Identification of Human CYP2C19 Residues that Confer *S*-Mephenytoin 4′-Hydroxylation Activity to CYP2C9†

Cheng-Chung Tsao,‡ Michael R. Wester,§ Burhan Ghanayem,‡ Sherry J. Coulter,‡ Brian Chanas,‡ Eric F. Johnson,§ and Joyce A. Goldstein*,‡

Laboratory of Pharmacology and Chemistry, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709, and Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, California 92037

Received July 19, 2000; Revised Manuscript Received October 11, 2000

ABSTRACT: CYP2C19 is selective for the 4′-hydroxylation of *S*-mephenytoin while the highly similar CYP2C9 has little activity toward this substrate. To identify critical amino acids determining the specificity of human CYP2C19 for *S*-mephenytoin 4′-hydroxylation, we constructed chimeras by replacing portions of CYP2C9 containing various proposed substrate recognition sites (SRSs) with those of CYP2C19 and mutating individual residues by site-directed mutagenesis. Only a chimera containing regions encompassing SRSs 1−4 was active (30% of wild-type CYP2C19), indicating that multiple regions are necessary to confer specificity for *S*-mephenytoin. Mutagenesis studies identified six residues in three topological components of the proteins required to convert CYP2C9 to an *S*-mephenytoin 4′-hydroxylase (6% of the activity of wild-type CYP2C19). Of these, only the I99H difference located in SRS 1 between helices B and C reflects a change in a side chain that is predicted to be in the substrate-binding cavity formed above the heme prosthetic group. Two additional substitutions, S220P and P221T residing between helices F and G but not in close proximity to the substrate binding site together with five differences in the N-terminal portion of helix I conferred *S*-mephenytoin 4′-hydroxylation activity with a K_M similar to that of CYP2C19 but a 3-fold lower K_{cat}. Three residues in helix I, S286N, V292A, and F295L, were essential for *S*-mephenytoin 4′-hydroxylation activity. On the basis of the structure of the closely related enzyme CYP2C5, these residues are unlikely to directly contact the substrate during catalysis but are positioned to influence the packing of substrate binding site residues and likely substrate access channels in the enzyme.

The human CYP2C subfamily includes four structurally related enzymes (*1*). Among the human CYP2C isoforms, CYP2C9 and CYP2C19 are the most highly conserved, exhibiting 91% structural identity (*2*). Despite the fact that only 43 of 490 amino acids differ between CYP2C9 and CYP2C19, they have different substrate selectivities and/or different regio-selectivities for some common substrates. For example, CYP2C9 metabolizes acidic substrates such as diclofenac (*3*), ibuprofen (*4*), and 7-hydroxylates *S*-warfarin (*5*, *6*), but these pathways are not carried out by CYP2C19. Instead, CYP2C19 is the major enzyme responsible for the 4′-hydroxylation of *S*-mephenytoin (*1*, *6*, *7*), is highly selective for 5′-hydroxylation of omeprazole (*7*), and shows different stereo- and regio-specific preferences for warfarin hydroxylation (*5*, *6*). Tolbutamide is a common substrate but CYP2C9 exhibits a lower K_M than the other CYP2Cs for this substrate (*8*).

Many genetic polymorphisms that affect *S*-mephenytoin metabolism in humans have been characterized (*9*). Studies in our laboratory have identified nine genetic defects in CYP2C19 (*1*, *5*, *6*, *10*−*13*), which account for approximately 100% of Oriental and 95% of Caucasian poor metabolizers (*13*) of *S*-mephenytoin. Understanding the structural features of CYP2C19 that determine substrate specificity and efficient metabolism would be of benefit in the rational design of drugs whose clearance is less likely to be affected by the polymorphic forms of CYP2C19.

In the present study, we have identified critical amino acid residues that contribute to the marked specificity of human CYP2C19 for *S*-mephenytoin by constructing chimeras and mutant enzymes to confer activity to the closely related CYP2C9. These constructs were expressed in bacteria and assayed for their capacity to catalyze *S*-mephenytoin 4′-hydroxylation. Initial chimeras replaced regions of CYP2C9 with the corresponding segment of CYP2C19 so that discontinuous regions of the two proteins that form the substrate-binding site were systematically exchanged between CYP2C19 and CYP2C9. Site-directed mutagenesis was then applied to identify key residues within the critical regions identified in the chimeras that are responsible for the substrate specificity of CYP2C19. These studies identified six amino acid residues in three discontinuous regions of the protein that confer *S*-mephenytoin 4′-hydroxylation to CYP2C9.

† This study is supported in part by USPHS Grant GM31001 (E.F.J.)

* To whom correspondence should be addressed. Phone: (919) 541-4495. Fax: (919) 541-4107. E-mail: goldste1@niehs.nih.gov.

‡ National Institute of Environmental Health Sciences.

§ The Scripps Research Institute.

[1] Abbreviations: CYP, cytochrome P450 protein; DMSO, dimethyl sulfoxide; HPLC, high-performance liquid chromatograph; SRS, substrate recognition site.

10.1021/bi001678u CCC: $20.00
Published on Web 01/25/2001

Table 1: Oligonucleotides Employed in the Mutagenesis of CYP2C9[a]

codon change	sequence
I99H	5′-GAAGAGG**CA**TTTCCCACTG-3
S220P/P221T	5′-CAATAATTTT**C**CC**A**CTATCATTGATAC-3′
S286N	5′-TTCACTATTGAAA**A**CTTGGAAAACACT-3′
E288V	5′-ACTATTGAAAGCTTGG**T**AAACACTGCA-3′
N289I	5′-ATTGAAAGCTTGGAAA**T**CACTGCAGTT-3′
S286N/N289I	5′-TTCACTATTGAAA**A**CTTGGAAA**T**CACTGCAGTTGAC-3′
S286N/E288V/N289I	5′-TTCACTATTGAAA**A**CTTGG**T**AA**T**CACTGCAGTTGAC-3′
V292A	5′-ACTGCAG**C**TGACTTGTTTGGA
F295L	5′-GTTGACTTG**C**TTGGAGCTGGG-3′

[a] Nucleotides in bold indicate mismatched bases in CYP2C19. Codons for the changed amino acids are underlined.

