Studies of Binding Modes of (S)-Mephenytoin to Wild Types and Mutants of Cytochrome P450 2C19 and 2C9 Using Homology Modeling and Computational Docking

Akifumi Oda[1,2,3], Noriyuki Yamaotsu[2], and Shuichi Hirono[2]

Received June 9, 2004; accepted August 22, 2004

Purpose. This study investigated the structural features of CYP2C19 complexed with (S)-mephenytoin, using computational methods. In addition to wild-type CYP2C19 proteins (1A and 1B), which have selective 4 -hydroxylase activities of (S)-mephenytoin, CYP2C19 mutants were also studied, together with a wild type and artificial mutants of CYP2C19.

Methods. Three-dimensional structures of wild-type and mutant proteins of CYP2C19 and CYP2C9 were estimated from homology modeling using the crystal structure of rabbit CYP2C5 as a reference. The binding mode of (S)-mephenytoin to CYP2C19 was investigated using computational docking.

Results. The results reproduced the specific bindings between (S)-mephenytoin and the wild types of CYP2C19. Our findings suggest that Asp293 of CYP2C19 plays an important role in the binding of (S)-mephenytoin, which was surrounded by Val113 and Ala297, and points the phenyl ring at the heme iron. In addition the wild types of CYP2C19, the computational docking studies also accounted for the experimental activities of CYP2C19 mutants, and wild-type and mutant CYP2C19 proteins.

Conclusions. These results confirm that the predicted three-dimensional structure of the CYP2C19-(S)-mephenytoin complex is reasonable, and that this strategy is useful for investigating complex structures. Virtual screening for drug discovery can also be carried out using these methods.

KEY WORDS: binding mode; computational docking; cytochrome P450 2C19; homology modeling.

INTRODUCTION

Cytochrome P450 (CYP) [1]plays important roles in the metabolism of a wide variety of xenobiotic and endogenous compounds, including clinically important drugs. CYP2C19 (1) is one of the major enzymes that metabolizes drugs in the human liver and has an important role in the metabolism of (S)-mephenytoin (2,3). In 1984, Küpfer and co-workers observed that urinary recovery of 4′-hydroxylated (S)-mephenytoin was not detected for approximately 3–5% of Caucasians, which attracted attention to the identification of polymorphisms of CYP2C19 (4). A number of gene mutations in CYP2C19 were found by analyzing the genes of poor metabolizers (PMs) (5–11) (Table I). Almost all Oriental PMs can be described by two defective alleles (*CYP2C19*2A* and *CYP2C19*3*), but other mutant alleles are also important for Caucasian PMs. CYP2C9 is another enzyme that plays a significant role in drug metabolism (12). Although it metabolizes a variety of drugs, its 4′-hydroxylase activity of (S)-mephenytoin is negligible. In order to identify the key residues of CYP2C19 for 4′-hydroxylase activity, Tsao and co-workers constructed chimeras by replacing portions of CYP2C9 with those of CYP2C19, mutating individual residues by site-directed mutagenesis and assessing (S)-mephenytoin 4′-hydroxylase activity (13). They found that the mutation of six residues of CYP2C9 to imitate CYP2C19 resulted in 6% of the activity of wild-type CYP2C19.

To date, the three-dimensional structure of CYP2C19 has not yet been determined. The mechanism by which (S)-mephenytoin complexes with CYP2C19 is also unknown. However, the ligand-free crystal structure of one of the enzymes belonging to the CYP2C subfamily (rabbit CYP2C5) has been obtained (14). This provides a model upon which to base structurally unknown human CYPs and their ligands using computational methods.

The homology modeling method is the most powerful computational approach to predict the three-dimensional structures of proteins based on sequence similarities with structurally known proteins. This method estimates a three-dimensional model of a protein from the known structures of homologous proteins, and it is based on the assumption that a three-dimensional structure of one protein is analogous to the others, which have similar sequences. In addition to the structure prediction, computational ligand docking might also be used for structure-based drug design (15). This is one of the optimization problems, in which atoms of the ligand molecule are positioned into points in ligand-binding pockets of the target biomolecule. The difficulties of these problems exponentially increase in terms of the number of ligand atoms. Even if only the grid points of three-dimensional space are considered, these problems are NP-hard. For this problem, several approaches—for example, probabilistic search methods and sophisticated data structures—were proposed (16–18).

Previous structures that have been proposed for CYP2C19 complexed with (S)-mephenytoin using homology modeling and computational docking methods (19,20) were based on CYP102 (P450-BM3). The homology between CYP102 and CYP2C19 is lower than between CYP2C19 and CYP2C5. Although the homology modeling and docking with ligands for CYP2C19 were carried out using CYP2C5 as template (21–23), the docking of (S)-mephenytoin has not yet been examined. Furthermore, these studies only looked at wild-type CYP2C19 and not related mutants.

In this study, a three-dimensional structure of wild-type CYP2C19 complexed with (S)-mephenytoin was constructed using homology modeling and computational docking procedures. In addition to wild-type CYP2C19, homology modeling and docking studies were also attempted for mutant forms of the protein. Docking studies of a wild type and mutants of CYP2C9 were also used to investigate important structural properties for (S)-mephenytoin docking. Homology modeling calculations were generally carried out only for wild-type proteins, and this is the first study of the homology modeling and

[1] Discovery Laboratories, Toyama Chemical Co. Ltd., 2-4-1 Shimookui, Toyama 930-8508, Japan
[2] School of Pharmaceutical Sciences, Kitasato University, 5-9-1 Shirokane, Minato-ku, Tokyo 108-8641, Japan
[3] To whom correspondence should be addressed. (e-mail: AKIFUMI_ODA@toyama-chemical.co.jp)

0724-8741/04/1200-2270/0 © 2004 Springer Science+Business Media, Inc.

Table I. CYP2C19 Alleles

Allele	Enzyme activity	Effect of nucleotide changes	Name of protein
*2C19*1A*	Active	—	2C19.1A
*2C19*1B*	Active	Ile331Val	2C19.1B
*2C19*2A*	Inactive	Splicing defect	—
*2C19*2B*	Inactive	Splicing defect	—
*2C19*3*	Inactive	Stop codon	—
*2C19*4*	Inactive	Initial codon	—
*2C19*5A*	Inactive	**Arg433Trp**[b]	2C19.5A
*2C19*5B*	Inactive	Ile331Val, **Arg433Trp**[b]	2C19.5B
*2C19*6*	Inactive	**Arg132Gln,**[b] Ile331Val	2C19.6
*2C19*7*	Inactive	Exon skipping	—
*2C19*8*	Active/inactive[a]	**Trp120Arg**[b]	2C19.8

[a] 2C19.8 is active *in vitro* although it is inactive *in vivo*.
[b] Mutations without Ile331Val, which does not affect enzyme activities in 2C19.1B, are described by bold letter.

docking of mutant CYPs in order to elucidate the binding mode of the substrate. The availability of the modelings of mutants for the binding mode prediction is also discussed.

