Efficient Assembly of Functional Cytochrome P450 2C2 Requires a Spacer Sequence between the N-terminal Signal Anchor and Catalytic Domains

(Received for publication, May 7, 1997, and in revised form, June 26, 1997)

Ci-Di Chen‡, Balraj Doray§, and Byron Kemper¶
From the Departments of ‡Molecular and Integrative Physiology and §Cell and Structural Biology, University of Illinois College of Medicine, Urbana, Illinois 61801

Cytochromes P450 (P450) are anchored to the endoplasmic reticulum membrane by an N-terminal transmembrane sequence with the catalytic domain facing the cytoplasmic side. Within the peptide sequence linking these two domains in P450 2C2 is a glycine-rich region from residues 22 to 28. To examine the role of this region, deletion and substitution mutations were constructed, and the activities and spectral properties were determined for the mutant proteins expressed in COS-1 cells, insect cells, and bacteria. Deletion of residues 22 to 28 or substitution of 7 valines for this region inactivated the proteins in COS-1 cells, and no P450 species was detected for these mutations in bacteria or insect cells. Substitution of the three glycine residues with alanine or proline or the entire sequence from 22 to 28 with 7 alanines did not reduce lauric acid hydroxylase activity of the proteins expressed in COS-1 cells. Reducing the number of alanines substituted to 4, 3, and 2 progressively decreased activity in COS-1 cells to undetectable levels when 2 alanines were substituted. The loss of activity in COS-1 cells correlated with decreased expression of hemoprotein with a reduced difference spectrum of 450 nm (P450 species) and a corresponding increase in the inactive P420 species in insect cells and bacteria. The activities expressed per nanomole of P450 in insect microsomes were similar for P450 2C2 and the alanine substitution mutants, including the mutant with 2 alanines followed by 7 valines for this region inactivated the proteins in insect cells and bacteria. The signal anchor sequence in P450 2C2 can be divided arbitrarily into a Gly-rich region of residues 22–28, which includes three Gly residues, followed by a Pro-rich region from 30 to 37, which includes the highly conserved sequence PPGP. Pro and Gly tend to disrupt α-helices and are often present in bends in proteins (10). A bend or flexible segment may be required so that the catalytic domain can achieve the proper orientation with respect to the membrane for interaction with its substrate or its membrane-bound redox partner, NADPH:cytochrome P450 reductase. Alternatively, a bend or linker sequence may be required for the proper folding and assembly of P450 in the microsomal membrane. Substitutions of Ala for Pro residues in the PPGP sequence result in mutants that are expressed as apoproteins in yeast but do not exhibit reduced CO difference spectra at 450 nm suggesting that this region is important for folding (11). Substitutions for the first and last Pro of PPGP diminish activity of the mutants expressed in COS-1 cells, whereas substitutions, including Glu, for the second Pro have little effect on activity (12). These results suggest that the structure of this region is not a simple bend but are more consistent with a polyproline II helix structure, which is a structure commonly involved in protein-protein interactions (13, 14), so that intra- or inter-protein interactions of this region may be important for folding.

The sequence that immediately follows the N-terminal signal anchor and links it to the cytoplasmic domain, like the signal anchor, does not align with sequences in soluble bacterial P450s (9) suggesting it may have a role related to membrane binding. This sequence in P450 2C2 can be divided arbitrarily into a Gly-rich region of residues 22–28, which includes three Gly residues, followed by a Pro-rich region from 30 to 37, which includes the highly conserved sequence PPGP. The signal anchor sequence is not required for expression of fully functional mammalian P450s in bacteria (reviewed in Ref. 8). The membrane insertion function of this sequence thus appears to be important for the assembly of P450 in the membrane, but once the protein is folded, catalytic activity is independent of the sequence.

The sequence that immediately follows the N-terminal signal anchor and links it to the cytoplasmic domain, like the signal anchor, does not align with sequences in soluble bacterial P450s (9) suggesting it may have a role related to membrane binding. This sequence in P450 2C2 can be divided arbitrarily into a Gly-rich region of residues 22–28, which includes three Gly residues, followed by a Pro-rich region from 30 to 37, which includes the highly conserved sequence PPGP. The signal anchor sequence is not required for expression of fully functional mammalian P450s in bacteria (reviewed in Ref. 8). The membrane insertion function of this sequence thus appears to be important for the assembly of P450 in the membrane, but once the protein is folded, catalytic activity is independent of the sequence.