EXPERIMENTAL PROCEDURES

Chemicals and Reagents. Standards of *S*-mephenytoin, 4′-hydroxymephenytoin, and radiolabeled *S*-[4-^{14}C]mephenytoin were purchased from Amersham Corp. (Arlington Heights, IL). Recombinant human NADPH−P450 oxidoreductase and human cytochrome b_5 were purchased from Oxford Biomedical Research Inc. (Oxford, MI) and 1,2-didodecanoyl-sn-glycero-3-phosphocholine was purchased from Sigma Chemical Co. (St. Louis, MO).

Site-Directed Mutagenesis and Construction of Chimeras. CYP2C9 and CYP2C19 were previously modified for expression in bacteria (*13*). The chimeras were constructed using the *Sma*I, *Sph*I, *Pst*I, *Sty*I, and *Nde*I sites that are shared by both cDNAs. An *Eco*RI site present in the CYP2C19 cDNA was introduced into CYP2C9 by site-directed mutagenesis as previously described (*14*). Site-directed amino acid changes were performed using a Chameleon double strand, site-directed mutagenesis kit (Stratagene, La Jolla, CA) with synthetic oligonucleotides containing the desired point mutations and a second primer (Table 1), which abolished a unique restriction site in the pCW expression plasmid. All of the constructs were confirmed by automated sequencing (Applied Biosystems Inc., Foster City, CA) prior to protein expression in DH5α cells transformed with the recombinant pCW clones. The P450-rich bacterial membranes were isolated and partially purified using the methods described by Richardson et al. (*15*) and the P450 contents were determined by dithionite-reduced carbon monoxide difference spectra.

S-Mephenytoin Hydroxylation Assays. Assays for the 4′-hydroxylation of *S*-mephenytoin were performed with partially purified recombinant proteins using previously established procedures (*1*). Briefly, purified recombinant proteins (10−25 pmol) were incubated with 0.3 μg/*p*mol 1,2-didodecanoyl-*sn*-glycero-3-phosphocholine, P450 reductase (4 *p*mol/*p*mol P450), and human cytochrome b_5 (2 *p*mol/*p*mol P450). The reconstituted mixture was preincubated for 5 min at 37 °C and then placed on ice. A final concentration of 0.4 mM (32 μCi/μmol) radiolabeled *S*-mephenytoin was added to assay buffer consisting of 20 mM HEPES (pH 7.4), 0.1 mM EDTA, and 1.25 mM $MgCl_2$ in the final volume of 125 μL reaction mixture. The mixture was then incubated at 37 °C with shaking for 3 min, and the reaction was started by adding 50 mM NADPH (final concentration 2 mM) and terminated after 15 min with an equal volume of methanol. The incubation mixture was centrifuged at 10000*g* for 10 min and an aliquot of the supernatant was assayed directly by HPLC without extraction. The HPLC system consisted of a reversed-phase C18 Ranin Microsorb mv, 250 mm × 4.6 mm column (Varian, Walnut Creek, CA), using an isocratic solvent system consisting of methanol/water (45:55) at a flow rate of 1 mL/min for 25 min. Detection of radioactive peaks was accomplished using an in-line Flow-One radiochemical detector (Packard Instruments Co., Tampa, FL).

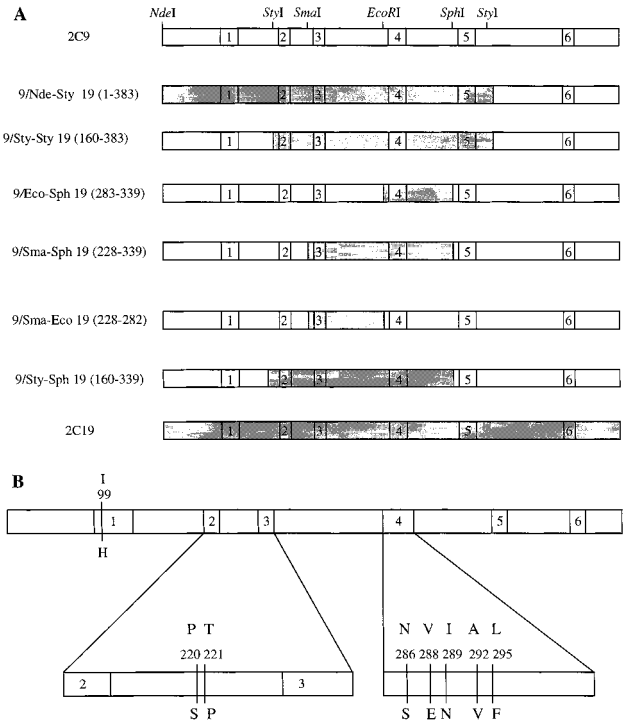

FIGURE 1: Schematic representation of the chimeric constructs (A) and the site-directed mutants of CYP2C9 and CYP2C19 (B). (A) The restriction enzyme sites used to construct chimeras and the putative SRSs (numbered 1−6) are indicated at the top. The white bars represent the segments of CYP2C9 sequence and the shaded bars represent sequence derived from CYP2C19. (B) The relative locations and residue number of site-directed mutations of CYP2C9 are depicted. The amino acids shown above the construct represent residues in CYP2C9 that were mutated to the corresponding residues of CYP2C19 shown below the constructs.

