used to derive the models. While the protein model described here focuses on a combination of a homology model with an imbedded pharmacophore model to predict the potential sites of metabolism in CYP2C9 substrates, the model by Afzelius et al.[63] combines a homology model with 3D-QSAR to predict the K_i's for CYP2C9 inhibitors.

Pharmacophore Models. The first pharmacophore model of CYP2C9 suggested a hydrogen-bond donor or acceptor within the active site.[64] This model was updated by overlaying eight substrates and one inhibitor, using phenytoin as the template molecule and ensuring that sites of metabolism and the hydrogen-bond donor heteroatoms were superimposed.[21] A second pharmacophore model was based on 10 compounds, using phenytoin and warfarin as templates, and was supported by data for 10 additional tienilic acid analogues.[22] This model superimposed the substrates using their sites of oxidation and brought all anionic heteroatoms in the substrates within 3.5−4.8 Å of a (hypothetical) cationic interaction site within the CYP2C9 protein.[22] The orientations of the anionic heteroatoms differed between substrates, and therefore, a hydrogen bond to the protein was suggested to be unlikely and a purely cation−anion interaction was presumed instead. This model was updated, leading to the inclusion of a hydrophobic zone between the hydroxylation site and the cationic site on the protein.[23] Evidence for the importance of an anionic site in the substrates was recently strengthened by site-directed mutagenesis experiments highlighting the importance of arginine residues within the CYP2C9 active site.[50]

The currently derived pharmacophore is suggested to interact with Arg[108] for several compounds. This is in line with the work by Ridderström et al.,[50] which suggested that Arg[108] is one of the important amino acids in the CYP2C9 active site. Comparison of the current pharmacophore with the pharmacophore derived by Mancy et al.[22,23] suggests that Arg[108] is the cationic site in the latter model.

CoMFA/QSAR Based Models. One QSAR model for CYP2C9 overlaid 27 compounds by aligning phenyl rings, benzylic carbon atoms, and attached hydrogen atoms within the molecules but not the sites of oxidation. 9(*S*),11(*R*)-Cyclocoumarol was used as a template, while the model assumed the presence of an aromatic binding region and a sterically forbidden region.[24] This CoMFA model predicted two cationic enzyme binding sites and an aromatic region as important features,

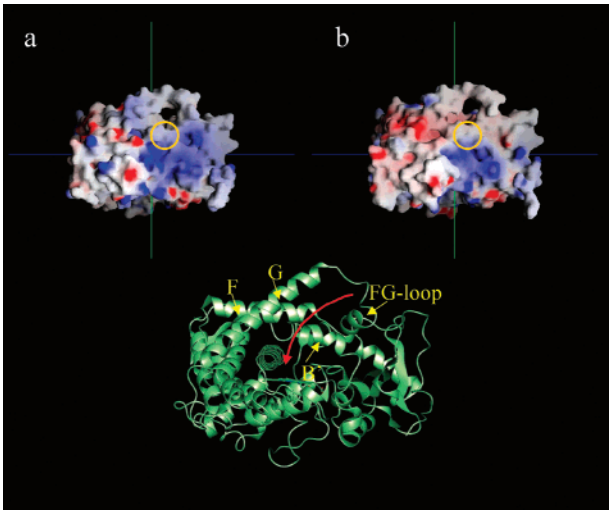

Figure 6. Charge distribution around the proposed substrate access channel in (a) CYP2C9 and (b) CYP2C19. The figure was generated using GRASP,[98] with the surface potential ranging from −15 (red) to +15 (blue) kcal/mol. The approximate location of the entrance of the access channel as described in bacterial crystal structures is indicated by a circle. The orientation shown here is similar to that used by Lüdemann et al. describing various pathways in CYP101 and CYP102.[62,68,69]

together with a steric region near the site of metabolism (possibly the I-helix). This model was refined with 14 additional compounds with starting geometries derived from calculations within a homology model of CYP2C9,[26] suggesting that Phe[110] and Phe[114] might be involved in aromatic interactions and that Arg[105]and Asp[293] could be responsible for electrostatic interactions.[26] The initial model predicted the K_i's of 13 of the 14 compounds correctly, and inclusion of these extra compounds did not change this model appreciably.