The sequence that immediately follows the N-terminal signal anchor and links it to the cytoplasmic domain, like the signal anchor, does not align with sequences in soluble bacterial P450s (9) suggesting it may have a role related to membrane binding. This sequence in P450 2C2 can be divided arbitrarily into a Gly-rich region of residues 22–28, which includes three Gly residues, followed by a Pro-rich region from 30 to 37, which includes the highly conserved sequence PPGP. The signal anchor sequence is not required for expression of fully functional mammalian P450s in bacteria (reviewed in Ref. 8). The membrane insertion function of this sequence thus appears to be important for the assembly of P450 in the membrane, but once the protein is folded, catalytic activity is independent of the sequence.
EXPERIMENTAL PROCEDURES

**Plasmid Construction**—Construction of plasmid pc2A containing P450 2C2 cDNA in pT7Z18R has been described (4). A KpnI/BamHI fragment of pc2A, which contains the P450 2C2 cDNA, was inserted into the mammalian expression vector pCMV5 to produce pc2A-CMV. Single strand DNA mutagenesis was performed as described (15). The following oligonucleotides were used as primers and uracil-containing single stranded DNA added from plasmid pc2A grown in Escherichia coli strain C2326 was used as template. 5'-CAGACCGCATGAGCTGGAAGCTTCTCTGCTG-3' (C2AA). 5'-GGGCAAGAGGGACATCGGAGGGCGCGCGTCGTCGAGCTGCAATGCGCTTCTCCTGGCCG-3' (C2AT). 5'-CTTCTAGATGACGTCGACGCTCTTCTCTCGGCG-3' (C2AT). 5'-GTCTGACACTTCGTGGCAGTGTGCTTCCTG-3' (C22). 5'-CTTACACTTGGGAACTGGTGTGTTGATGGTGTCCTTCCTG-3' (C2VT).

To construct mutants in the pCMV5 mammalian expression vector, each pc2A mutant was digested with KpnI and ApaI, and the isolated fragment containing a mutation in the Gly-rich region was substituted for the corresponding fragment in pc2A-CMV. The mutations were confirmed by sequencing the KpnI-ApaI segment. The signal anchor sequence for Δ[22–28]C22 was derived from P450 2C1 which differs only by a Pro for Leu substitution at residue 3 that is on the luminal side of the membrane and should not affect targeting or topology of the protein (16).

**COS-1 Cell Culture and Assay of Laurate Hydroxylase Activity**—Culture of COS-1 cells and transfection of the cells with plasmid DNA were as described (17). Laurate hydroxylase activity was assayed in whole cell lysates of transfected COS-1 cells, and lauric acid metabolites were separated by high performance liquid chromatography as described (17) except that the reaction was incubated for 30 min.

**Biocatalytic Labeling of COS-1 Cells and Immunoprecipitation**—Forty-eight h after transfection, cells were precultivated for 30 min in Met- and Cys-free minimal essential medium. To estimate the relative expression levels of the P450 mutants, cells were incubated for 4 h with 50 μCi/μl Tran35S-label in Met- and Cys-free minimal essential medium. For studies of the stabilities of the proteins, the cells were labeled for 30 min followed by a chase incubation in complete medium. The cells were lysed, and radioactive proteins were immunoprecipitated from cell lysates and analyzed by SDS-PAGE as described previously (18).

**Expression of P450 in Insect Cells**—The wild-type and mutant P450 cDNAs were cloned into baculoviral genomic DNA downstream from the baculovirus polyhedrin gene promoter. Each mutant in a pTZ vector was digested with SfiI and XhoI, and the isolated cDNA fragment was ligated to the pFASTBACT1 vector (Life Technologies, Inc.) cut with SfiI and XhoI. The resulting plasmids were then transfected into DH10Bac cells on plates containing 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (isopropyl β-D-thiogalactopyranoside) (100 μl). Cells were lysed by a single pass through a French pressure cell at 18,500 p.s.i., E. coli membranes were isolated, and the proteins were partially purified on a hydroxylapatite column as described for P450 2C1 and P450 2C2 (20).

**Reduced CO Difference Spectra of P450s**—P450s expressed in COS-1 cells were not detectable by determining the reduced CO difference spectra so that expression in bacteria or insect cells was required to obtain sufficient amounts for this purpose. To obtain a reduced CO difference spectrum, the insect cells containing expressed protein were diluted in the glycerol buffer, and cells were lysed by sonication with two 2-s bursts at a setting of 4 with a Branson Sonifier, model 200. The difference spectrum after reduction with sodium hydrosulfite and binding of CO was obtained, and the amount of P450 was calculated by the difference between absorbance at 450 and 490 nm as described (22). Difference spectra were determined on the partially purified bacterially expressed proteins by the same procedure except that 1 μM methyl viologen was added to accelerate the formation of the reduced CO-bound proteins (23).

**Determination of the Amount of P450 Protein by Western Blotting**—Twenty μl of insect cell lysate in the glycerol buffer was mixed with SDS-PAGE sample buffer. The protein was separated by SDS-PAGE and transferred to nitrocellulose membrane (24). The membrane was then blocked in 5% skim milk and incubated with antibody raised against the P450 2C2 (25). A second antibody, alkaline phosphatase-conjugated rabbit anti-porcine IgG, was used against the first antibody, and the membrane was incubated with the 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium system (Sigma) to detect the second antibody.