RESULTS

To determine which of the 43 amino acid differences between CYP2C9 and CYP2C19 contribute to the difference in *S*-mephenytoin metabolism, a series of chimeras was generated to localize specific portions of CYP2C19 that could confer specificity for the substrate (Figure 1). The chimeras were derived from the two enzymes by inserting putative

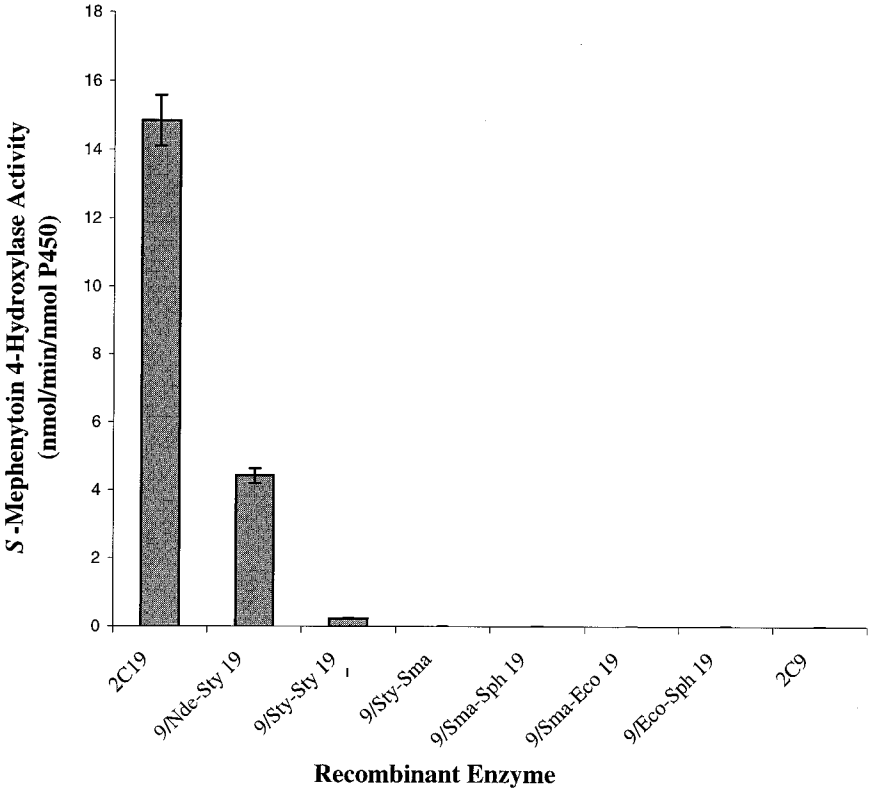

FIGURE 2: *S*-Mephenytoin 4′-hydroxylation activity (nmol/min/nmol P450) of partially purified chimeras of CYP2C9 and CYP2C19. Values represent the mean ± SE ($n = 3$).

substrate recognition sites (SRSs) proposed by Gotoh (*16*) from CYP2C19 into CYP2C9. These segments encode discontinuous portions of P450 structures that form the substrate binding site (*16*). These constructs replaced either residues 1−383 (containing SRS 1, 2, 3, 4 and 5) of CYP2C9 with those of CYP2C19 (9/*Nde-Sty* 19), residues 160−383 (containing SRS 2, 3, 4, and 5) of CYP2C19 (9/*Sty-Sty* 19), residues 283−339 (containing SRS 4) of CYP2C19 (9/*Eco-Sph* 19), residues 228−339 (containing SRS 3 and 4) of CYP2C19 (9/*Sma-Sph* 19), residues 228−282 (SRS 3) of CYP2C19 (9/*Sma-Eco* 19) or residues 160−339 (containing SRS 2, 3, and 4) of CYP2C19 (9/*Sty-Sph*).

In contrast to CYP2C19, CYP2C9 does not exhibit *S*-mephenytoin 4′-hydroxylation activity. Like CYP2C9, constructs 9/*Sty-Sph* 19, 9/*Sma-Sph* 19, 9/*Sma-Eco* 19, and 9/*Eco-Sph* 19 did not display *S*-mephenytoin 4′-hydroxylation activity. A CYP2C9 chimera containing the *Sty*I−*Sma*I fragment of CYP2C19 (residues 160−227 containing SRS 2) was also inactive (data not shown). An earlier study in a yeast cDNA expression system established that the region from 338 to the C-terminus was also inactive (*12*). In the present study, only the 9/*Nde-Sty* chimera, containing SRSs 1−5 of CYP2C19, conferred *S*-mephenytoin 4′-hydroxylation activity equivalent to approximately 30% of the activity of CYP2C19 (Figure 2). These results indicate residues in each of the 4 regions (0−160, 160−227, and 228−339) are necessary for conferring minimal *S*-mephenytoin 4′-hydroxylation activity to CYP2C9.

Analyzing these regions sequentially, at least one residue located in the first 160 N-terminal residues is obviously necessary to confer minimal *S*-mephenytoin activity to CYP2C9 in combination with other residues from 160 to 339. Of these, only I99H is located in close proximity to the substrate-binding site based on the structure of CYP2C5 (*17*). In a previous study, we demonstrated that adding I99H alone to CYP2C9 conferred a dramatic increase in omeprazole 5′-hydroxylation activity but surprisingly did not confer *S*-mephenytion hydroxylation (*7*). In contrast, the reverse mutation in CYP2C19 results in a significant decrease in the catalytic efficiency of CYP2C19 for *S*-mephenytoin 4′-hydroxylation and is found in an allelic variant of CYP2C19 (E.F.J., unpublished data). In this study, we found that the substitution I99H could confer *S*-mephenytoin hydroxylation activity to a chimera containing additional residues derived from CYP2C19 between the *Sty*I−*Sph*I restrictions sites (SRSs 2−4) (Figure 3). In contrast, the I99H substitution did not confer detectable *S*-mephenytoin 4′-hydroxylation activity to chimeras containing only the *Sty*I-*Sma*I (160−227) or *Sma*I-*Sph*I (residues 228−339) fragments of CYP2C19, indicating that one or more differences in each of these regions are also required for activity. Our previous study on the differences between the two proteins that confer omeprazole hydroxylation to CYP2C9 indicated that the mutations S220P and P221T in the *Sty*I−*Sma*I segment as well as I99H conferred 100% of the omeprazole 5′-hydroxylation activity of CYP2C19 to CYP2C9 in a yeast cDNA expression system (*7*). We therefore examined the effects of mutating residues 220 and 221 of CYP2C9 to correspond to those in CYP2C19. The addition of I99H, S220P, and P221T to the chimeras 9/*Sma-Sph* 19 (containing SRSs 3 and 4 of CYP2C19) and 9/*Eco-Sph* 19 (containing only SRS 4 of CYP2C19) resulted in detectable *S*-mephenytoin 4′-hydroxylation activity corresponding to 30 and 17%, respectively, of that of CYP2C19 (Figure 2). The addition of I99H, S220P, and P221T to a 9/*Sma*−*Eco* construct (containing SRS 3 of CYP2C19) did not confer detectable