MATERIALS AND METHODS

Proteins

There are two ways to validate the availability of docking methods. One is the way in which several proteins are used, and another is the way in which several small molecules are used. Because one of the central purposes of this study is the proposal of the method for predicting the complex structures by using various mutant proteins, identical ligand molecule, (S)-mephenytoin, was docked into various target proteins. Furthermore, multi proteins were desired also for investigations of the availability of homology modeling for mutants.

Two wild-type and four mutant CYP2C19 proteins, along with one wild type and four artificial mutants of CYP2C9 constructed by Tsao and co-workers, were used for homology modeling and computational docking studies (Tables I and II). Wild-type enzymes that are expressed from alleles *CYP2C19*1A* and *CYP2C19*1B* are described in Roman script (that is, CYP2C19.1A and CYP2C19.1B). The proteins expressed from the mutant alleles *CYP2C19*5A*, *CYP2C19*5B*, *CYP2C19*6* and *CYP2C19*8* are described as CYP2C19.5A, CYP2C19.5B, CYP2C19.6, and CYP2C19.8, respectively. Four artificial mutants of CYP2C9 were used: protein **1**, which had four mutant residues (2C9/I99H, S220P, P221T and S286N); protein **2**, which had five mutant residues (2C9/I99H, S220P, P221T, S292A and F295L); protein **3**, which had five mutant residues (2C9/I99H, S220P, P221T, S286N and F295L); and protein **4**, which had six mutant residues (2C9/I99H, S220P, P221T, S286N, V292A and F295L). All mutants of CYP2C9 are not naturally occurring. In Tables I and II, amino-acid mutations are also described (note that although CYP2C19.1B has an Ile331 to Val mutation, there is no loss of (S)-mephenytoin 4′-hydroxylase activity. In Table I, bold type letters illustrate mutations without Ile331Val). (S)-Mephenytoin 4′-hydroxylase activities were experimentally observed only for CYP2C19.1A, CYP2C19.1B, CYP2C19.8, and protein **4**. Although the enzyme activity of protein **4** is 6% of CYP2C19, it is much larger than the other mutants (e.g. enzyme activity of CYP2C19.5A is about 0.37% of that of CYP2C19.1A). In this study, because only binding modes are discussed, we treat the protein **4** as active enzyme.

Homology Modeling

In this study, we used the crystal structure of rabbit CYP2C5 (PDB code: 1DT6) (14) as the template for homology modeling of CYP2C19 and CYP2C9. The basic local alignment search tool (BLAST) based on finite automaton was used for the sequence alignments of the target (CYP2C19 and CYP2C9) against template (CYP2C5). BLOSUM62 was used as alternate scoring matrix for BLAST. The alignment produced by BLAST was used as the input for the FAMS program (24). In FAMS, the modeling of C_α atoms, main-chain atoms and side-chain atoms are carried out sequentially. For modeling calculations, FAMS can fully automatically perform both the database search and optimization steps by using the simulated annealing (SA) method. Even if the positions of side-chain atoms extracted from databases are not adequate—for example, if they have a conflict of main chain atoms—FAMS can refine the structure by using SA. Although homology models for wild types and mutants of CYP2C19 and CYP2C9 were obtained by FAMS, there were no hydrogen atoms in these model structures because hydrogens were not observed in the template of modeling (the experimental structure of CYP2C5 produced by X-ray crystallography). In this study, constructed homology models were protonated by the "protonate" module in AMBER 6.0 (25), and their structures were refined by AMBER PARM94 force fields. Only 1000 cycle calculations were carried out for refinements, because these calculations were performed only for reducing structural distortions. The explicit water molecules and cyclic-boundary conditions were not adopted for these refinements; a cutoff distance of 12 Å for non-bonded interactions and a distance-dependent dielectric coefficient $\varepsilon = 80R$ (R: distance between two interacting atoms) were used. For these calculations, the steepest-descent minimizer was used in the first 10 steps, and the rest of the calculations were carried out using the conjugate-gradient method.

Table II. CYP2C9 Mutants

Construct	Enzyme activity	Abbreviated name of protein
CYP2C9	Inactive	CYP2C9
CYP2C9/I99H, S220P, P221T, S286N	Inactive	**1**
CYP2C9/I99H, S220P, P221T, V292A, F295L	Inactive	**2**
CYP2C9/I99H, S220P, P221T, S286N, F295L	Inactive	**3**
CYP2C9/I99H, S220P, P221T, S286N, V292A, F295L	Active	**4**

Generation of a Conformer Set for the Ligand Molecule

The structure of the ligand molecule (S)-mephenytoin optimized by *ab initio* HF/6-31G* method was used for computational docking studies. Atomic charges of the ligand molecule were obtained by ESP method, which can reproduce the electrostatic potential calculated by *ab initio* HF/6-31G* (26). Gaussian 98 was used for quantum chemical calculations (27). In the complex structure, because the conformation of a ligand molecule is not always the most stable conformer, we considered many metastable states of ligands for computational docking studies. For this purpose, a calculation of only the most stable conformation is not sufficient, and a conformer set consisting of many metastable conformers of a ligand is required. In this study, CAMDAS, which is a program developed by this laboratory (28) for the generation of ligand conformer sets, was used to calculate a conformer set for (S)-mephenytoin. Molecular dynamics (MD) simulations at high temperatures were carried out by this program, and the conformers obtained in the trajectory of this simulation were sampled. The clustering procedure was carried out for all sampled conformers, and the representative conformers were picked up from the clusters. In the CAMDAS calculation, a modified MMFF force field was used (29), in which electrostatic interactions were omitted and the weights of both angle bending and dihedral terms equal to 0.8 were included. The temperature of MD simulation by CAMDAS was 1200 K and 1 ns (1,000,000 steps) calculation was carried out. Chiral inversions of (S)-mephenytoin were not allowed through the MD simulation. The two dihedral angles shown in Fig. 1 were used as explanatory variables for the clustering. For (S)-mephenytoin, the conformer set in which 31 conformers were included was obtained by CAMDAS calculation.