**Determination of Laurate Hydroxylase Activity**—P450 reductase was expressed in bacteria and purified by 2’5’-ADP-agarose affinity chromatography (26), and the specific activity was determined by oxidation of cytochrome c as described (27). One unit is defined as the amount of enzyme that reduces 1 μmol of cytochrome c/min. Laurate hydroxylase activity was assayed in whole cell lysates of infected insect cells containing 40 pmol of P450 as described (17) except that the reaction was incubated for 15 min, and 8 units of P450 reductase were added so that P450 was limiting. The reaction under these conditions was linear for time with 15 min and with amount of P450 added (not shown). For the bacterially expressed enzymes, the activity was assayed in a reconstituted reaction as described (20) except that 0.2 units of P450 reductase per pmole of P450 were used.

**RESULTS**

**Sequence Requirements for Laurate (ω-1)-Hydroxylase Activity of P450 2C2 in the Gly-rich Region**—In previous studies, a chimera of P450, with amino acids 1–21 deleted and fused to the luminal and transmembrane domains of epidermal growth factor receptor, was active as P450 2C2 for laurate hydroxylase activity on transfected COS-1 cells (28). A similar chimera, when used amino acid 28 of P450 2C2 was inactive, suggesting that the region from 22 to 28 might be required for activity. To examine the functional significance of this region, residues 22–28 of P450 2C2, Δ[22–28]P2C2, were deleted. Laurate hydroxylase activity was not detected in COS-1 cells expressing this mutant (Fig. 1), and the amounts of immunoreactive P450 were de-
Fig. 1. Laurate (ω-1)-hydroxylase activity of mutants in the Gly-rich region of P450 2C2 expressed in COS-1 cells. Cell lysates from COS-1 cells transfected with P450 2C2 and the mutants were assayed for laurate hydroxylase activity as described under "Experimental Procedures." The designation for each mutant is shown at the right, and the sequences from residues 21 to 30 of P450 2C2 are shown with substituted amino acids in italics. The means of three to eight independent transfections and the S.E. of the mean are shown.

creased (Figs. 2 and 3). The region 22–28 contains an unusual triglycerine sequence that might be required for flexibility or a nonhelical structure of this segment. The three Gly were replaced with Ala (C2AAA), commonly in α-helices, or Pro (C2PPP), which has much less conformational freedom than Gly. Both of these mutants were equal to or higher in activity compared with P450 2C2 when expressed in COS-1 cells (Fig. 1), which indicates that there are no strict sequence requirements for Gly at these positions. To examine whether other amino acids in this region are important for enzymatic activity, mutant C2A7 with seven Ala substituted for residues 22–28 was constructed. The activity of C2A7 was 89% that of wild type (Fig. 1). In contrast, if seven Val (C2V7) were substituted for amino acid residues from 22 to 28, no laurate hydroxylase activity was detected in the COS-1 cells. These results suggest that the structural requirements of the region from 22 to 28 are relaxed since Ala can be substituted for each amino acid without substantially affecting activity. However, the inactivity if 7 Val are substituted suggests that either increased hydrophobicity or decreased flexibility is not tolerated since Val is more hydrophobic than Ala and has a tertiary β-carbon that increases the stiffness of the polypeptide backbone. Substitution of 3 Pro for Gly did not substantially affect activity which suggests that hydrophobicity may be more critical than flexibility since Pro restricts the conformation but is less hydrophobic than Ala.

Dependence of Laurate (ω-1)-Hydroxylase Activity on the Length of the Gly-rich Region—Relaxed structural requirements and flexibility suggest that the Gly-rich region might be a spacer sequence, in which case a minimum length for the sequence would be expected. To examine this property, either 2 (C2A2), 3 (C2A3), or 4 (C2A4) Ala were substituted for the seven residues from 22 to 28. The shortest construction, like deletion of the entire sequence, resulted in no detectable activity when expressed in COS-1 cells. However, the activities of C2A3 and C2A4 progressively increased to about 40 and 65%, respectively, that of P450 2C2 (Fig. 1). These results suggest that the length of the Gly-rich region must be more than two amino acids to produce a functional P450 2C2, and four or more amino acids are required for maximal expression of active protein.