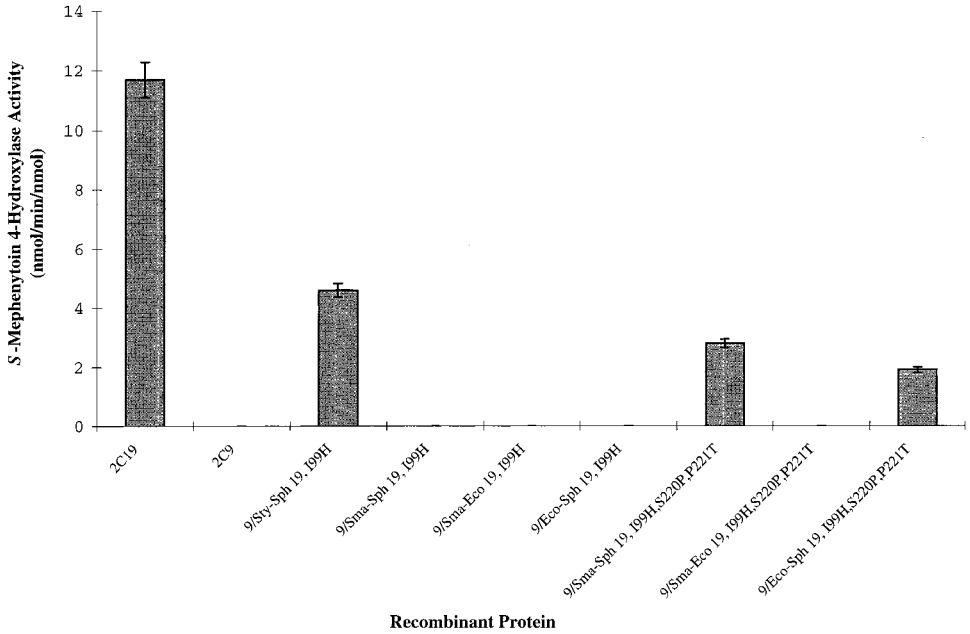

FIGURE 3: *S*-Mephenytoin 4′-hydroxylation activity (nmol/min/nmol P450) of selected site-directed mutants of the chimeras listed in Figure 1A. Values represent the mean ± SE ($n = 3$).

Table 2: *S*-Mephenytoin 4-Hydroxylation Activity of Partially Purified CYP2C9 Mutants

construct	*S*-mephenytoin 4′-hydroxylation activity (nmol/min/nmol)[a]	% activity of CYP2C19
CYP2C19	5.0 ± 0.4	100
CYP2C9	ND[b]	
9/*Eco-Sph*, I99H, S220P, P221T	0.9 ± 0.2	17
9/I99H, S220P, P221T, S286N	ND	
9/I99H, S220P, P221T, E288V	ND	
9/I99H, S220P, P221T, N289I	ND	
9/I99H, S220P, P221T, S286N, N289I	ND	
9/I99H, S220P, P221T, S286N, E288V, N289I	ND	
9/I99H, S220P, P221T, V292A, F295L	ND	
9/I99H, S220P, P221T, S286N, V292A, F295L	0.3 ± 0.1	6.3
9/I99H, S220P, P221T, E288V, V292A, F295L	ND	
9/I99H, S220P, P221T, N289I, V292A, F295L	ND	
9/I99H, S220P, P221T, S286N, N289I, V292A, F295L	0.4 ± 0.1	8.3
9/I99H, S220P, P221T, S286N, V292A	NE[c]	ND
9/I99H, S220P, P221T, S286N, F295L	ND	
9/I99H, S220P, P221T, S286N, N289I, V292A	ND	ND
9/I99H, S220P, P221T, S286N, N289I, F295L	ND	

[a] Values represent means ± SE ($n = 3$) [b] None detectable, due to the radioactivity levels are less than the level of detection (<0.13 nmol/min/nmol P450). [c] NE, no P450 protein expression.

S-mephenytoin 4′-hydroxylation activity, indicating that CYP2C19 residues in this region are not essential for activity. Taken together, these results indicate that amino acid differences residing in three topologically distinct regions of the protein are the minimal critical determinants for the distinct catalytic specificity of CYP2C19 for *S*-mephenytoin 4′-hydroxylation and that amino acids in other regions further increase the turnover number for mephenytoin.

A comparison of the amino acid sequences encoded by the *Eco*RI−*Sph*I region indicated that CYP2C19 differs from CYP2C9 at five residues that would occur in helix I: 286, 288, 289, 292, and 295 based on the structure of CYP2C5 (*17*). We therefore sequentially mutated various combinations of amino acids in this region to those of the corresponding residues in CYP2C19 in a construct consisting of CYP2C9 that contained the mutations I99H, S220P, and P221T. None of the single substitutions in the I helix were able to confer *S*-mephenytoin 4′-hydroxylation activity, suggesting that more than one residue in this region is necessary for this activity. To determine which combinations of residues were required for *S*-mephenytoin activity, the additional mutants summarized in Table 2 were constructed. The double mutants 286/289 and 292/295 or triple mutants, 286/288/289, 288/292/295, or 289/292/295, were inactive. However, the triple mutant 286/292/295 in combination with the corresponding CYP2C19 residues at positions 99, 220, and 221, resulted in the conversion of CYP2C9 to an *S*-mephenytion 4′-hydroxylase exhibiting 6% of the activity of CYP2C19. These results suggest that these six residues provide a minimal framework for endowing CYP2C9 with *S*-mephenytoin 4′-hydroxylation activity (Table 2). The addition of the remaining two residues in SRS 4 (288 and 289) further enhanced this activity to 17% of that of wild-type CYP2C19.