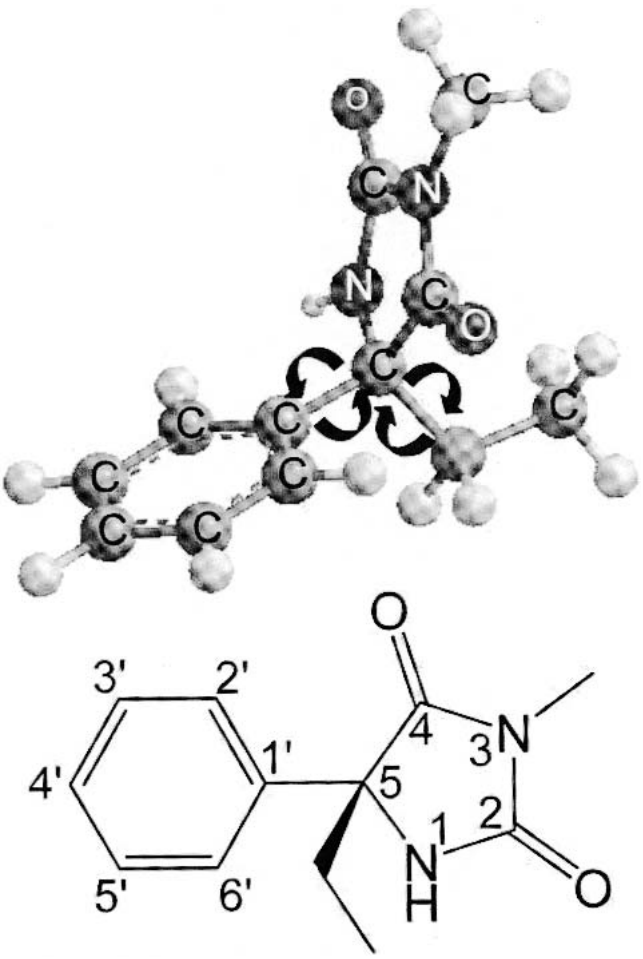


Fig. 1. Two torsional angles for clustering of (S)-mephenytoin conformers. The two torsional angles illustrated in this figure were used as criteria for the clustering of conformers.

Computational Docking

For computational ligand docking, ligand-binding pockets in the homology models of the proteins were explored using the SiteID module in the SYBYL program package (Tripos Inc., St. Louis, MO, USA). SiteID can search for potential binding sites within or on the proteins by using the floodfill-solvation technique. The grid method with default parameters was used in this study. MOLCAD surfaces for the amino-acid residues within a 6 Å radius of explored binding pockets were calculated using the MOLCAD program (Tripos Inc.). These surfaces were used as steric constraints for computational docking trials. In the obtained ligand-binding pocket of the proteins, hydrogen-bond acceptor and donor site queries were defined by the UNITY program (Tripos Inc.). In this procedure, some queries were deleted because of VDW bumps and we used only the remaining queries for computational docking in the next step. We adopted tolerances of 1.5 Å for the spatial constraints of all hydrogen-bond queries. Using the steric constraint and hydrogen-bond queries, each conformer from the conformer set of (S)-mephenytoin generated by CAMDAS was docked into the homology models of the CYPs by the UNITY 3D program. For wild-type CYP2C19.1A, docking models that fulfilled at least one of the hydrogen-bond queries ("partial match") were adopted. For the other proteins, only one query that plays an important role in docking models of wild-type CYP2C19.1A was considered. The steric constraint was constructed from the MOLCAD surface with default parameters. Using these strategies, many model structures of the complex were obtained for one ligand-protein system; for example, 31 models were generated by the docking calculation of (S)-mephenytoin with CYP2C19.1A. In computational docking studies, we need to determine which models can be adopted as candidates for the complex structure. Usually, the scoring functions that calculate the binding free energies of ligand–protein systems were used for this purpose, and the models that had smaller scores (that is, smaller binding free energy) were accepted as more reasonable models. Although many types of scoring functions were introduced (16,18,30–32), each type has different strengths and weaknesses. Recently, a method for evaluating docking models called "consensus scoring" has been introduced (33,34). In consensus scoring, multiple scoring functions were used to cover the shortcomings of each method. In this study, the validities of the calculated docking models were evaluated using CScore (Tripos Inc.) (34), which is one of the consensus-scoring programs. It includes five scoring functions: FlexX score (16), Gold score (18), PMF score (30), Dock score (31), and ChemScore (32). These five scores were independently calculated for docking models, and awarded marks out of five for the validities of the models. The models that are given high scores by CScore are

the possible candidates of the structures of protein–ligand complexes. However, when docking models with high CScore values were not obtained, we assumed that the computational docking trials for the ligand–protein systems could not be successfully carried out.

MD Simulation

The homology model the docking results of which were not reasonable were further analyzed by structural refinement by the MD simulation with explicit water molecules. The rectangular solid box in which the complex was included was set up and the box was filled with TIP3P water molecules (35). Sodium ions were used as counter ions for setting the total charge of the system to 0. The 1 ns MD simulation at 300 K was carried out using the particle-mesh Ewald method for calculations of electrostatic interactions under cyclic-boundary conditions by the AMBER 6.0 program. The bond lengths between two atoms were constrained by SHAKE (36). The cutoff distance for van der Waals interactions was 9.0 Å, and the MD simulation was carried out under isothermal and isochoric conditions. After the MD calculation, 100,000 steps of structure refinement were performed using steepest-descent and conjugate-gradient optimizers. The first 5000 steps of refinement were calculated by the steepest-descent method, and the rest of the calculation was carried out using the conjugate-gradient method, until the energy gradient become less than 0.1 kcal/molÅ. Explicit TIP3P waters were used in this refinement with the particle-mesh Ewald method under the cyclic-boundary condition. A cutoff distance of 9.0 Å for van der Waals interactions and no constraint of bond lengths were used for the refinement. After this calculation, an additional computational docking study was carried out for refined structure. The same computational conditions as previous docking calculations were used in the additional docking.

For FAMS calculations we used a personal computer with one Pentium3 600 MHz processor and Red Hat Linux 6.1. For the other calculations, an Octane2 workstation by SGI with dual MIPS R12000 400 MHz processors and IRIX release 6.5 was used.