Effect of Mutations on the Level of Expression in Transfected Cells—A simple explanation for the effects of mutations in the Gly-rich region is that the mutations affect the rate of synthesis or degradation of stable protein in the COS-1 cells. To examine the levels of the proteins expressed in the COS-1 cells, transfected cells were labeled for 4 h with Tran[35S]-label, and P450 2C2 and the mutant proteins were immunoprecipitated and analyzed by SDS-PAGE. In mock-transfected cells, only a weak band comigrating with P450 was detected, and a second protein migrating more rapidly was present which served as a useful internal control for labeling and immunoprecipitation (Fig. 2). Since the half-life of the P450 expressed in COS-1 cells is less than 1 h (17), labeling the cells for 4 h provides a reasonable measure of the steady-state levels of the protein. Differences of only 2-fold or less in the amount of radioactive protein immunoprecipitated for all of the mutants and wild-type P450 2C2 were detected in two or three independent experiments except for [Δ22–28]C2 (Fig. 2). Therefore, the loss of activity in some of the mutant proteins did not result from different levels of protein synthesis or degradation in COS-1 cells. The steady-state level of [Δ22–28]C2 expressed was about one-third of wild type, so that [Δ22–28]C2 was either poorly expressed or more rapidly degraded after expression, but the decreased amount of protein does not completely account for the undetectable enzymatic activity.

To examine whether synthesis or degradation of [Δ22–28]C2 was affected, a pulse-chase experiment was performed. Forty-eight h after transfection, COS-1 cells were labeled for 30 min and subsequently chased in complete medium for 3 h. Immunoprecipitated proteins were analyzed by SDS-PAGE and autoradiography. The original level of [Δ22–28]C2 was less than wild-type P450 2C2 (Fig. 3, lanes 3 and 8), and during the chase [Δ22–28]C2 was more rapidly degraded than wild-type P450 2C2 and the degradation may be multiphasic. As estimated by densitometry, the initial half-life of wild-type P450 2C2 is about twice as long as that of [Δ22–28]C2. The data indicate that [Δ22–28]C2 is expressed in COS-1 cells and more rapidly degraded after expression than wild type which largely accounts for the decreased amounts of expressed protein.

Effects of Mutations in the Gly-rich Region on the Expression of Hemoprotein in Insect Cells and Bacteria—The decreased activity in COS-1 cells of mutants in the Gly-rich region that had normal levels of apoprotein could result from decreased expression of properly assembled P450 2C2 or to direct effects of the mutations on the catalytic activity of the enzyme. The levels of expression in the COS-1 cells are not sufficient to determine the level of functional P450 expressed by reduced CO difference spectra, for which a maximum of 450 nm is characteristic of a normal conformation in the environment of the heme and the ability to bind CO, an analog of the substrate, O2. To obtain larger amounts of the mutant proteins, the mutants were expressed either in insect cells by infection with recombinant baculovirus or in E. coli. The spectral properties of the mutants were determined either in a natural membrane context in lysed insect cells or in solubilized, partially purified preparations for the bacterially expressed proteins.

In the insect cells, the amount of functional P450 produced correlated with activity of the mutants in COS-1 cells. Ala
substitution mutants with 3–7 Ala retained 40 to 100% laurate hydroxylase activity compared with P450 2C2 in COS-1 cells. These mutants contained similar amounts of functional P450 as wild type (Fig. 4, Table I). In contrast, the amount of the P450 species in insect cells was reduced by 80% for C2A2, which had undetectable activity in COS-1 cells. All of the mutants with Ala substituted had an increase in the amount of P420 that was expressed, and in the case of C2A2, the increase in P420 correlated with the decrease in P450 so that the total amount of hemoprotein expressed was similar to that of P450 2C2. Likewise, for these mutants, the amount of immunoreactive protein, relative to P450 2C2, was similar (Table I) so that the ratio of hemoprotein to apoprotein was similar for each variant which indicates that the efficiency of heme incorporation was not affected by the mutations. These results indicate that Ala substitution mutations in the 22–28 region are affecting the ability of the protein to fold into the correct functional conformation, but heme is incorporated since the total amount of P450 plus P420 and the ratio of hemoprotein to immunoreactive protein are about the same for the mutants and wild type.

Deletion of the entire region from 22 to 28 or substitutions of Val had the most severe effects. For the deletion mutant, no P450 or P420 hemoprotein nor immunoreactive protein was detectable in the insect cells. The P420 species, but not P450, was detected for the mutant with 7 Val substituted, consistent with undetectable activity of this mutant in COS-1 cells. In addition, the total amount of hemoprotein and immunoreactive protein for this mutant was reduced. Deletion of the entire region thus prevents the protein from folding into a stable protein, and reduced amounts of both hemoprotein and immunoreactive protein for the C2V7 mutant suggest that the stability of the protein is reduced in this case as well.

Mutants expressed in E. coli had properties similar to the proteins expressed in insect cells (Fig. 4 and Table II), although the mutants were generally expressed at lower levels relative to wild type. The amount of P450 expressed per 500 ml of cultured cells decreased progressively as the length of the linker region decreased. Nevertheless, a small amount of P450 was present even for the C2A2 that was inactive in COS-1 cells. As in the insect cells, the decreasing amounts of P450 were accompanied by increased amounts of P420. In contrast to the insect cells, the increases in P420 did not compensate entirely for the loss of P450 so that total hemoprotein decreased as the number of Ala substituted was decreased to 2. Hemoprotein is not detectable in insect cells for the [A22–28]C2 mutant, but the P420 species is observed in bacteria at a level 15% total hemoprotein of P450 2C2. These differences presumably reflect varying effects of the different environments for folding in the bacteria and insect cells. Nevertheless, a qualitative correlation between the expression of the P450 species and the activity of the mutants expressed in COS-1 cells is present in the bacterial system.