Detailed characterization of the kinetic properties of the CYP2C9 mutants (9/*Sma-Sph*, I99H, S220P, and P221T, and 9/*Eco-Sph*, I99H, S220P, P221T) revealed that the mutants

Table 3: Kinetic Characteristics for *S*-Mephenytoin Metabolized by CYP2C19, CYP2C9 Chimeric Proteins 9/*Sma-Sph*, I99H, S220P, P221T, and 9/*Eco-Sph*, I99H, S220P, P221T[a]

protein	K_{cat} (nmol/min/nmol P450)	K_M (μM)	K_{cat}/K_M
CYP2C19	13.26 ± 2.65	49.77 ± 3.96	0.27
9/*Sma-Sph*, I99H, S220P, P221T	6.86 ± 0.2 (52%)	81.70 ± 5.03	0.08 (30%)
9/*Eco-Sph*, I99H, S220P, P221T	1.36 ± 0.71 (10%)	46.20 ± 3.20	0.03 (11%)

[a] K_{cat} and K_M values were estimated by nonlinear regression and analysis of the velocity versus substrate concentration plots (concentration range: 3−500 μM) and are indicated in the table ± standard error ($n = 3$). Percentage values relative to CYP2C19 are shown in parentheses.

had K_M values similar to wild-type CYP2C19, but their K_{cat} values were 11 and 50% of those of wild-type CYP2C19, respectively (Table 3). Their catalytic efficiencies (K_{cat}/K_M) were 30 and 11% of that of CYP2C19.

DISCUSSION

Chimeras between CYP2C9 and the closely related CYP2C19 were systematically constructed in order to identify the minimal amino acid differences that are required for CYP2C19 selectivity for *S*-mephenytoin 4′-hydroxylation. A combination of six differences, I99H, S220P, P221T, S286N, V292A, and F295L were required to confer *S*-mephenytoin 4′-hydroxylation activity to CYP2C9 indicating that multiple differences between CYP2C9 and CYP2C19 contribute to the unique substrate specificity of CYP2C19. The H99I and S220P substitutions were shown in our previous work to confer omeprazole 5′-hydroxylation activity to CYP2C9 (*7*). Furthermore, when a second nonconservative mutation was made at residue 221 (P221T) in conjunction with S220P and I99H, the omeprazole 5′-hydroxylation activity of this triple mutant was further enhanced and almost equivalent to that of CYP2C19. In both the previous and present studies, this triple mutant did not exhibit *S*-mephenytoin 4′-hydroxylation activity. However, in the present study, when these mutations were added to a construct containing five additional substitutions from helix I of CYP2C19, *S*-mephenytoin 4′-hydroxylation activity (17% of CYP2C19) was conferred to CYP2C9. A construct containing only three of the five divergent residues in helix I, residues 99, 220, 221, 286, 292, and 295 of 2C19 displayed 6% of the wild-type CYP2C19 *S*-mephenytoin 4′-hydroxylation activity. Kinetic studies showed that these mutants exhibited K_Ms similar to that of wild-type CYP2C19 but they exhibit lower K_{cat}s. Their intrinsic clearances were 30 and 11% of that of wild-type CYP2C19. This is consistent with our modeling studies suggesting that these residues are not part of the binding sites for the substrate but that they influence substrate access or binding indirectly. The further 3-fold increase in activity conferred by the inclusion of CYP2C19 residues 288 and 289 suggests that these residues affect the microenvironment of the binding site. Attempts to express the 9/I99H, S220P, P221T, S286N, V292A mutant, were unsuccessful, precluding assessment of the contribution of residue 295 to *S*-mephenytoin 4′-hydroxylation activity. Interestingly, two other studies (*3*, *18*) have demonstrated significantly increased activities toward CYP2C9 substrates such as warfarin and diclofenac when mutating residue 286 of CYP2C19 to that of CYP2C9 (N286S). This mutation as well as the I289N substitution conferred high-affinity binding of the selective inhibitor of CYP2C9, sulfaphenazole, to CYP2C19 (*18*). In the current study, we also found that residue 286 appears to be necessary for the *S*-mephenytoin substrate specificity of CYP2C19, although other residues are also required. Taken together these results suggest an important role for residue 286 in determining substrate specificity for CYP2Cs.

The recently obtained structure of a mammalian P450 in the same subfamily, CYP2C5 (*17*), provides a basis for localizing the residues identified by mutagenesis in the three-dimensional structure of the enzyme and allows interpretation of the likely mechanisms that lead to differences in substrate specificity. The amino acid sequence identity between CYP2C9 and CYP2C19 and CYP2C5 is >75% at the amino acid level, and the sequence alignment of the three enzymes is straightforward. On the basis of the structure of CYP2C5 (*17*), homology models of CYP2C9 and CYP2C19 were constructed using MODELLER (*19*). The program builds the models based on the backbone and side-chain orientations found in the template structure provided by CYP2C5. Energy minimization and simulated annealing by molecular dynamics are then used to optimize the side-chain packing and stereochemistry balanced by restraints to conform to the template structure. The program AUTODOCK was then used to place a molecule of *S*-mephenytoin into the active site of the model of CYP2C19. A single low-energy orientation and position of the substrate was found for 50 runs after randomly seeding the substrate in the model of CYP2C19. The resulting orientation positions the 4′-hydrogen of *S*-mephenytoin approximately 5.8 Å from the heme iron. This value is similar to estimates for hydroxylation sites obtained for other CYP2C9 substrates based on NMR relaxation studies (*20*).

