Preliminary Characterization and Crystal Structure of a Thermostable Cytochrome P450 from Thermus thermophilus

Received for publication, July 2, 2002, and in revised form, September 4, 2002
Published, JBC Papers in Press, October 24, 2002, DOI 10.1074/jbc.M205682000

Jason K. Yano†, Francesca Blasco§, Huiying Li†, Rolf D. Schmid§, Anke Henne†, and Thomas L. Poulos‡‡

From the †Department of Molecular Biology and Biochemistry, the Department of Physiology and Biophysics, and the Program in Macromolecular Structure, University of California, Irvine, California 92697-3900, the §Institut für Technische Biochemie, Universität Stuttgart, Allmandring 31, D-70569 Stuttgart, Germany, and the ‡Institute of Microbiology and Genetics, University of Göttingen, Grisebachstrasse 8, 37077 Göttingen, Germany

The second structure of a thermophile cytochrome P450, CYP175A1 from the thermophilic bacterium Thermus thermophilus HB27, has been solved to 1.8-Å resolution. The overall P450 structure remains conserved despite the low sequence identity between the various P450s. The CYP175A1 structure lacks the large aromatic network found in the only other thermostable P450, CYP119, thought to contribute to thermal stability. The primary difference between CYP175A1 and its mesophile counterparts is the investment of charged residues into salt-link networks at the expense of single charge interactions. Additional factors involved in the thermal stability increase are a decrease in the overall size, especially shortening of loops and connecting regions, and a decrease in the number of labile residues such as Asn, Gln, and Cys.

Cytochromes P450 (P450s) catalyze the monooxygenation of a vast array of organic compounds in the following reaction.

\[ \text{NAD(P)H} + H^+ + R-H + O_2 \rightarrow \text{NAD(P)^+} + H_2O + ROH \]

**REACTION 1**

P450s are ubiquitous enzymes essential for steroid biosynthesis, catabolism of drugs, utilization of carbon compounds as an energy source in bacteria, and in the production of various macrolide antibiotics. P450s are found throughout the biosphere and various genome projects have revealed a remarkable number of P450s in a variety of organisms. For example, humans have ~180, Caenorhabditis elegans 81, Drosophila 90 (83 functional 7 pseudogenes) (1), and Arabidopsis 273 (drenelson.utmem.edu/CytochromeP450.html). Several bacteria also have P450s, but normally bacteria produce only one or at most a few P450s.

Another surprising occurrence of a P450 was demonstrated in 1996 when the first P450 from a hyperthermophile, CYP119, was discovered (2). CYP119 was identified in the acidotherophilic archaeon Sulfolobus solfataricus. This was not only the first thermostable P450 to be discovered but also the first demonstration of an archaeon P450. Subsequent cloning, expression, and purification of CYP119 led to the finding that CYP119 exhibits a melting temperature of ~90 °C compared with ~55 °C for mesophilic bacterial P450 (3). This was soon followed by our determination of the CYP119 crystal structure (4), followed by an independent structure determination by Park et al. (5). The enhanced thermal stability of CYP119 was attributed to an unusually large clustering of aromatic side chains not found in all other known P450 structures (3-6). The ever increasing data base of known structures of proteins from both mesophiles and their thermophilic counterparts reveals that aromatic clustering does, indeed, correlate with enhanced thermal stability, but that this is only one of several possible mechanisms for enhancing thermal stability (7). Whether or not aromatic clustering is a universal feature of thermostable P450s will require additional crystal structures. Fortunately, the second P450 from a thermophile was recently discovered, CYP175A1 from Thermus thermophilus.

The present report describes cloning, expression, purification, crystallization, and structure determination of CYP175A1, a cytochrome P450 from the thermophilic bacterium Thermus thermophilus HB27. The 44-kDa protein is soluble and displays an absorption spectra in the reduced, oxidized, and carbonyl adduct analogous to those of other P450 enzymes. We also demonstrate that this enzyme exhibits thermostability with a melting temperature 30 °C higher than that of P450. To understand whether CYP175A1 and CYP119 share common structural features related to thermal stability, the two crystal structures are compared.

**MATERIALS AND METHODS**

Gene Extraction, Expression, and Purification of P450 (CYP175A1) from Thermus thermophilus HB27 (Tt)—Plasmid DNA containing CYP175A1 was a kind gift from Dr. Hans-Joachim Fritz (University of Göttingen, Göttingen, Germany). Polymerase chain reaction (PCR) was used to engineer EcoRI and PstI restriction sites into the CYP175A1 gene. The oligonucleotide primers used for amplification of the new CYP175A1 gene from the original plasmid were the following: primer A, 5'-CCGGAATTCATGAAGCGCCTTTCCCTGAGG; primer B, 5'-CCGGAGCTCGAGCTCAGGAGGGCAGGACGAGGCTTCCTCCTGAGG.

The new restriction sites, EcoRI and PstI, are underlined. The reaction mixture consisted of template DNA (100 ng), 2.5 units of Pfu DNA polymerase (Stratagene), 5 μl reaction buffer, 5 μl of dNTPs, and H2O up to 50 μl. PCR cycling parameters were: 95 °C for 1 min; 30 cycles of 95 °C for 1 min; 55 °C for 1 min; 72 °C for 1 min; and 72 °C for 4 min. The CYP175A1 gene sequence was confirmed by DNA sequencing.