Specific Activity of Mutants in the Gly-rich Region—In addition to effects of the mutations on assembly of P450, it is possible that once a mutant protein has folded, its activity is also decreased. Since measurable amounts of P450 are produced in the insect cells and bacteria, the enzymatic activity can be expressed on the basis of the amount of functional P450 present. In the insect cells, the specific activities of the mutants with different numbers of Ala substituted for residues 22–28 ranged from 0.25 to 0.53 nmol/min/nmol and were similar to or greater than wild type (Table I). This result is particularly striking for C2A2, which is expressed in insect cells to a level only about 20% that of P450 2C2 and is not detectably active in COS-1 cells. Likewise, solubilized and partially purified mutant proteins expressed in bacteria had specific activities considerably greater than expected, based on the COS-1 cells studies. The specific activity of the bacterially expressed C2A4 and C2A3 was about 50% wild type and C2A2 was about 20% (Table II). The assay of the bacterially expressed, partially purified and solubilized P450s in a reconstituted system rather than in natural membranes as in the insect cell systems may be more sensitive to mutations in the protein. Nevertheless, the 20% level of specific activity of C2A2 contrasts with the undetectable activity in COS-1 cells. These data indicate that substitution of Ala within the Gly-rich region and shortening its length has little influence on the catalytic activity of P450 2C2 in insect microsomal membranes and modest effects on the activity of solubilized bacterially expressed P450 once the proteins have folded into the functional P450 species.

Stabilities of the P450 Species of Gly-rich Region Mutants—The increase in the amount of P420 in all the Ala substitution mutants suggests that the P450 species may be less stable and is converted to P420 after synthesis in vivo or during manipulations after lysis of cells. The similar specific activities of the Gly-rich region mutants and P450 2C2 in lysates of insect cells, however, indicate that the stabilities of the P450 species are similar at least during the time of the reaction. A variety of
conditions, including high temperature, have been shown to convert P450 to P420 (29). To examine whether the Gly region mutants were more susceptible to conversion from P450 to P420, the effects of incubation at 48 °C on the amounts of P450 and P420 were determined for partially purified and bacterially expressed P450 2C2, C2A7, and C2A3. The two mutants have sufficient P450 to obtain measurements and have increased amounts of P420, and C2A3 has about a 60% reduction in activity compared with P450 2C2 in COS-1 cells. The rate of loss of P450 resulting from incubation of the proteins at 48 °C was similar for all three proteins (Fig. 5). The half-lives of P450 loss for P450 2C2 and C2A7 were about 20 min, and C2A3 was slightly less stable with a half-life of about 15 min (Fig. 5B), which is not sufficient to explain the increase in P420 expressed in bacteria or insect cells. The kinetics of P420 were more complex and thus are not shown on the plot but increased at early times, reached a plateau, and then decreased slightly. These kinetics reflect conversion of P450 to P420 and degradation at the high temperature, and there were no obvious differences among P450 2C2 and the mutants. These data suggest that the effects of mutations in the Gly-rich region do not increase the rate of conversion of P450 to P420 but affect assembly of the protein so that increased amounts of P420 are formed directly.

DISCUSSION

These studies indicate that the Gly-rich region of P450 2C2 between the N-terminal signal anchor and the catalytic domain functions as a spacer or a linker. Sequence requirements are relaxed as evidenced by functional mutant protein when 7 Ala are substituted for this region. Substitution of 7 Val, however, results in inactive enzyme. Val results in decreased flexibility of the peptide backbone, but substitution of 3 Pro for the 3 Gly, which substantially reduces the conformational freedom of the chain, does not result in reduced activity. Val is also more hydrophobic than Ala whereas Pro is less. These results suggest that the increased hydrophobic character of the Val is responsible for the loss in activity. Relatively high hydrophobicity of a linker near the membrane/cytoplasmic interface might be detrimental since it would favor partition into the lipid interior of the membrane. In addition, a minimum requirement for the length of the spacer is indicated by the loss of activity in COS-1 cells when less than 3 Ala are substituted for this region. Increasing the length of this region by one amino acid by insertion of an Ala between residues 29 and 30 did not affect activity of P450 2C2 expressed in COS-1 cells (12). Relaxed sequence requirements and a minimum length are expected characteristics of a linker segment.