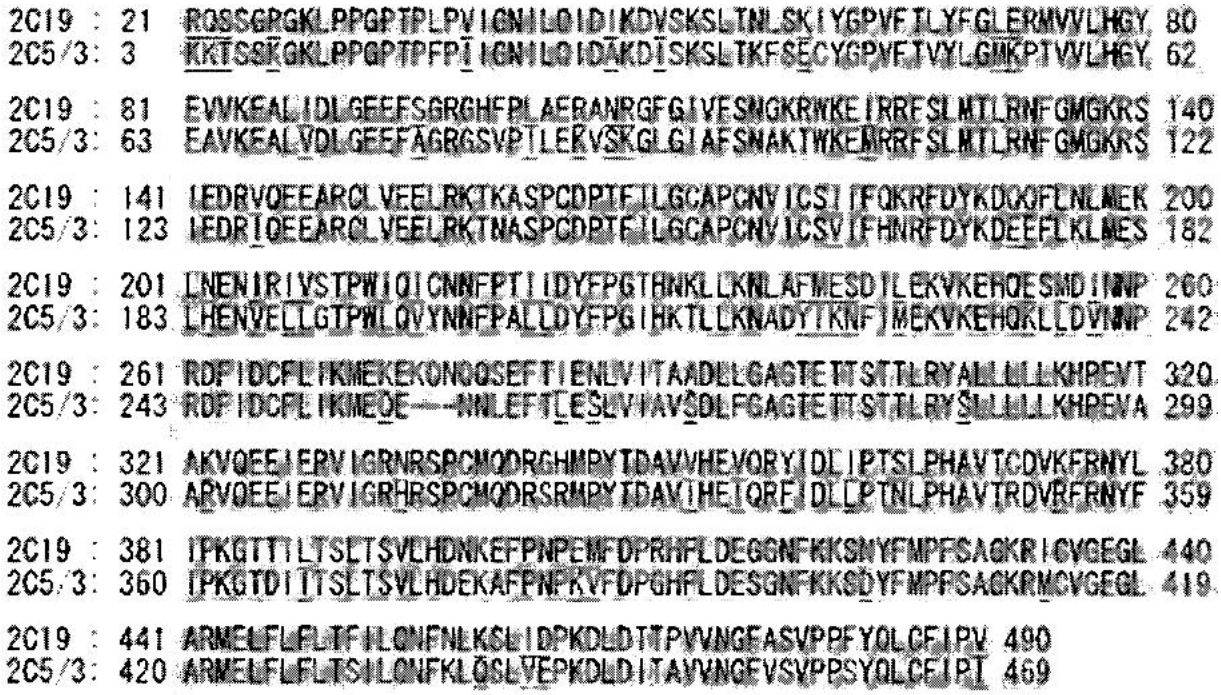
RESULTS

Sequence alignments of CYP2C19.1A and CYP2C9 against CYP2C5/3LVdH using BLAST are illustrated in Fig. 2. Similarities in amino-acid sequences are shown in Table III. In aligned sequences, identical amino acids were defined as "identity", and the pairs of residues whose score were positive in BLOSUM62 were defined as "positive". Although there are three insertions between helix H and I in both CYP2C19.1A and CYP2C9, they have a high degree of homology with CYP2C5/3LVdH. In this study, these alignments were used for homology modeling calculations of not only wild types but also mutants of CYP2C19 and CYP2C9. In Fig. 3, structurally optimized homology models for wild types and mutants of CYP2C19 are illustrated. The wild-type CYP2C9 and artificial mutant protein **4** are also shown in Fig. 4. In these figures, mutated residues in proteins without wild-type CYP2C19.1A and CYP2C9 are also shown.

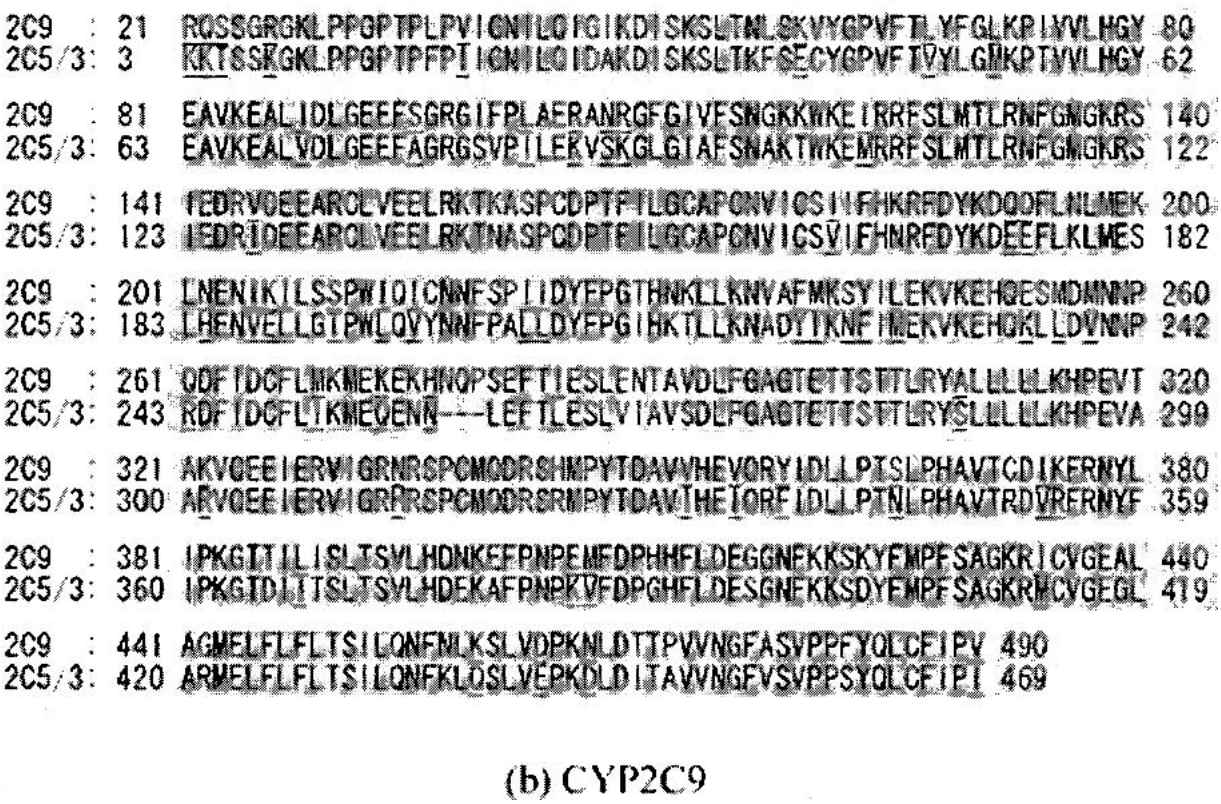
The ligand-binding pocket found by SiteID for the CYP2C19.1A homology model is illustrated in Fig. 5. Although several candidates of binding pockets were obtained by SiteID, the adjacent pocket of heme is shown in this figure because we know that active sites of CYPs are near the heme region. As shown in this figure, the ligand-binding pocket in the homology model was mainly composed of hydrophobic amino acids such as Val, Ile, Phe, Leu and Ala. The hydrophobic feature of the pocket of CYP2C19.1A was consistent with the known crystallographic structures of other CYPs and it is intimately related with the property of CYPs, which bind with hydrophobic substrates and oxidatively metabolize them. The hydrogen-bond queries in this region, which were three hydrogen-bond acceptor sites, were also illustrated. They were associated with the main-chain and side-chain oxygen atoms of Asp293 and the main-chain carbonyl oxygen of Gly296, respectively. Although all of them were directed to the heme moiety of CYP2C19.1A, the query from the carbonyl oxygen atom in the main chain of Asp293 was closest to the heme iron. Some of these queries might play important roles in the docking of (S)-mephenytoin in CYP2C19.