The position of the substrate resulting from the automated docking is depicted in Figure 4A together with the locations in the model of the six critical residues of CYP2C9 and CYP2C19 that were required to confer the highest *S*-mephenytoin 4′-hydroxylation activity to CYP2C9. As can be seen none of the critical residues appear to be in direct contact with the substrate, and they generally lie outside of the primary shell of contact residues that form the active-site cavity. In fact, with two exceptions, residues 208 and 362, all of the side chains that form the primary shell of the substrate-binding cavity above the heme are identical between CYP2C19 and CYP2C9. The two divergent residues have not been implicated as determinants of substrate specificity in this or previous studies. Although two of the modeled residues identified in this work lie close to the substrate-binding cavity, the side chain of residue 295 is oriented away from the active site in CYP2C5. A Phe is found at this position in CYP2C5 as well as in CYP2C9, whereas a Leu is found in CYP2C19.

Another amino acid difference, I99H that confers *S*-mephenytoin 4′-hydroxylation activity to CYP2C9 resides close to the substrate-binding site. This residue maps to the

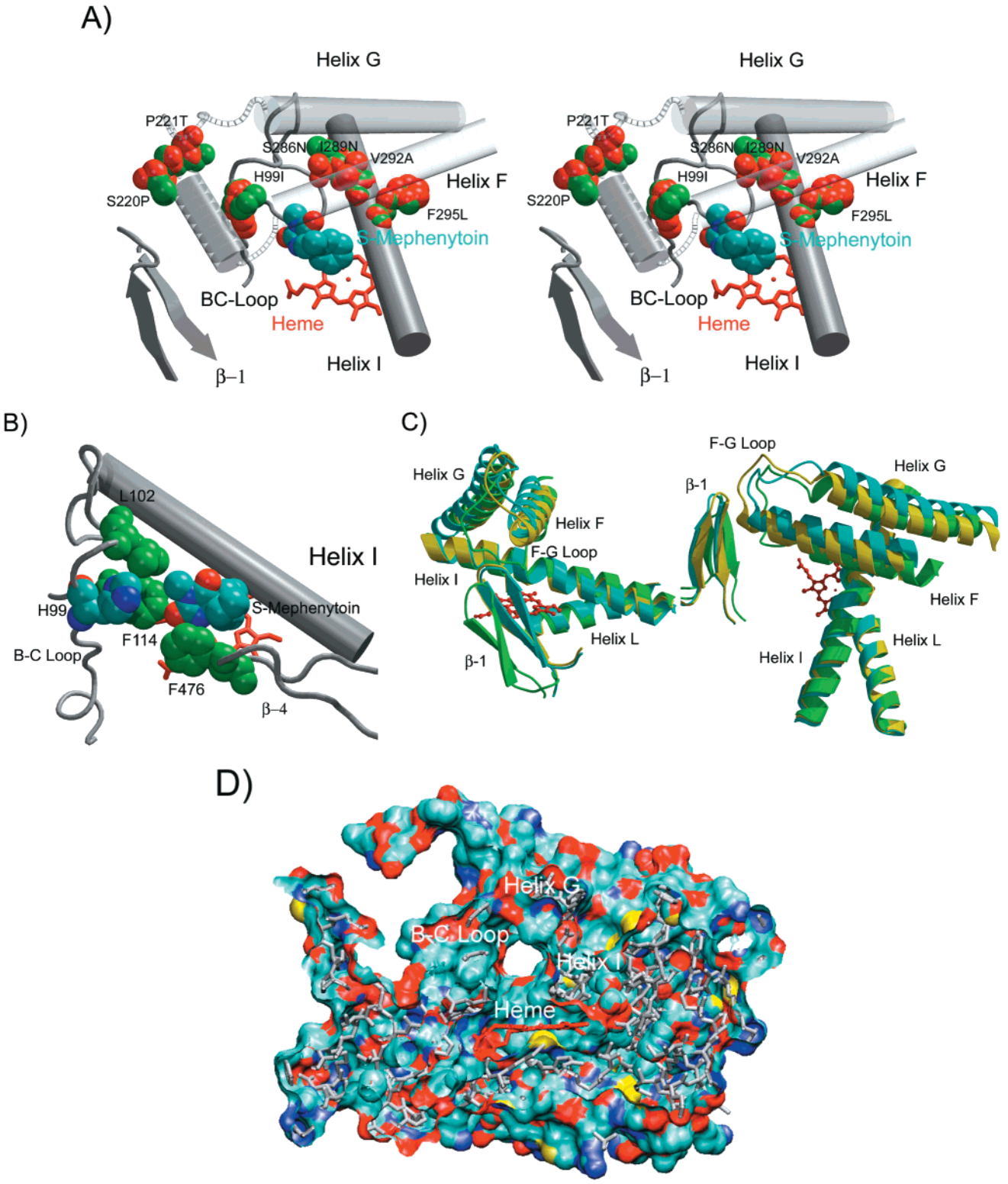

FIGURE 4: Localization of residues that confer *S*-mephenytoin 4′-hydroxylation activity in the CYP2C19 modeled on the structure of CYP2C5. (A) Several topological features of a model for CYP2C19 based on the structure of CYP2C5 are shown in a stereoview. The position of *S*-mephenytoin predicted by AUTODOCK is depicted in a space filling model. The critical amino acid residues are shown in green for CYP2C19 and in red for CYP2C9. For this model, the F−G loop was modeled as a helix as described in the text. (B) A view of H99 in the model of CYP2C19 showing the placement of neighboring residues that block direct interactions with the substrate. (C) A comparison of the locations of helices F, G, and I and β-sheet 1 for the substrate-free (cyan) structure of P450 BM3, the substrate-bound form of BM3 (yellow), and the structure of CYP2C5 (green). The longer F−G loop in CYP2C5 passes through the substrate access channel seen for P450 BM3. (D) A view of an alternative substrate access channel in CYP2C5. A solvent accessible surface rendering of the structure of CYP2C5 without the F−G loop. The portion of the molecule and surface in front of the heme iron is removed to reveal a channel passing from the substrate-binding site to the back surface of the protein.