Short sequences that link domains in soluble proteins usually have a coil or turn secondary structure and a higher abundance of Gly, Thr, Ser and other small hydrophilic amino acids and a lower abundance of larger charged residues, Glu and Arg, and hydrophobic residues (30). The secondary structure and amino acid composition of the Gly-rich segment of P450 2C2 are in agreement with these expected properties of an oligopeptide linker. The secondary structure of the Gly-rich region is predicted to be either a turn or coil by several methods (31–34). In mammalian family 2 P450s, Gly, Ser, and the hydrophilic amino acids, Gln, Lys, Arg, and His, are in high abundance and hydrophobic amino acids, Val, Leu, Ile, Phe, and Met, are in low abundance (Table III). The bias against

---

TABLE I

| P450 variant | P450 | P420 | Total P450 + P420 | P450/P420 | Laurate hydroxylase activity |
|--------------|------|------|------------------|-----------|-----------------------------|
| P450 2C2 (wt) | 1.29 ± 0.18 | 100 | 0.46 ± 0.31 | 100 | 2.80 | 100 |
| C2A7 | 0.81 ± 0.16 | 63 | 0.65 ± 0.07 | 141 | 1.25 | 96 |
| C2A4 | 1.31 ± 0.03 | 101 | 1.07 ± 0.34 | 233 | 1.22 | 123 ± 7 |
| C2A3 | 0.85 ± 0.20 | 66 | 0.89 ± 0.35 | 193 | 0.96 | 139 ± 4 |
| C2A2 | 0.28 ± 0.11 | 22 | 1.33 ± 0.60 | 289 | 0.21 | 86 ± 4 |
| (Δ22–28)C2 | 0 | 0 | 0 | 0 | 0 | 0 |
| C2V7 | 0 | 0 | 6.1 ± 0.19 | 132 | 35 | 26 ± 2 |

* Total P450 or P420/100-mm culture dish.

a nmol/min/nmol of P450.

b S.E. of the mean with n = 3.

c Determined by Western blotting.

d Laurate hydroxylase activity.

---

TABLE II

Expression of apoprotein, P450, and P420 in E. coli cells and laurate hydroxylase activities of proteins partially purified from solubilized membranes by hydroxylapatite chromatography.

The proteins were expressed in E. coli and partially purified as described under “Experimental Procedures.” Reduced CO difference spectra were recorded to estimate the amount of hydroxylated [14C]lauric acid formed in 15 min in a reconstituted reaction as described under “Experimental Procedures.” Values are averages of two independent preparations of P450.

| P450 variant | P450 | P420 | Total P450 + P420 | P450/P420 | Laurate hydroxylase activity |
|--------------|------|------|------------------|-----------|-----------------------------|
| P450 2C2 (wt) | 7.1 | 100 | 0.25 | 100 | 100 | 0.53 |
| C2A7 | 4.0 | 57 | 0.52 | 208 | 62 | 7.8 |
| C2A4 | 1.3 | 18 | 1.02 | 408 | 31 | 1.2 |
| C2A3 | 1.1 | 16 | 1.08 | 432 | 30 | 1.0 |
| C2A2 | 0.5 | 8 | 1.02 | 408 | 21 | 0.51 |
| (Δ22–28)C2 | 0 | 0 | 1.07 | 428 | 15 | 0 |
| C2V7 | 0 | 0 | 0.84 | 336 | 12 | 0 |

a Total P450 or P420/500 ml of cultured cells.

b nmol/min/nmol of P450.

c Laurate hydroxylase activity.

---

These studies indicate that the Gly-rich region of P450 2C2 between the N-terminal signal anchor and the catalytic domain functions as a spacer or a linker. Sequence requirements are relaxed as evidenced by functional mutant protein when 7 Ala are substituted for this region. Substitution of 7 Val, however,
hydrophobic amino acids in linking peptides is consistent with the inactivity of P450 2C2 when the more hydrophobic Val is substituted. There are also some striking differences in the amino acid abundances between the Gly-rich region of P450s and linking peptides of the soluble proteins. The bulky basic amino acids, Arg and His, are in high abundance in P450 sequences, acidic residues are rare, and some of the subfamilies have substantial numbers of Trp (Table III). These differences may reflect the presumed location of the linker at the membrane/cytoplasm interface. Basic amino acids may provide interactions with the mostly negatively charged phospholipids of the membrane, and the side chain of Trp has both hydrophobic and hydrophilic character, is the most favored amino acid for binding in the interfacial region (35), and is often near the interfacial boundary in membrane proteins (36).

The length of the Gly-rich linker region in P450 2C2 is 7 if the amino acids following the first charged residue at the C-terminal end of the signal anchor sequence to the residue preceding the highly conserved hydrophobic amino acid before the Pro-rich region are considered. In a survey of 161 other mammalian P450s,3 the length of the linker region ranges from 4 to 17 residues, and the most frequently occurring length (55 of 161) is seven amino acids which is close to the average length of 6.5 for oligopeptide linking domains (30). Notably none of the P450s has a linker of less than four, and three is the minimal length required to express a functional P450 2C2 in COS-1 cells. Thus, even though the actual sequence in this region is not conserved, segments with minimal lengths which consist of largely hydrophilic amino acids are present in nearly all P450s which indicates the requirements for this region determined for P450 2C2 are likely to be important for other P450s as well.