Using the hydrogen-bond queries and steric constraints calculated from the MOLCAD surface, a computational docking trial of (S)-mephenytoin was carried out for wild-type CYP2C19.1A. The conformer set of (S)-mephenytoin included 31 conformers, which were used for docking studies. Figure 6 shows one of the docking models with CScore values equal to 5 (that is, full points). The hydrogen bond between the amide hydrogen adjacent to the nitrogen atom of (S)-mephenytoin and the carbonyl oxygen in the main chain of Asp293 in CYP2C19.1A played a significant role in forming this docking model. The 4′-hydrogen of (S)-mephenytoin in this model was positioned near the heme iron. (S)-Mephenytoin is surrounded by hydrophobic amino acids such as Val113, Ile205, Ala292, Ala297, Leu366, and Phe476.

Computational docking trials for proteins other than CYP2C19.1A with (S)-mephenytoin were also carried out by using SiteID and UNITY (Table IV). These protein models contained ligand-binding pockets that were mainly composed of hydrophobic amino acids, which were similar to CYP2C19.1A. Although Asp 293 was conserved in all proteins, hydrogen-bond queries constructed from the carbonyl oxygen of Asp293 in CYP2C19.6, CYP2C9, proteins **1** and **3** were omitted by the VDW bump check. The bumps were caused by the differences of the shapes of ligand-binding sites between CYP2C19.1A and these proteins, and the shapes of the sites were affected by the mutations (CYP2C19.6) or differences in amino-acid sequences between CYP2C19 and CYP2C9 (CYP2C9, proteins **1** and **3**). These proteins were all experimentally observed to be inactive for (S)-mephenytoin 4′-hydroxylation. Computational docking calculations for another six proteins, including the hydrogen-bond queries of Asp293, were carried out using the same procedure of CYP2C19.1A. The obtained docking models were evaluated by CScore. When no models had high CScore values for a certain complex system, we considered that the docking trial had failed, even if docking models were obtained. For four of these six proteins (CYP2C19.1B, CYP2C19.8, proteins **2** and **4**) computational docking models were obtained in which 4′-hydrogens of (S)-mephenytoin were positioned nearby heme irons and with CScore values of 5 (that is, full points). For three of these four proteins (CYP2C19.1B, CYP2C19.8 and protein **4**) the enzyme activities were experimentally observed and computational docking studies seemed able to account for the experimental results. In these complex models,



(a) CYP2C19



(b) CYP2C9

Fig. 2. Sequence alignments of CYP2C19 and CYP2C9 against CYP2C5/3LVdH. The identical residues were expressed by meshed letters and the underlined residues were positives.

the ligand (S)-mephenytoin was surrounded by hydrophobic amino acids such as Val113, Ile205, Ala297 and Phe476, which were conserved in all wild types and mutants in this study. Furthermore, proteins for which the solutions of CScore = 5 were not obtained by computational docking trials based on the hydrogen-bond query from carbonyl oxygen of Asp293 (that is, CYP2C19.5A and CYP2C19.5B) were experimentally inactive mutants.

Table III. Identities and Positives[a] Between Amino-Acid Sequences of CYP2Cs

	CYP2C5/3LVdH	CYP2C19	CYP2C9
CYP2C5/3LVdH	100%	75%	75%
CYP2C19	87%	100%	92%
CYP2C9	87%	96%	100%

[a] Identities and positives were described in upper right and lower left portions, respectively.

By contrast, for the experimentally inactive protein **2**, docking models with CScore values of 5 were also obtained. Thus, we carried out more structure refinement for ligand-free protein **2** using MD simulation, steepest-descent and conjugate-gradient optimizers. After the refinement, we also performed additional computational docking for the refined structure with (S)-mephenytoin. Figure 7 shows the three-dimensional structures of protein **2** before and after the refinement. This reveals that structural features, such as the position of heme and the shape of the ligand-binding pocket around Asp293, were changed by MD calculations. After the structure refinement, no solutions of the docking trials of (S)-mephenytoin with protein **2** were obtained under the assumption that the hydrogen bond between carbonyl oxygen of Asp293 and (S)-mephenytoin are indispensable to ligand docking.

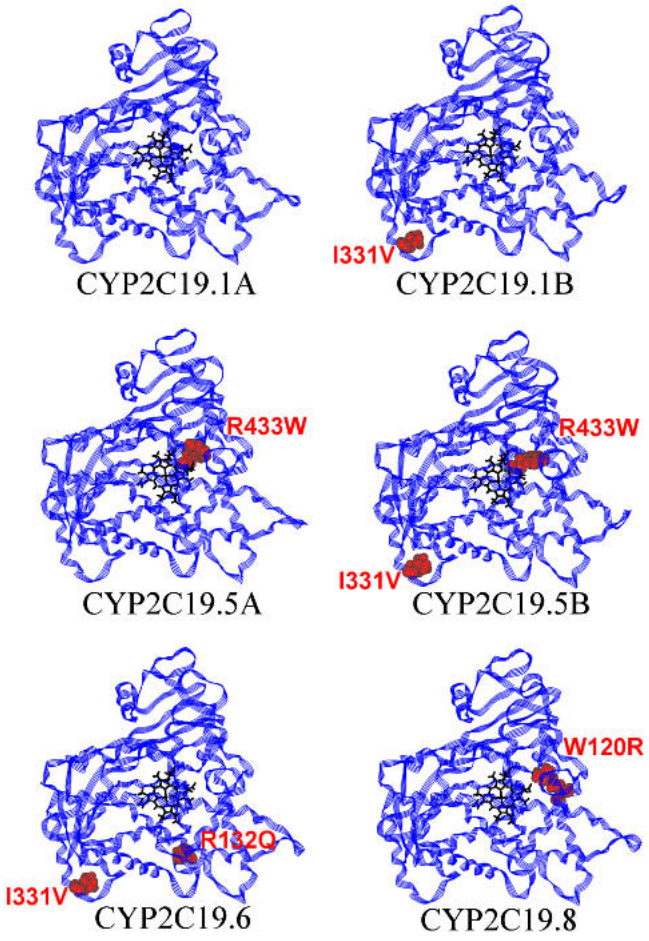


Fig. 3. Homology model for wild types and mutants of CYP2C19.

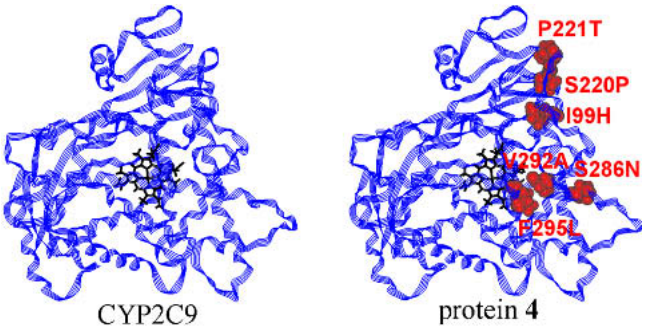


Fig. 4. Homology model for wild types and mutants of CYP2C9.