3D- and 4D-QSAR models for CYP2C9[29] have been constructed using Catalyst[27] and PLS MS-WHIM[28] using three different data sets. All models had at least one hydrophobic region and one hydrogen-bond acceptor, similar to the results obtained for the CoMFA models described above.[24,26]

The CYP102-based homology model[50] was recently updated to a new model[63] based on the crystal structure of CYP2C5/3[20] and extended with a 3D-QSAR model for CYP2C9 inhibitors including conformers selected using this updated CYP2C9 model. The resulting model was capable of predicting the K_i's of eight compounds not used to derive the model to within 0.5 log units.[63]

Comparison with CYP2C19. The differences between CYP2C9 and CYP2C19 are relatively small (95.7% homology[65] with only 43 out of 490 amino acids differing, of which 18 are conservative changes; see Table 2 in Lewis et al.[32]). In the active-site region, only three conservative amino acid differences are present between these two enzymes (Leu[208], Val[292], and Leu[362] in CYP2C9 (highlighted in Figure 2) are replaced by Val[208], Ala[292], and Ile[362] in CYP2C19[66]). This makes the differences in substrate specificity difficult to explain between these two 2C-family members (e.g., CYP2C19 does not metabolize acidic substrates[51]). When the entrances of the substrate access channels of CYP2C9 and CYP2C19 are compared (Figure 6), it is obvious that the proposed entrance of the CYP2C19 access channel between the F−G loop and the B′-region is more negatively charged and would therefore repel negatively charged substrates. This entrance channel has been described for CYP101,[6] CYP102,[11,67] and CYP2C5/3[20] and was described as pathway 2 by Lüdemann et al. as found by thermal motion pathway analysis,[68] various molecular dynamics (MD) techniques,[62,68,69] and adiabatic mapping of MD trajectories.[69] Although acidic substrates can be accommodated within the active site of CYP2C19, these substrates will not easily reach the active site because of the repulsion at the entrance of the substrate access channel. This argument not only is supported by SDM data (see above) but also is supported by the recent crystal structure of CYP51 for *Mycobacterium tuberculosis*,[70] in which the amino acids responsible for azole resistance are located in the substrate access channel and not in the active site.

Summary and Conclusions

This paper describes the first combined model comprising a pharmacophore for a diverse set of substrates of CYP2C9 within a homology model based on a mammalian crystal structure.[20] A combined model has the advantages of both types of models and results in a well-defined model that will take into account not only the chemical reactivities and the similarities between the various substrates but also the steric and electronic properties of the surrounding enzyme.

Our model identifies Arg[108], Phe[114], and Phe[476] as key residues responsible for substrate binding and substrate selectivity. Arg[108] and Phe[476] are suggested to be involved in binding interactions with various substrates through polar and hydrophobic interactions, respectively. The suggestion that Arg[108] is a key residue in binding polar and acidic substrates is in line with the SDM results from Ridderström et al.[50] Phe[476] was also suggested as a residue responsible for hydrophobic interactions in earlier homology models.[34,63] Phe[114], a residue that influences warfarin binding,[49] is suggested to influence the orientation of Arg[108], thereby influencing the binding of warfarin (a concept recently used to explain (*S*)-mephenytoin 4′-hydroxylation[51]), although a direct interaction can be formed with certain substrates.

Molecular orbital calculations can identify the chemically most likely intermediates and products based on internal energies. However, the steric and electronic influences of the protein, which distinguish the catalyzed reaction from the chemical reaction, are ignored. For the prediction of the possible metabolism of newly designed compounds, these calculations are very useful in indicating the most reactive parts of the compound and indicate possible sites of metabolism. Our model successfully predicted the sites of metabolism in CYP2C9 substrates not used to derive the combined model.

Comparison of charge distributions on the surfaces and the active sites of CYP2C9 and CYP2C19 homology models suggest that although acidic substrates can be accommodated within the active site of CYP2C19, these substrates will not easily reach the active site because of the repulsion at the entrance of the substrate access channel, when compared with CYP2C9.

Acknowledgment. The authors thank Valerie A. Horne for useful discussions and feedback.

References

(1) de Groot, M. J.; Vermeulen, N. P. E. Modeling the active site of cytochrome P450s and glutathione *S*-transferases, two of the most important biotransformation enzymes. *Drug Metabol. Rev.* **1997**, *29*, 747−799.

(2) Ekins, S.; de Groot, M. J.; Jones, J. P. Pharmacophore and three-dimensional quantitative structure activity relationship methods for modeling cytochrome P450 active sites. *Drug Metab. Dispos.* **2001**, *29*, 936−944.