A possible model for the interaction of the signal anchor sequence and the Gly-rich linker region with the membrane at the cytoplasm/membrane interface may be derived from these studies. Near the C-terminal end of the hydrophobic core of the signal anchor, a Trp is often conserved in family 2 P450s which could form interactions with both the phospholipid head groups and apolar carbon chains near the interface. The Trp is followed by a Lys and the Gly-rich region which lies along the hydrophilic membrane interfacial region. Polar and basic amino acids and Trp residues within the linker region could interact with the phospholipid headgroups of the membrane and the water molecules in the cytoplasm. An extended conformation and exposure to the aqueous environment expected of a linker sequence is supported by the ability of thrombin to cleave at a site introduced into P450 1A2 at a position equivalent to the P450 2C2 Gly-rich region (37). This linker region would lead to the first sequence that is part of the folded catalytic domain. Determination of the exact boundary between the linker and the catalytic folded protein will require additional experiments, but the highly conserved Pro-rich region that may form interactions with other peptides (12) is a reasonable candidate for the beginning of the folded protein.

The major function of the “linker” region is related to the

---

3 Survey was compiled by D. R. Nelson and is available at [http://www. drnelson.utmem.edu/nelsonhomepage.html](http://www.drnelson.utmem.edu/nelsonhomepage.html).
assembly of the protein into a functional form. Deletion of the entire region from 22 to 28 resulted in no stable immunoreactive protein in insect cells and no detectable activity and less stable immunoreactive protein in COS-1 cells consistent with a major disruption of protein folding. Substitution of 7 Val, likewise, resulted in no functional P450 and reduced amounts of total hemoprotein insect cells and bacteria. Among the Ala substitution mutants, the C2A2 mutant had the most dramatic effect in COS-1 cells, in which no activity was observed. In insect cells, although the amount of P450 was decreased, the total amount of P450 and P420 expressed was similar to that of P450 2C2. This suggests that protein folds correctly and incorporates heme but that a subtle folding defect or conformational change occurs that alters the interaction between the heme and protein since the spectra change from 450 nm to 420 nm is thought to result from an altered interaction of the thiolate ligand with the heme (29). More importantly, the specific activity of C2A2 in insect cell membranes was similar to that of P450 2C2. This result establishes that the major effect of shortening the linker is on the assembly of functional protein and neither the conformation of the folded catalytic domain of P450 nor its interaction with the membrane is affected in a functionally significant way.

The increased expression of the P420 species for several of the mutants could be due either to initial formation of P450 followed by conversion to P420 or to direct folding into the P420 species. Increased formation of P420 from P450 in the cells remains a possibility but is not supported by the observation that loss of P450, when partially purified preparations are incubated at high temperature, is similar for C2A7, C2A3, and P450 2C2. Assuming stability during incubation at high temperature in vitro accurately reflects the stability of P450 within the cell, the most likely effect of the mutants is to alter the assembly process so that there is a higher probability of the protein folding into an inactive P420 species rather than P450. Once folded, both forms would be stable.

It is possible that the requirement for a linker is related to the assembly of the functional protein at the membrane-cytoplasm interface. Although deletion of the region from 22 to 28 of P450 2C2 resulted in no detectable P450 in bacteria, deletion of the signal anchor as well as this region resulted in expression of active enzyme as has been observed for P450 2B4 and P450 2E1 (38). The lack of sequences in soluble bacteria P450s that align with the 22–28 region would also be consistent with its role in assembly at the membrane.

Acknowledgments—We thank Dr. Todd Porter for the pNIImpA3 expression vector for P450 reductase; Drs. Toby Richardson and Eric Johnson for the pCW expression vectors for P450 2C1 and P450 2C2; Drs. Jeng-Shong Chen and Mary Schuler for help in baculovirus expression systems; and Jocelyn McDonald for the initial construction of the deletion mutant [A22–28/C2]. Secondary structure predictions were done using the Biology Workbench of the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign.