DISCUSSION

Homology models illustrated in Figs. 3 and 4 suggest that even if the mutated portions are not near the active sites (i.e., in adjacent regions of heme) mutations of one or two residues are still capable of inactivating CYP2C19. For example, Ile331Val, which does not affect the enzyme activity in CYP2C19.1B, and also mutations that play important roles for inactivities of mutants of CYP2C19 (Arg132Gln in CYP2C19.6), were positioned far from the ligand-binding pocket. Furthermore, although Arg433Trp in both CYP2C19.5A and CYP2C19.5B is located near by heme moiety, this residue is far away from the iron atom and does not seem to be directly implicated in the ligand docking. This indicates that mutations change the conformations of ligand-binding pockets of proteins, and affect the enzyme activities even if the mutations are located far from the ligand-binding pocket. By contrast, Ile99His, Ser286Asn, Val292Ala and Phe295Leu in protein **4** existed near the active site in our homology model. In particular, Ser286Asn, Val292Ala and Phe295Leu belong to the helix I, and these positions are also members of the substrate-recognition site 4 (SRS-4) (37). Therefore, these mutations might be directly implicated in (S)-mephenytoin docking for protein **4**.

In the model of CYP2C19.1A complex with (S)-

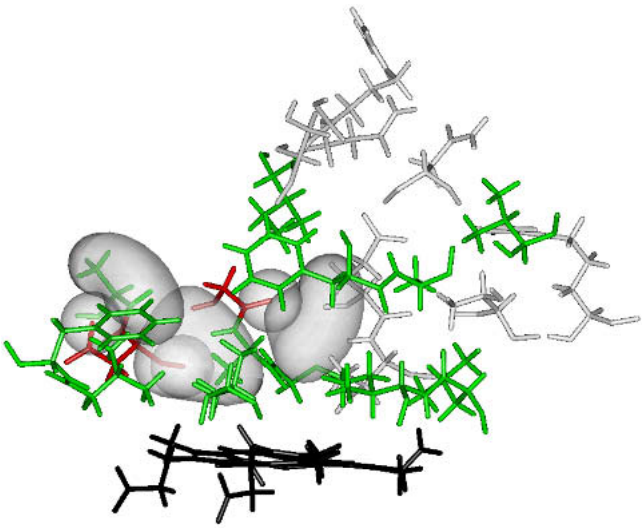


Fig. 5. Active-site residues explored by Site ID and hydrogen-bond queries in CYP2C19.1A. Hydrophobic amino acids are colored green; Asp293 and Gly296 are illustrated in red. Three hydrogen-bond queries were described with spatial constraint.

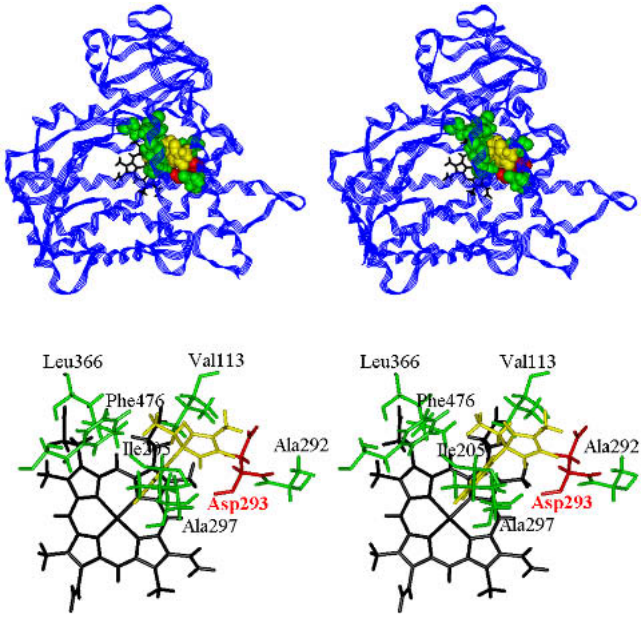


Fig. 6. CYP2C19.1A complex with (S)-mephenytoin. (S)-Mephenytoin, Asp293, and hydrophobic amino acids were illustrated in yellow, red, and green, respectively.

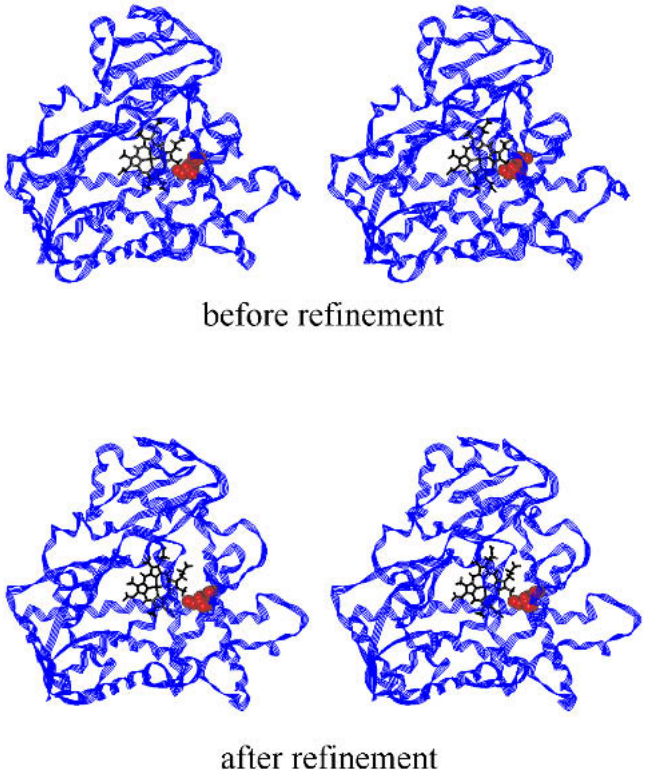


Fig. 7. Three-dimensional structures of protein **2** before and after refinement by MD simulation, steepest-decent and conjugate-gradient optimizers. The two structures were aligned by the fitting of their main chains. The RMSD of the protein **2** structure before and after refinement was 3.524 Å. The red colored residue was Asp293.

mephenytoin shown in Fig. 6, the orientation of (S)-mephenytoin whose 4′-hydrogen was located near the heme iron is consistent with (S)-mephenytoin 4′-hydroxylase activity of CYP2C19.1A and suggests that the three-dimensional structure of the calculated docking model is reasonable. The result that (S)-mephenytoin was surrounded by hydrophobic amino acids indicates that this hydrophobicity plays an important role for (S)-mephenytoin docking in the binding pocket of CYP2C19.1A. Significantly, Ala297 and Phe476 were located nearby the phenyl ring of ligand, suggesting that the CH-π interaction between the phenyl ring and Ala297, and/or the π-π stacking between ligand and Phe476, are important for the location of (S)-mephenytoin in CYP2C19.1A. Furthermore, Asp293 and Ala297, which seem to be important residues for ligand docking, are members of helix I and SRS-4. This is consistent with the results of Tsao *et al.* (11)

Table IV. The Results of Computational Docking

Protein	Experimental 4′-hydroxylase activity	Hydrogen bond query on Asp293[a]	Maximum CScore value[b]
2C19.1A	Active	OK	5
2C19.1B	Active	OK	5
2C19.5A	Inactive	OK	3
2C19.5B	Inactive	OK	4
2C19.6	Inactive	Bump	—
2C19.8	Active (*in vitro*)	OK	5
2C9	Inactive	Bump	—
1	Inactive	Bump	—
2	Inactive	OK	5
3	Inactive	Bump	—
4	Active	OK	5

[a] The result of VDW bump checks for hydrogen-bond query on Asp293.
[b] The best CScore value of those for all docking models.