(3) Miners, J. O.; Birkett, D. J. Cytochrome P4502C9: an enzyme of major importance in human drug metabolism. *Br. J. Clin. Pharmacol.* **1998**, *45*, 525−538.

(4) Poulos, T. L. Cytochrome P450. *Curr. Opin. Struct. Biol.* **1995**, *5*, 767−774.

(5) Poulos, T. L.; Finzel, B. C.; Gunsalus, I. C.; Wagner, G. C.; Kraut, J. The 2.6-Å crystal structure of *Pseudomonas putida* cytochrome P-450. *J. Biol. Chem.* **1985**, *260*, 16122−16130.

(6) Poulos, T. L.; Finzel, B. C.; Howard, A. J. Crystal structure of substrate-free *Pseudomonas putida* cytochrome P-450. *Biochemistry* **1986**, *25*, 5314−5322.

(7) Poulos, T. L.; Howard, A. J. Crystal structures of metyrapone- and phenylimidazole-inhibited complexes of cytochrome $P450_{cam}$. *Biochemistry* **1987**, *26*, 8165−8174.

(8) Poulos, T. L.; Finzel, B. C.; Howard, A. J. High-resolution crystal structure of cytochrome P450cam. *J. Mol. Biol.* **1987**, *195*, 687−700.

(9) Raag, R.; Poulos, T. L. Crystal structures of cytochrome $P\text{-}450_{CAM}$ complexed with camphane, thiocamphor, and adamantane: factors controlling P-450 substrate hydroxylation. *Biochemistry* **1991**, *30*, 2674−2684.

(10) Li, H.; Narasimhulu, S.; Havran, L. M.; Winkler, J. D.; Poulos, T. L. Crystal structure of cytochrome $P450_{cam}$ complexed with its catalytic product, 5-*exo*-hydroxycamphor. *J. Am. Chem. Soc.* **1995**, *117*, 6297−6299.

(11) Ravichandran, K. G.; Boddupalli, S. S.; Hasemann, C. A.; Peterson, J. A.; Deisenhofer, J. Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450s. *Science* **1993**, *261*, 731−736.

(12) Li, H.; Poulos, T. L. Modeling protein−substrate interactions in the heme domain of cytochrome $P450_{BM-3}$. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1995**, *51*, 21−32.

(13) Ito, S.; Matsuoka, T.; Watanabe, I.; Kagasaki, T.; Serizawa, N.; Hata, T. Crystallization and preliminary X-ray diffraction analysis of cytochrome $P450_{sca-2}$ from *Streptomyces carbophilus* involved in production of pravastatin sodium, a tissue-selective inhibitor of HMG-CoA reductase. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *55*, 1209−1211.

(14) Cupp-Vickery, J. R.; Li, H. Y.; Poulos, T. L. Preliminary crystallographic analysis of an enzyme involved in erythromycin biosynthesis: Cytochrome P450eryF. *Proteins: Struct., Funct., Genet.* **1994**, *20*, 197−201.

(15) Cupp-Vickery, J. R.; Poulos, T. L. Structure of cytochrome P450eryF involved in erythromycin biosynthesis. *Nat. Struct. Biol.* **1995**, *2*, 144−153.

(16) Hasemann, C. A.; Ravichandran, K. G.; Peterson, J. A.; Deisenhofer, J. Crystal structure and refinement of cytochrome $P450_{terp}$ at 2.3 Å resolution. *J. Mol. Biol.* **1994**, *236*, 1169−1185.

(17) Nakahara, K.; Shoun, H.; Adachi, S.; Iizuka, T.; Shiro, Y. Crystallization and preliminary X-ray diffraction studies of nitric oxide reductase cytochrome P450nor from *Fusarium oxysporum*. *J. Mol. Biol.* **1994**, *239*, 158−159.

(18) Park, S.-Y.; Shimizu, H.; Adachi, S.; Shiro, Y.; Iizuka, T.; Shoun, H. Crystal structure of nitric oxide reductase cytochrome P-450nor from *Fusarium oxysporum*. *Keio Univ. Symp. Life Sci. Med.* **1998**, *1* (Oxygen Homeostasis and Its Dynamics), 147−155.