REFERENCES
1. Nelson, D. R., Kamataki, T., Wazan, D. J., Guengerich, F. P., Estabrook, R. W., Feyerisen, R., Gonzalez, F. J., Coon, M. J., Guengerich, F. P., and Gotoh, O., Okuda, K., and Nebert, D. W. (1993) DNA Cell Biol. 12, 1–51
2. Nelson, D. R., and Strobel, H. W. (1987) Mol. Biol. Ecol. 4, 572–585
3. Szczesna-Skorupa, E., Browne, N., Mead, D., and Kemper, B. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 738–742
4. Szczesna-Skorupa, E., and Kemper, B. (1989) J. Cell Biol. 108, 1237–1243
5. Sakaguchi, M., Mihara, K., and Sato, R. (1987) EMBO J. 6, 2425–2431
6. Sakaguchi, M., Mihara, K., and Sato, R. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 3361–3364
7. Menier, S., Luc, P. V., Kreibich, G., Sabatini, D. D., and Adesnik, M. (1988) J. Cell Biol. 107, 457–470
8. von Wachenfeldt, C., and Johnson, E. F. (1995) in Cytochrome P450-Structure, Mechanism, and Biochemistry (Ortiz de Montellano, P., ed), 2nd Ed., pp. 183–223, Plenum Publishing Corp., New York
9. Hasemann, C. A., Kurumbail, R. G., Boddupalli, S. S., Peterson, J. A., and Deisenhofer, J. (1995) Structure 2, 41–62
10. Engelman, D. M., Streitz, T. A., and Goldman, A. (1986) Annu. Rev. Biophys. Biophys. Chem. 15, 321–353
11. Yamasaki, S., Sato, K., Suhara, K., Sakaguchi, M., Mihara, K., and Ozuma, T. (1989) J. Biochem. (Tokyo) 114, 652–657
12. Chen, C.-D., and Kemper, B. (1990) J. Biol. Chem. 271, 28607–28612
13. Cicchetti, P., Mayer, B. J., Thiel, G., and Baltimore, D. (1992) Science 257, 803–806
14. Ren, R., Mayer, B. J., Cicchetti, P., and Baltimore, D. (1993) Science 259, 1157–1161
15. Mead, D. A., Szczesna-Skorupa, E., and Kemper, B. (1986) Protein Eng. 1, 97–74
16. Szczesna-Skorupa, E., and Kemper, B. (1993) J. Biol. Chem. 268, 1757–1762
17. Straub, P., Lloyd, M., Johnson, E. F., and Kemper, B. (1995) J. Biol. Chem. 268, 21997–22003
18. Szczesna-Skorupa, E., Strobel, H. W., and Kemper, B. (1993) Arch. Biochem. Biophys. 340, 170–175
19. Gharey, J., Kimura, H., Takahara, M., Hsu, I., Masui, Y., and Iouye, M. (1984) EMBO J. 3, 2437–2442
20. Richardson, T. H., Jung, P., Griffin, K. J., Wester, M., Raucy, J. L., Kemper, B., Barnhime, L. M., Hassett, C., Omiecinski, C. J., and Johnson, E. F. (1995) Arch. Biochem. Biophys. 323, 87–96
21. Barnes, H. J., Arlotti, M. P., and Waterman, M. R. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 5597–5601
22. Omura, T., and Sato, R. (1964) J. Biol. Chem. 239, 2370–2378
23. Richardson, T. H., Hu, M.-H., Cronbach, T., Barnes, H. J., Chan, G., Waterman, M. R., Kemper, B., and Johnson, E. F. (1993) Arch. Biochem. Biophys. 300, 510–516
24. Towbin, H., Staehelin, T., and Gorgon, J. (1979) Proc. Natl. Acad. Sci. U. S. A. 76, 4550–4554
25. Kronbach, T., Kemper, B., and Johnson, E. F. (1991) Biochemistry 30, 6097–6102
26. Yasukochi, Y., and Masters, B. S. S. (1976) J. Biol. Chem. 251, 5337–5344
27. Vermilion, J. L., and Coon, M. J. (1978) J. Biol. Chem. 253, 2694–2704
28. Szczesna-Skorupa, E., Ahn, K., Chen, C.-D., Doray, B., and Kemper, B. (1995) J. Biol. Chem. 270, 24327–24333
29. Martinis, S. A., Blanken, S. R., Hager, L. P., Sligar, S. G., Hui Bon Hoa, G., Rux, J. J., and Dawson, J. H. (1996) Biochemistry 35, 14530–14536
30. Argo, P. (1990) J. Mol. Biol. 211, 943–958
31. Chou, K. C. (1995) Proteins Struct. Funct. Genet. 21, 319–344
32. Levin, J. M., and Garner, J. (1988) Biochim. Biophys. Acta 95, 283–295
33. Mao, B., Chou, K. C., and Zhang, C. T. (1994) Protein Eng. 7, 319–330
34. King, R. D., and Sternberg, J. E. (1993) J. Mol. Biol. 216, 441–457
35. Wimley, W. C., and White, S. H. (1996) Nat. Struct. Biol. 3, 842–848
36. Michel, H., and Deisenhofer, J. (1990) Curr. Top. Membr. Transp. 36, 53–59
37. Dong, M. S., Bell, L. C., Guo, Z. Y., Phillips, D. R., Blair, I. A., and Guengerich, F. P. (1996) Biochemistry 35, 10631–10640
38. Pernecky, S. J., and Coon, M. J. (1996) Methods Enzymol. 272, 25–34