region between the B and C helices of CYP2C5 that is often referred to as SRS 1. The side chain of H99 of CYP2C19 has the capacity for hydrogen bonding that would allow interaction with both omeprazole and *S*-mephenytoin, whereas

I99 of CYP2C9 lacks this capability (*7*). The corresponding residue S99 of CYP2C5 is oriented into the substrate-binding cavity above the heme. If H99 of CYP2C19 and isoleucine 99 of CYP2C9 adopt an orientation similar to that seen for S99 of CYP2C5, their side chains would extend into the active site (Figure 4A) and both would contribute to the active-site surface. However, the CYP2C19 model suggests that adjacent residues would shield H99 from direct interaction with the substrate as shown in Figure 4B. In this model, other residues that pack against H99, such as L102, F114, and F476 (Figure 4B), are closer to *S*-mephenytoin than H99 and are likely to prevent a closer approach of the substrate to H99. In the event that a closer approach were possible, the substrate would be more distant from the heme, which would reduce the likelihood for oxidation. Also, the orientation of the side chain and the substrate would not readily permit hydrogen bonding. Other generally allowed rotamers for H99 that could yield alternative models would place the side chain of H99 at a greater distance from the substrate but could allow productive hydrogen bonding with the backbone of the protein. Such interactions could potentially alter the packing of the SRS-1 region that lies in the loop between helices B and C. As the other residue differences that help to confer *S*-mepehenytoin activity are more distant from the substrate-binding site, it is likely that the distinct substrate specificities of the two enzymes arise from indirect affects modifying the packing of the side chains that form the primary shell or that alter substrate access to the active site. This hypothesis is consistent with kinetic studies showing that the K_M's of the active CYP2C9 mutants for *S*-mephenytoin 4′-hydroxylation were similar to those of CYP2C19, but the k_{cat}'s of the mutants were lower than that of CYP2C19.

The location of the residues that confer *S*-mephenytoin 4′-hydroxylation activity is consistent with changes that alter substrate access. The residues that confer *S*-mephenytoin 4′-hydroxylation activity to CYP2C9 are found in regions that exhibit differences in their positions and/or interactions with other residues when the structures of the substrate bound and substrate free forms of P450 BM3 are compared. In contrast to most CYPs that have been crystallized with substrates present, P450 BM3 has been crystallized without a bound substrate, and it exhibits a clear channel from the substrate-binding cavity above the heme to the surface that passes between the F−G loop and the top of the loop preceding Helix A and the top β-strand 1−2 (*21*). This passage is closed when the protein is crystallized with the substrate analogue palmitoleic acid present in the active site (*22*). The closure largely reflects the movement of the F and G helices and the connecting loop between them, Figure 4C, coupled with additional changes that extend through helix H to the N-terminal end of helix I as well as regions that contact this portion of helix I. This region is contained in the *Sty*I−*Sph*I fragment used in this study. The results of this study indicate that while several of these residues are necessary for activity, additional differences in this region also contribute to optimal activity.

One of the most striking changes that occurs in BM3 upon binding substrate occurs in the N-terminal end of the F helix that translates across the surface of the helix I and pivots, Figure 4C. The place where the F helix crosses the I helix in CYP2C5 corresponds closely to that seen for the substrate bound form of BM3, Figure 4C, and this may reflect the presence of a foreign compound in the active site of CYP2C5 when it was crystallized. This position occurs precisely where two of the differences, V292A and F295L, are located in helix I of CYP2C9 and CYP2C19. It is potentially significant that a large volume difference is evident for the two residues at position 292, and although the difference is not as large as that at 295, the aromatic Phe in CYP2C9 is bulkier than the Leu in CYP2C19. In addition, the Phe is more highly constrained with respect to favorable rotamer conformations. If similar changes occur during substrate binding to the mammalian P450s, it seems likely that differences that occur at various sites of interaction between secondary structural elements could influence the changes that occur, the extent of conformational changes, and the orientation of the axis of the F-helix. As can be seen in Figure 4C, the orientation and length of the F-helix differs from that of the substrate bound form of BM3 and differences in the packing with the I helix are likely to contribute to this difference.

The potential substrate access channel seen in BM3 is closed in CYP2C5. This reflects, in part, the positioning of the F-helix as described in the previous paragraph as well as the presence of a significant insertion of amino acids in the region between the F and G helices. Although the structure of this loop could not be accurately modeled in the structure of CYP2C5, the location of electron density and crystal packing constraints indicate that it fills the region where the substrate access channel occurs in BM3. The location of the loop was modeled as a helix in Figure 4A and is shown as a coil in Figure 4B. By modeling this region as a helix in the model of CYP2C19 shown in Figure 4A, the two additional critical differences at residues 220 and 221 are placed in closest proximity to the substrate binding site. Although they do not contact the substrate, they do occur at the contact between the F−G and B−C loops close to residue 99 where the I99H difference occurs. This loop closes the substrate access channel seen in the substrate-free structure of BM3. The poor resolution of this region in the crystal structure for CYP2C5 suggests that it may be conformationally labile. Halpert and co-workers (*23*) also noted that residue 221 was one of seven residues in CYP2B4 which changed the enzymatic phenotype of CYP2B4 to that of CYP2B5, even though modeling studies suggested that this residue was located far from the active site.