(19) Yano, J. K.; Koo, L. S.; Schuller, D. J.; Li, H.; Ortiz de Montellano, P. R.; Poulos, T. L. Crystal structure of a thermophylic cytochrome P450 from the Archaeon *Sulfolobus solfataricus*. *J. Biol. Chem.* **2000**, *275*, 31086−31092.

(20) Williams, P. A.; Cosme, J.; Sridhar, V.; Johnson, E. F.; McRee, D. E. Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity. *Mol. Cell* **2000**, *5*, 121−131.

(21) Jones, B. C.; Hawksworth, G.; Horne, V. A.; Newlands, A.; Morsman, J.; Tute, M. S.; Smith, D. A. Putative active site template model for cytochrome P4502C9 (Tolbutamide hydroxylase). *Drug Metab. Dispos.* **1996**, *24*, 260−266.

(22) Mancy, A.; Brotto, P.; Dijols, S.; Dansette, P. M.; Mansuy, D. The substrate binding site of human liver cytochrome P450 2C9: an approach using designed tienilic acid derivatives and molecular modeling. *Biochemistry* **1995**, *34*, 10365−10375.

(23) Mancy, A.; Dijols, S.; Poli, S.; Guengerich, F. P.; Mansuy, D. Interactions of sulfaphenazole derivatives with human liver cytochromes P450 2C: Molecular origin of the specific inhibitory effects of sulfaphenazole on CYP2C9 and consequences for the substrate binding site topology of CYP2C9. *Biochemistry* **1996**, *35*, 16205−16212.

(24) Jones, J. P.; He, M.; Trager, W. F.; Rettie, A. E. Three-dimensional quantitative structure−activity relationship for inhibitors of cytochrome P4502C9. *Drug Metab. Dispos.* **1996**, *24*, 1−6.

(25) He, M.; Korzekwa, K. R.; Jones, J. P.; Rettie, A. E.; Trager, W. F. Structural forms of phenprocoumon and warfarin that are metabolized at the active site of CYP2C9. *Arch. Biochem. Biophys.* **1999**, *372*, 16−28.

(26) Rao, S.; Aoyama, R.; Schrag, M.; Trager, W. F.; Rettie, A.; Jones, J. P. A refined 3-dimensional QSAR of cytochrome P450 2C9: computational predictions of drug interactions. *J. Med. Chem.* **2000**, *43*, 2789−2796.

(27) *Catalyst*, version 4.0; Molecular Simulations Inc.: San Diego, CA, 1998.

(28) Bravi, G.; Gancia, E.; Mascagni, P.; Pegna, M.; Todeschini, R.; Zaliani, A. MS-WHIM, new 3D theoretical descriptors derived from molecular surface properties. A comparative 3D-SAR study in a series of steroids. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 79−92.

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

(30) Korzekwa, K. R.; Jones, J. P. Predicting the cytochrome P450 mediated metabolism of xenobiotics. *Pharmacogenetics* **1993**, *3*, 1−18.

(31) Wiseman, H.; Lewis, D. F. V. The metabolism of tamoxifen by human cytochromes P450 is rationalized by molecular modelling of the enzyme−substrate interactions: potential importance to its proposed anti-carcinogen/carcinogen actions. *Carcinogenesis* **1996**, *17*, 1357−1360.

(32) Lewis, D. F. V.; Dickins, M.; Weaver, R. J.; Eddershaw, P. J.; Goldfarb, P. S.; Tarbit, M. H. Molecular modelling of human CYP2C subfamily enzymes CYP2C9 and CYP2C19: rationalization of enzyme specificity and site-directed mutagenesis experiments in the CYP2C subfamily. *Xenobiotica* **1998**, *28*, 235−268.

(33) Lewis, D. F. V. The CYP2 family: models, mutants and interactions. *Xenobiotica* **1998**, *28*, 617−661.

(34) Payne, V. A.; Chang, Y.-T.; Loew, G. H. Homology modeling and substrate binding study of human CYP2C9 enzyme. *Proteins: Struct., Funct., Genet.* **1999**, *37*, 176−190.

(35) de Groot, M. J.; Ackland, M. J.; Horne, V. A.; Alex, A. A.; Jones, B. C. Novel approach to predicting P450 mediated drug metabolism. The development of a combined protein and pharmacophore model for CYP2D6. *J. Med. Chem.* **1999**, *42*, 1515−1524.

(36) de Groot, M. J.; Ackland, M. J.; Horne, V. A.; Alex, A. A.; Jones, B. C. 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.* **1999**, *42*, 4062−4070.