For other proteins than CYP2C19.1A, hydrogen-bond queries constructed from the carbonyl oxygen of Asp293 in CYP2C19.6, CYP2C9, proteins **1** and **3** were omitted by the VDW bump check, and there were no docking models whose CScore values were 5 in CYP2C19.5A and CYP2C19.5B. For the refined model of protein **2**, no solutions were obtained by the computational docking. Because all these proteins were experimentally observed enzyme inactivities without exceptions, these results support that Asp293 plays an important role in CYP2C19–(S)-mephenytoin complex. Furthermore, for CYP2C19.1B, CYP2C19.8 and proteins **4** computational docking calculations gave the solutions of CScore = 5, and these proteins were experimentally active. These results suggest that our methods applied to this study are reasonable for binding mode predictions.

Only recently, the three-dimensional structures of CYP2C5 complex with sulfaphenazole (38), ligand-free CYP2C9 and CYP2C9 complexed with warfarin (39), were determined by X-ray crystallographic analyses. Because the sizes of the ligands in these complexes were different from (S)-mephenytoin, these structures were not able to give critical clues for the investigation of the complex structure of CYP2C19 with (S)-mephenytoin. Furthermore, for the observed structure of the CYP2C9 complex with warfarin, the ligand molecule was not located in the usual position of CYP substrate-binding pockets. Although this is useful for investigating the mechanisms of inhibitions and allosteric effects of CYPs, it does not aid research into the binding modes of normal CYP substrates, such as (S)-mephenytoin. Nonethe-

less, when these data are used together with the results of computational docking studies in this paper, it might give some support to the predictions of complex structures. Comparing the structures of CYP2C19.1A with (S)-mephenytoin in this study, and CYP2C5 with sulfaphenazole in crystallographic data, it was found that Asp290 of CYP2C5 (the residue corresponding to Asp293 in CYP2C19.1A) was in the substrate-binding pocket and interacted with the ligand sulfaphenazole. This indicates that both Asp290 of CYP2C5 and Asp293 of CYP2C19.1A play significant roles in binding of substrates, and supports the validity of our model structure of CYP2C19 complex with (S)-mephenytoin.

CONCLUSIONS

In this study, we found that models produced by homology modeling and computational docking procedures of CYP complexed with (S)-mephenytoin were obtained for experimentally active enzymes but not for inactive proteins. The binding modes of all active enzymes were similar to wild-type CYP2C19.1A. For inactive proteins, hydrogen bonds between the carbonyl oxygen in the main chain of Asp293 and the ligand were not able to form, or the scores of the resulting docking models were low. These results indicate that the computational methods used in this study are useful for the investigation of the structures of ligand–protein complexes. Although more detailed confirmations of the importance of Asp293 for (S)-mephenytoin binding into CYP2C19 by using structural biologic methods such as X-ray diffraction or NMR spectroscopy were preferable, the role of Asp293 was supported by the results that all calculations for wild types and mutants are consistent with experimentally observed enzyme activities without exception.

Three-dimensional structures of biopolymers are very important for the rational drug design. Though the determinations of structures by experimental methods (i.e., X-ray diffraction or NMR) give the useful information for drug design, these experiments are expensive and time-consuming. Therefore, it is preferable that the three-dimensional structures of structurally unknown proteins were computationally predicted using known structures of homologous proteins. Especially, when many complexes have to be considered for drug design (i.e., the situation that there are many mutants such as CYP2C19), homology modeling play an important role in drug design studies because all structures can not be determined by experiments within reasonable costs. In this study, even if the mutated parts of the mutants were far from the substrate-binding pockets, such as CYP2C19.6, the computational results were consistent with the experimental activities. This suggests that the computational methods used in this study, especially for FAMS, are promising not only for wild-type enzymes but also for mutants. We expect that virtual screening trials can be carried out for several proteins using these methods.