An examination of the structure of CYP2C5 suggests that an alternative route for substrate access seems likely. As shown in Figure 4D, an open channel is seen between the substrate binding site and the exterior of the protein. This channel could allow substrate access and egress between helix I, the B−C loop and helix G. The B−C loop exhibits relatively high B-values in the structure of CYP2C5, and it may be sufficiently flexible to open this channel further for the passage of substrates and products to or from the active site. The additional residues identified in this study reside along helix I where helix G passes over it and along the base of this channel where it contacts the B−C loop, Figure 4A. In addition, a number of studies have shown that mutations in the C-terminal portion of this segment where it exits from the substrate-binding site and passes along helix I alter the substrate specificities of members of the CYP2A, CYP2B, and CYP2C subfamilies (*24*−*29*). Also, the substitutions at residues 286 and 289 could potentially alter

hydrogen-bonding networks that may occur between helix I and the B−C loop. This scenario suggests that the critical differences defined in this study influence substrate specificity indirectly by altering the conformational changes that occur during substrate access and/or binding. In this regard, it is interesting to note that several mutants of P450 cam that confer altered rates and regiospecificities for hydrogen peroxide driven hydroxylation of naphthalene reflect amino acid substitutions at sites distal from the substrate binding site of the enzyme (*30*).

REFERENCES

1. de Morais, S. M. F., Wilkinson, G. R., Blasisdell, J., Nakamura, K., Meyer, U. A., and Goldstein, J. A. (1994) *J. Biol. Chem. 269*, 15419−15422.
2. Romkes, M., Faletto, M. B., Blaisdell, J. A., Raucy, J. L., and Goldstein, J. A. (1991) *Biochemistry 30*, 3247−3235.
3. Klose, T. S., Ibeanu, G. C., Ghanayem, B. I., Pedersen, L. G., Li, L., Hall, S., and Goldstein, J. A. (1998) *Arch. Biochem. Biophys. 357*, 240−248.
4. Hamman, M. A., Thompson, G. A., and Hall, S. D. (1997) *Biochem. Pharmacol. 54*, 33−41.
5. Goldstein, J. A., and de Morais, S. M. F. (1994) *Pharmacogenetics 4*, 285−299.
6. de Morais, S. M. F., Wilkinson, G. R., Blaisdell, J., Nakamura, K., Meyer, U. A., and Goldstein, J. A. (1994) *Mol. Pharmacol. 46*, 954−958.
7. Ibeanu, G. C., Ghanayem, B. I., Linko, P., Li, L., Pedersen, L. G., and Goldstein, J. A. (1996) *J. Biol. Chem. 271*, 12496−12501.
8. Wester, M. R., Lasker, J. M., Johnson, E. F., Raucy, J. L. (2000) *Drug Metab. Dispos. 28*, 354−359.
9. Wilkinson, G. R., Guengerich, F. P., and Branch, R. A. (1989) *Pharmacol. Ther. 43*, 53−76.
10. Ferguson, R. J., de Morais, S. M. F., Benhamou, S., Bouchardy, C., Blasisdell, J., Ibeanu, G. C., and Goldstein, J. A. (1998). *J. Pharmacol. Exp. Ther. 284*, 356−361.
11. Ibeanu, G. C., Blaisdell, J., Ghanayem, B. I., Beyeler, C., Benhamou, S., Bouchardy, C., Wilkinson, G. R., Dayer, P., Daly, A. K., and Goldstein, J. A. (1998) *Pharmacogenetics 8*, 129−135.
12. Ibeanu, G. C., Goldstein, J. A., Meyer, U., Benhamou, S., Bouchardy, C., Dayer, P., Ghanayem, B. I., and Blaisdell, J. (1998) *J. Pharmacol. Exp. Ther. 286*, 1490−1495.
13. Ibeanu, G. C., Blaisdell, J., Ferguson, R. J., Ghanayem, B. I., Brøsen, K., Benhamou, S., Bouchardy, C., Wilkinson, G. R., Dayer, P., and Goldstein, J. A. (1999) *J. Pharmacol. Exp. Ther. 290*, 635−640.
14. Deng, W. P., and Nickloff, J. A. (1992) *Anal. Biochem. 200*, 81−88.
15. Richardson, T. H., Jung, F., Griffin, K. J., Wester, M., Raucy, J. L., Kemper, B., Bornheim, L. M., Hassett, C., Omiecinski, C. J., and Johnson, E. F. (1995) *Arch. Biochem. Biophys. 323*, 87−96.
16. Gotoh, O. (1992) *J. Biol. Chem. 267*, 83−90.
17. Williams, P. A., Cosme, J., Sridhar, V., Johnson, E. F., and McRee, D. E. (2000) *Mol. Cell 5*, 121−131.
18. Jung, F., Griffin, K. J., Song, W., Richardson, T. H., Yang, M., and Johnson, E. F. (1998) *Biochemistry 37*, 16270−16279.
19. Sali, A., Potterton, L., Yuan, F., van Vlijmen, H., and Karplus, M. (1995) *Proteins 23*, 318−326.
20. Poli-Scaife, S., Attias, R., Dansette, P. M., and Mansuy, D. (1997) *Biochemistry 36*, 12672−12682.
21. Hasemann, C. A., Kurumbali, R. G., Boddupalli, S. S., Peterson, J. A., and Diesenhofer, J. (1995) *Structure 3*, 41−62.
22. Li, H. Y., and Poulos, T. L. (1997) *Nat. Struct. Biol. 4*, 140−146.
23. He, Y. Q., Harlow, G. R., Szklarz, G. D., and Halpert, J. R. (1998) *Arch. Biochem. Biophys. 350*, 333−339.
24. Lindberg, R., and Negishi, M. (1989) *Nature 339*, 632−634.
25. Lin, D., Black, S. M., Nagahama, Y., and Miller, W. M. (1993) *Endocrinology 132*, 2498−2506.
26. Kronbach, T., Larabee, T. M., and Johnson, E. F. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8262−8265.
27. Kronbach, T., Kemper, B., and Johnson, E. F. (1991) *Biochemistry 30*, 6097−6102.
28. Uno, T., and Imai, Y. (1992) *J. Biochem. 112*, 155−162.
29. Johnson, E. F., Kronbach, T., and Hsu, M.-H. (1992) *FASEB J. 6*, 700−705.
30. Joo, H., Lin, Z., and Arnold, F. H. (1999) *Nature 339*, 670−673.

BI001678U