(37) Conformational search using Monte Carlo sampling with the SYBYL force field. All resulting conformers within 25 kcal/mol of the lowest energy conformer were geometry-optimized using AM1 (optcycle = 1000, maxcycle = 1000, convergence flag on).

(38) *Spartan, SGI IRIX 6.2*, version 5.1.1 Open GL; Wavefunction Inc.: Irvine, CA, 1998.

(39) *SYBYL*, version 6.7; Tripos Inc.: St. Louis, MO, 2000.

(40) Because of the high amino acid homology between CYP2C9 and CYP2C5/3 (75.9% identity), the alignment between these two 2C-family members could be obtained manually. The crystal structures of CYP2C5/3 and CYP102 were structurally superimposed using SYBYL. The coordinates of the amino acids of the B′-region of CYP102 were then used as a template for the homology modeling while ensuring structural integrity (alignment in Figure 2). Using a cytochrome P450 from the same subfamily as a main template for homology modeling results in a more reliable alignment and consequently a more reliable model compared to the use of (several) bacterial P450s as templates (de Groot, unpublished results).

(41) Sali, A.; Blundell, T. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779−815.

(42) *Modeller*, version 3; Rockefeller University: New York, 1995.

(43) *Quanta97*, version 97.0305; Molecular Simulations Inc.: San Diego, CA, 1998.

(44) Laskowski, R. A.; McArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283−291.

(45) Rettie, A. E.; Wienkers, L. C.; Gonzalez, F. J.; Trager, W. F.; Korzekwa, K. R. Impaired (*S*)-warfarin metabolism catalyzed by the R144C allelic variant of CYP2C9. *Pharmacogenetics* **1994**, *4*, 39−42.

(46) Ibeanu, G. C.; Ghanayem, B. I.; Linko, P.; Li, L.; Pedersen, L. G.; Goldstein, J. A. Identification of residues 99, 220, and 221 of human cytochrome P450 2C19 as key determinants of omeprazole hydroxylase activity. *J. Biol. Chem.* **1996**, *271*, 12496−12501.

(47) Jung, F.; Griffin, K. J.; Song, W.; Richardson, T. H.; Yang, M.; Johnson, E. F. Identification of amino acid substitutions that confer a high affinity for sulphaphenazole binding and a high catalytic efficiency for warfarin metabolism to P450 2C19. *Biochemistry* **1998**, *37*, 16270−16279.

(48) Klose, T. S.; Ibeanu, G. C.; Ghanayem, B. I.; Pedersen, L. G.; Li, L.; Hall, S. D.; Goldstein, J. A. Identification of residues 286 and 289 as crucial for conferring substrate specificity of human CYP2C9 for diclofenac and ibuprofen. *Arch. Biochem. Biophys.* **1998**, *357*, 240−248.

(49) Haining, R. L.; Jones, J. P.; Henne, K. R.; Fisher, M. B.; Koop, D. R.; Trager, W. F.; Rettie, A. E. Enzymatic determinants of the substrate specificity of CYP2C9: Role of B′−C loop residues in providing the p-stacking anchor for warfarin binding. *Biochemistry* **1999**, *38*, 3285−3292.

(50) Ridderström, M.; Masimirembwa, C.; Trump-Kallmeyer, S.; Ahlefelt, M.; Otter, C.; Andersson, T. B. Arginines 97 and 108 in CYP2C9 are important determinants of the catalytic function. *Chem. Biol. Res. Commun.* **2000**, *270*, 983−987.

(51) Tsao, C.-C.; Wester, M. R.; Ghanayem, B.; Coulter, S. J.; Chanas, B.; Johnson, E. F.; Goldstein, J. A. Identification of human CYP2C19 residues that confer *S*-mephenytoin 4′-hydroxylation activity to CYP2C9. *Biochemistry* **2001**, *40*, 1937−1944.

(52) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.

(53) The protein model for CYP2C9 was reoriented in space using Quanta97. The plane of the heme was defined in the *XY*-plane with the center of the heme at the origin, the proximal side of the heme in the $+Z$ direction, and the propionate side chains of the heme in the $+X$ direction.