REFERENCES

1. M. Romkes, M. B. Faletto, J. A. Blaisdell, J. L. Raucy, and J. A. Goldstein. Cloning and expression of complementary DNAs for multiple members of the human cytochrome P450IIC subfamily. *Biochemistry* **30**:3247–3255 (1991).
2. S. A. Wrighton, J. C. Stevens, G. W. Becker, and M. VandenBranden. Isolation and characterization of human liver cytochrome P450 2C19: correlation between 2C19 and S-mephenytoin 4′-hydroxylation. *Arch. Biochem. Biophys.* **306**: 240–245 (1993).
3. J. A. Goldstein, M. B. Faletto, M. Romkes-Sparks, T. Sullivan, S. Kitareewan, J. L. Raucy, J. M. Lasker, and B. I. Ghanayem. Evidence that CYP2C19 is the major (S)-mephenytoin 4′-hydroxylase in humans. *Biochemistry* **33**:1743–1752 (1994).
4. A. Küpfer, P. Desmond, R. Patwardhan, S. Schenker, and R. A. Branch. Mephenytoin hydroxylation deficiency: kinetics after repeated doses. *Clin. Pharmacol. Ther.* **35**:33–39 (1984).
5. S. M. F. de Morais, G. R. Wilkinson, J. Blaisdell, K. Nakamura, U. A. Meyer, and J. A. Goldstein. The major genetic defect responsible for the polymorphism of S-mephenytoin in humans. *J. Biol. Chem.* **269**:15419–15422 (1994).
6. S. M. F. de Morais, G. R. Wilkinson, J. Blaisdell, K. Nakamura, U. A. Meyer, and J. A. Goldstein. Identification of a new genetic defect responsible for the polymorphism of S-mephenytoin metabolism in Japanese. *Mol. Pharmacol.* **46**:594–598 (1994).
7. Z. S. Xiao, J. A. Goldstein, H.-G. Xie, J. Blaisdell, W. Wang, C.-H. Jiang, F.-X. Yan, N. He, S.-L. Huang, Z.-H. Xu, and H.-H. Zhou. Differences in the incidence of the CYP2C19 polymorphism affecting the S-mephenytoin phenotype in Chinese Han and Bai populations and identification of a new rare *CYP2C19* mutant allele. *J. Pharmacol. Exp. Ther.* **281**:604–609 (1997).
8. R. J. Ferguson, S. M. F. de Morais, S. Benhamou, C. Bouchardy, J. Blaisdell, G. Ibeanu, G. R. Wilkinson, T. C. Sarich, J. M. Wright, P. Dayer, and J. A. Goldstein. A novel defect in human *CYP2C19*: mutation of the initiation codon is responsible for poor metabolism of S-mephenytoin. *J. Pharmacol. Exp. Ther.* **284**:356–361 (1998).
9. G. C. Ibeanu, J. A. Goldstein, U. Meyer, S. Benhamou, C. Bouchardy, P. Dayer, and B. I. Ghanayem. and J. Blaisdell. Identification of new human *CYP2C19* alleles (*CYP2C19*6* and *CYP2C19*2B*) in a Caucasian poor metabolizer of mephenytoin. *J. Pharmacol. Exp. Ther.* **286**:1490–1495 (1998).
10. G. C. Ibeanu, J. Blaisdell, B. I. Ghanayem, C. Beyeler, S. Benhamou, C. Bouchardy, G. R. Wilkinson, P. Dayer, A. K. Daly, and J. A. Goldstein. An additional defective allele, *CYP2C19*5*, contributes to the S-mephenytoin poor metabolizer phenotype in Caucasians. *Pharmacogenetics* **8**:129–135 (1998).
11. G. C. Ibeanu, J. Blaisdell, R. J. Gerguson, B. I. Ghanayem, K. Brosen, S. Benhamou, C. Bouchardy, G. R. Wilkinson, P. Dayer, and J. A. Goldstein. A novel transversion in the intron 5 donor splice junction of *CYP2C19* and a sequence polymorphism in exon 3 contribute to the poor metabolizer phenotype for the anticonvulsant drug S-mephenytoin. *J. Pharmacol. Exp. Ther.* **290**:635–640 (1999).
12. S. Kimura, J. Pastewka, H. V. Gelboin, and F. J. Gonzalez. cDNA and amino acid sequences of two members of the human P450IIC gene subfamily. *Nucleic Acids Res.* **15**:10053–10054 (1987).
13. C.-C. Tsao, M. R. Wester, B. Ghanayem, S. J. Coulter, B. Chanas, E. F. Johnson, and J. A. Goldstein. Identification of human CYP2C19 residues that confer S-mephenytoin 4′-hydroxylation activity to CYP2C9. *Biochemistry* **40**:1937–1944 (2001).
14. P. A. Williams, J. Cosme, V. Sridhar, E. F. Johnson, and D. E. McRee. Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity. *Mol. Cell* **5**:121–131 (2000).
15. R. D. Taylor, P. J. Jewsbury, and J. W. Essex. A review of protein-small molecule docking methods. *J. Comput. Aided Mol. Des.* **16**:151–166 (2002).
16. M. Rarey, B. Kramer, T. Lengauer, and G. Klebe. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **261**:470–489 (1996).
17. T. J. Ewing. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *J. Comput. Aided Mol. Des.* **15**:411–428 (2001).
18. G. Jones, P. Willet, R. C. Glen, A. R. Leach, and R. Taylor. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **267**:727–748 (1997).
19. D. F. V. Lewis, M. Dickins, R. J. Weaver, P. J. Eddershaw, P. S. Goldfarb, and M. H. Tarbit. Molecular modeling of human CYP2C subfamily enzymes CYP2C9 and CYP2C19: rationalization of substrate specificity and site-directed mutagenesis experiments in the CYP2C subfamily. *Xenobiotica* **28**:235–268 (1998).
20. V. A. Payne, Y.-T. Chang, and G. H. Loew. Homology modeling

and substrate binding study of human CYP2C18 and CYP2C19 enzymes. *Proteins* **37**:204–217 (1999).
21. M. Ridderström, I. Zamora, O. Fjellström, and T. B. Andersson. Analysis of selective regions in the active sites of human cytochromes P450, 2C8, 2C9, 2C18 and 2C19 homology models using GRID/CPCA. *J. Med. Chem.* **44**:4072–4081 (2001).
22. D. F. V. Lewis. Modeling human cytochromes P450 involved in drug metabolism from the CYP2C5 crystallographic template. *J. Inorg. Biochem.* **91**:502–514 (2002).
23. M. J. de Groot, A. A. Alex, and B. C. Jones. Development of a combined protein and pharmacophore model for cytochrome P450 2C9. *J. Med. Chem.* **45**:1983–1993 (2002).
24. K. Ogata and H. Umeyama. An automatic homology modeling method consisting of database searches and simulated annealing. *J. Mol. Graphics Mod.* **18**:258–272 (2000).
25. D. A. Case, D. A. Pearlman, J. W. Caldwell, T. E. Cheatham, W. S. Ross, C. Simmerling, Y. Duan, J. Pitera, I. Massova, G. L. Seibel, U. C. Singh, P. Weiner, and P. A. Kollman. *AMBER 6*, University of California, San Francisco, CA 1999.
26. U. C. Singh and P. A. Kollman. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* **5**:129–145 (1984).
27. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, and J. A. Pople. *Gaussian 98 Revision A.7*, Gaussian, Inc, Wallingford, CT 1998.
28. H. Tsujishita and S. Hirono. CAMDAS: an automated conformational analysis system using molecular dynamics. *J. Comput. Aided Mol. Des.* **11**:305–315 (1997).
29. T. A. Halgren. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **17**:490–519 (1996).
30. I. Muegge and Y. C. Martin. A general and fast scoring function for protein–ligand interactions: a simplified potential approach. *J. Med. Chem.* **42**:791–804 (1999).
31. I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. Langridge, and T. E. Ferrin. A geometric approach to macromolecule–ligand interactions. *J. Mol. Biol.* **161**:269–288 (1982).
32. M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini, and R. P. Mee. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.* **11**: 425–445 (1997).
33. P. S. Charifson, J. J. Corkery, M. A. Murcko, and W. P. Walters. Consensus scoring: a method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J. Med. Chem.* **42**:5100–5109 (1999).
34. R. D. Clark, A. Strizhev, J. M. Leonard, J. F. Blake, and J. B. Matthew. Consensus scoring for ligand/protein interactions. *J. Mol. Graphics Mod* **20**:281–295 (2002).
35. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**:926–935 (1983).
36. J. P. Ryckaert, G. Cicotti, and H. J. C. Berensden. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **23**:327–341 (1977).
37. O. Gotoh. 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 (1992).
38. M. R. Wester, E. F. Johnson, C. Marques-Soares, P. M. Dansette, D. Mansuy, and C. D. Stout. Structure of a substrate complex of mammalian cytochrome P450 2C5 at 2.3 Å resolution: evidence for multiple substrate binding modes. *Biochemistry* **42**:6370–6379 (2003).
39. P. A. Williams, J. Cosme, A. Ward, H. C. Angove, D. M. Vinkovi, and H. Jhoti. Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* **424**:464–468 (2003).