(54) Minimization options: SYBYL force field, Powell minimization with standard minimize details. Termination on gradient <0.1 kcal/(mol·Å), with as many iterations as required for convergence. A dummy "oxygen" atom was placed at a distance of 1.67 Å from the iron atom and was kept in place with a distance constraint (1.67 Å, constant = 100) and angle constraints $N_{(porphyrin)}$−Fe−dummy (66.0°, constant = 10); see Figure 3. Additional constraint during minimization: dummy "oxygen"−site of oxidation, 1.75 Å (constant = 50).

(55) Brickmann, J.; Heiden, W.; Moekel, G.; Reiling, S.; Vollhardt, H.; Zachmann, C.-D. Interactive visualization of molecular scenarios with MOLCAD/SYBYL. *Data Visualization Mol. Sci.* **1995**, 83−97.

(56) Nelson, D. R.; Koymans, L.; Kamataki, T.; Stegeman, J. J.; Feyereisen, R.; Waxman, D. J.; Waterman, M. R.; Gotoh, O.; Coon, M. J.; Estabrook, R. W.; Gunsalus, I. C.; Nebert, D. W. P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* **1996**, *6*, 1−42.

(57) Nelson, D. R.; Strobel, H. W. Secondary structure prediction of 52 membrane-bound cytochromes P450 shows a strong structural similarity to $P450_{cam}$. *Biochemistry* **1989**, *28*, 656−660.

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

(59) Korzekwa, K. R.; Krishnamachary, N.; Shou, M.; Ogai, A.; Parise, R. A.; Rettie, A. E.; Gonzalez, F. J.; Tracy, T. S. Evaluation of atypical cytochrome P450 kinetics with two-substrate models: evidence that multiple substrates can simultaneously bind to cytochrome P450 active sites. *Biochemistry* **1998**, *37*, 4137−4147.

(60) Hutzler, J. M.; Hauer, M. J.; Tracy, T. S. Dapsone activation of CYP2C9-mediated metabolism: evidence for activation of multiple substrates and a two-site model. *Drug Metab. Dispos.* **2001**, *29*, 1029−1034.

(61) Yamazaki, H.; Inoue, K.; Shimada, T. Roles of two allelic variants (Arg144Cys and Ile359Leu) of cytochrome P4502C9 in the oxidation of tolbutamide and warfarin by human liver microsomes. *Xenobiotica* **1998**, *28*, 103−115.

(62) Lüdemann, S. K.; Lounnas, V.; Wade, R. C. How do substrates enter and products exit the buried active site of cytochrome P450cam? 1. Random expulsion molecular dynamics investigation of ligand access channels and mechanisms. *J. Mol. Biol.* **2000**, *303*, 797−811.

(63) Afzelius, L.; Zamora, I.; Ridderström, M.; Andersson, T. B.; Karlen, A.; Masimirembwa, C. M. Competitive CYP2C9 inhibitors: enzyme inhibition studies, protein homology modeling, and three-dimensional quantitative structure−actvity relationship analysis. *Mol. Pharmacol.* **2001**, *59*, 909−919.

(64) Jones, B. C.; Hawksworth, G.; Horne, V. A.; Newlands, A.; Tute, M. S.; Smith, D. A. Putative active site model for P4502C9 (tolbutamide hydroxylase). *Br. J. Clin. Pharmacol.* **1993**, *34*, P143−P144.

(65) Goldstein, J. A.; de Morais, S. M. F. Biochemistry and molecular biology of the human CYP2C subfamily. *Pharmacogenetics* **1994**, *4*, 285−299.

(66) A homology model for CYP2C19 was constructed in a similar manner as described here for CYP2C9, including the quality control checks (data not shown). The active site of CYP2C19 is identical to that of CYP2C9 with the exception of the amino acids highlighted in Figure 2.

(67) Li, H.; Poulos, T. L. The structure of the cytochrome P450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid. *Nat. Struct. Biol.* **1997**, *4*, 140−149.

(68) Lüdemann, S. K.; Carugo, O.; Wade, R. C. Substrate access to cytochrome P450cam: a comparison of a thermal motion pathway analysis with molecular dynamics simulation data. *J. Mol. Model.* **1997**, *3*, 369−374.

(69) Lüdemann, S. K.; Lounnas, V.; Wade, R. C. How do substrates enter and products exit the buried active site of cytochrome P450cam? 2. Steered molecular dynamics and adiabatic mapping of substrate pathways. *J. Mol. Biol.* **2000**, *303*, 813−830.

(70) Podust, L. M.; Poulos, T. L.; Waterman, M. R. Crystal structure of cytochrome P450 14α-sterol demethylase (CYP51) from *Mycobacterium tuberculosis* in complex with azole inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 3068−3073.

(71) Weaver, R. J.; Webster, D.; Melvin, W. T.; Dickins, M.; Burke, M. D. Human cytochrome P4502C metabolism. *Br. J. Clin. Pharmacol.* **1993**, *36*, 167P.

(72) Venkatakrishnan, K.; Greenblatt, D. J.; von Moltke, L. L.; Shader, R. I. Alprazolam is another substrate for human cytochrome P450-3A isoforms. *J. Clin. Physiopharmacol.* **1998**, *18*, 256.

(73) Engel, G.; Hofmann, U.; Heidemann, H.; Cosme, J.; Eichelbaum, M. Antipyrine as a probe for human oxidative drug metabolism: identification of the cytochrome P450 enzymes catalyzing 4-hydroxyantipyrine, 3-hydroxmethylypyrine, and norantipyrine formation. *Clin. Pharmacol. Ther.* **1996**, *59*, 613−623.

(74) Bauer, E.; Guo, Z.; Ueng, Y.-F.; Bell, L. C.; Zeldin, S.; Guengerich, F. P. Oxidation of benzo[*a*]pyrene by recombinant human cytochrome P450 enzymes. *Chem. Res. Toxicol.* **1995**, *8*, 136−142.

(75) Watanabe, K.; Matsunaga, T.; Yamamoto, I.; Funae, Y.; Yoshimura, H. Involvement of CYP2C in the metabolism of cannabinoids by human hepatic microsomes from an old woman. *Biol. Pharm. Bull.* **1995**, *18*, 1138−1141.

(76) Fang, J.; Baker, G. B.; Coutts, R. T.; McKenna, K. F. Elucidation of individual cytochrome P450 isoenzymes involved in the metabolism of clozapine. In *Proceedings of the Seventh North American ISSX Meeting*, San Diego, CA, 1996; International Society for the Study of Xenobiotics; p 356.

(77) Mitra, A. K.; Kalhorn, T. F.; Thummel, K. E.; Unadkat, J. D.; Slattery, J. T. Metabolism of arylamines by human liver microsomal cytochrome P450(s). In *Proceedings of the Fifth North American ISSX Meeting*, Tucson, AZ, 1993; International Society for the Study of Xenobiotics; p 148.

(78) Ono, S.; Hatanaka, T.; Miyazawa, S.; Tsutsui, M.; Aoyama, T.; Gonzalez, F. J.; Satoh, T. Human liver microsomal diazepam metabolism using cDNA-expressed cytochrome P450s: role of CYP2B6, 2C19 and the 3A subfamily. *Xenobiotica* **1996**, *26*, 1155−1166.

(79) Tracy, T. S.; Marra, C.; Wrighton, S. A.; Gonzalez, F. J.; Korzekwa, K. R. Studies of flurbiprofen 4′-hydroxylation. Additional evidence suggesting the sole involvement of cytochrome P450 2C9. *Biochem. Pharmacol.* **1996**, *52*, 1305−1309.

(80) Schmider, J.; Greenblatt, D. J.; von Moltke, L. L.; Karsov, D.; Vena, R.; Friedman, H. L.; Schader, R. I. Biotransformation of mestranol to ethinyl estradiol *in vitro*: the role of cytochrome P-450 2C9 and metabolic inhibitors. *J. Clin. Pharmacol.* **1997**, *37*, 193−200.

(81) Miners, J. O.; Coulter, S.; Tukey, R. H.; Veronese, M. E.; Birkett, D. J. Cytochromes P450 1A2, and 2C9 are responsible for the human hepatic O-demethylation of *R*- and *S*-naproxen. *Biochem. Pharmacol.* **1996**, *51*, 1003−1008.

(82) Hargreaves, J. A.; Howald, W. N.; Racha, J. K.; Levy, R. H. Identification of enzymes responsible for the metabolism of phenobarbital. In *Proceedings of the Seventh North American ISSX Meeting*, San Diego, CA, 1996; International Society for the Study of Xenobiotics; p 259.

(83) Miners, J. O.; Rees, D. L. P.; Valente, L.; Veronese, M. E.; Birkett, D. J. Human hepatic cytochrome P450 2C9 catalyzes the rate-limiting pathway of torsemide metabolism. *J. Pharmacol. Exp. Ther.* **1995**, *272*, 1076−1081.

(84) Machinist, J. M.; Mayer, M. D.; Shet, M.; Ferrero, J. L.; Rodrigues, A. D. Identification of the human liver cytochrome P450 enzymes involved in the metabolism of zileuton (ABT-077) and its N-dehydroxylated metabolite, Abbott-66193. *Drug Metab. Dispos.* **1995**, *23*, 1163−1174.
(85) Lang, D.; Rettie, A. E.; Böcker, R. H. Identification of enzymes involved in the metabolism of atrazine, terbuthylazine, ametryne, and terbutryne in human liver microsomes. *Chem. Res. Toxicol.* **1997**, *10*, 1037−1044.
(86) Waxman, D. J.; Chang, T. K. H.; Chen, G. Interactions of anticancer drugs with cytochromes P450 and glutathione *S*-transferases: drug activation and drug resistance mechanisms. In *Proceedings of the Fifth North American ISSX Meeting*, Tucson, AZ, 1993; International Society for the Study of Xenobiotics; p 2.
(87) Yasumori, T.; Murayama, N.; Yamazoe, Y.; Kato, R. Polymorphism in hydroxylation of mephenytoin and hexobarbital stereoisomers in relation to hepatic P-450 human-2. *Clin. Pharmacol. Ther.* **1990**, *47*, 313−322.
(88) Komatsu, T.; Yamazaki, H.; Asahi, S.; Gillam, E. M. J.; Guengerich, F. P.; Nakajima, M.; Yokoi, T. Formation of dihydroxy metabolite of phenytoin in human liver microsomes/cytosol: roles of cytochromes P450 2C9, 2C19, and 3A4. *Drug Metab. Dispos.* **2000**, *28*, 1361−1368.
(89) Yamazaki, H.; Shimada, T. Progesterone and testosterone hydroxylation by cytochromes P450 2C19, 2C9, and 3A4 in human liver microsomes. *Arch. Biochem. Biophys.* **1997**, *346*, 161−169.
(90) Komatsu, K.; Ito, K.; Nakajima, Y.; Kanamitsu, S.-I.; Imaoka, S.; Funae, Y.; Green, C. E.; Tyson, C. A.; Shimada, N.; Sugiyama, Y. Prediction of in vivo drug−drug interactions between tolbutamide and various sulfonamides in humans based on in vitro experiments. *Drug Metab. Dispos.* **2000**, *28*, 475−481.
(91) Wester, M. R.; Lasker, J. M.; Johnson, E. F.; Raucy, J. L. CYP2C19 participates in tolbutamide hydroxylation by human liver microsomes. *Drug Metab. Dispos.* **2000**, *28*, 354−359.
(92) Weber, C.; Gasser, R.; Hopfgartner, G. Absorption, excretion, and metabolism of the endothelin receptor antagonist bosentan in healthy male subjects. *Drug Metab. Dispos.* **1999**, *27*, 810−815.
(93) Lausecker, B.; Hess, B.; Fischer, G.; Mueller, M.; Hopfgartner, G. Simultaneous determination of bosentan and its three major metabolites in various biological matrices and species using narrow bore liquid chromatography with ion spray tandem mass spectroscopic detection. *J. Chromatogr., B* **2000**, *749*, 67−83.
(94) Gentest homepage: http//www.gentest.com.
(95) Warrington, J. S.; Shader, R. I.; von Moltke, L. L.; Greenblatt, D. J. In vitro biotransformation of sildenafil (Viagra): identification of human cytochromes and potential drug interactions. *Drug Metab. Dispos.* **2000**, *28*, 392−397.
(96) Hyland, R.; Roe, E. G. H.; Jones, B. C.; Smith, D. A. Identification of the cytochrome P450 enzymes involved in the *N*-demethylation of sildenafil. *Br. J. Clin. Pharmacol.* **2001**, *51*, 1−11.
(97) Bajpai, M.; Roskos, L. K.; Shen, D. D.; Levy, R. H. Roles of cytochrome P4502C9 and cytochrome P4502C19 in the stereoselective metabolism of phenytoin to its major metabolite. *Drug Metab. Dispos.* **1996**, *24*, 1401−1403.
(98) Nicholls, A.; Sharp, K. A.; Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct., Funct., Genet.* **1991**, *11*, 281−296.

JM0110791