After restriction enzyme digest of the gel-purified PCR product, the
CYP175A1 gene was cloned into the EcoRI/PsiI sites of the pKK233–3 plasmid (Amersham Biosciences). BL21 (DE3) Codon Plus Escherichia coli (Stratagene) were transformed with a pKK233–3 plasmid containing the gene for CYP175A1 from Tt. Single colonies were selected for overnight growth in 5-ml cultures, which were used to inoculate cultures (1 liter in each 2.8-liter flask). At an absorbance of 0.1, 500 μg/ml isopropyl β-D-thiogalactoside (United States Biochemical Corp.) was used to induce the expression of CYP175A1. The conditions for cell growth were 2 × YT (16 g of tryptone, 10 g of yeast extract, and 5 g of NaCl per liter), 0.1 g/liter ampicillin (Sigma), and 50 μg/ml chloramphenicol (Sigma) at 37 °C, with shaking at 220 rpm for the starter cultures and slow stirring (80–100 rpm) at 30 °C after induction. Forty-eight hours after induction, the cells were harvested and the cell paste was stored at −70 °C.

Cells were lysed in a French pressure cell in 50 mM potassium Pi buffer, pH 7.5, 0.1% Emulgen (Japan), and 1 mM phenylmethylsulfonyl fluoride (Sigma). Cell debris was removed by centrifugation, and the cell-free extract was brought to 35% ammonium sulfate concentration by adding solid ammonium sulfate. Precipitated proteins were removed by centrifugation, and the supernatant was brought to a concentration of 50% ammonium sulfate. The resulting precipitate was centrifuged, resuspended in minimal volume, and dialyzed against 50 mM potassium Pi buffer, pH 7.5, overnight (3 times buffer exchange). The dialyzed supernatant was passed over a DE52 column. Contaminant proteins stuck on the column, whereas CYP175A1 passed through because of its limited affinity for DE52 resin. The flow-through containing CYP175A1 had an absorbance ratio (A410 nm/A280 nm) of 1.3. The protein was concentrated further and diluted in 50 mM Tris, pH 7.5. The protein was then loaded onto a hydroxyapatite (Bio-Rad) column (2.5 × 20.0 cm). The protein was washed with two to three column volumes of 50 mM Tris, pH 7.5, followed by two to three column volumes of 50 mM potassium Pi buffer, pH 7.4. The protein eluted with a linear gradient of potassium Pi buffer, pH 7.4, from 50 to 300 mM over 10 column volumes at a flow rate of 0.5 ml/min. Fractions with an absorbance ratio (A410 nm/A280 nm) above 1.55 were pooled, concentrated, and exchanged into 50 mM Tris, pH 7.4, on a PD10 column (Amersham Biosciences). The protein was then concentrated to 40 mg/ml and stored at −30 °C. A viscosity coefficient of 0.034 ± 0.002 ml/g s was determined by light scattering. The percentage denatured was calculated by the appearance of a single band on a Coomassie Blue-stained SDS-page gel and had an absorbance ratio (A410 nm/A280 nm) of 1.80.

Melting Temperatures—The thermal melting temperatures were measured by recording the absorbance of the heme Soret peak at a linear temperature rate of 1 °C/min on a Cary 3E UV-visible spectrophotometer equipped with a temperature controller. All P450s were irreversibly denatured. P450BM-3 consists of the heme domain only (residues 1–455). The absorbance increase at the Soret peak was followed for CYP175A1 and CYP175A1, which was found to be 50% at 35 °C, 80% at 30 °C, and 100% at 25 °C. The absorbance at 37 °C was followed for all other P450s. Melting was introduced. In addition, simulated annealing omit maps were generated in AMoRe alone, because the rotation solution generated by finer rotational step size in BEAST was closer to the final orientation. AMoRe was used to find both the rotation and translation solution for the second molecule. The log likelihood gain score for the first rotation in BEAST was 20.07 with a mean of 7.50. The rigid body fitting routine in AMoRe was applied to the final solutions and had a pseudo R-factor score of 0.45 and a crystallographic R-factor of 52.7%. SOLVE (14) was used to interpret both the isomorphous difference Patterson from the platinum derivative and the native iron anomalous difference Patterson. SOLVE was able to locate two platinum sites and two iron sites in the asymmetric unit, in agreement with the predicted two molecules per asymmetric unit. The final Z score for the rotation was 14.08 with an overall figure of merit of 0.41. Data from 12 to 2.8 Å were used for phase combination calculations using SIGMAA (15). The combined phases were further improved by density modification using DM (16), including solvent flattening, histogram matching, and 2-fold noncrystallographic symmetry averaging to provide an interpretable electron density map.

Model Building and Refinement—Electron density map fitting beginning with the truncated P450BM-3 model was carried out with the graphic modeling package O (17). Several sections of the electron density map, especially the core near the heme group, clearly showed the identity of CYP175A1 side chains. Regions of the molecule that lacked clear backbone electron density were deleted prior to refinement in CNS (version 1.1) (18). Several rounds of model building followed by refinements with simulated annealing (at 3000 K) and minimization, continued until the R-factor was ~30%, and then isotropic B-factor refinement was introduced. In addition, simulated annealing omit maps were generated in CNS to assist in model building. When the R-factor reached ~25%, water molecules were automatically picked at the 3.5σ level in an Fo − Fc difference map using CNS and manually added in for the last few rounds of refinements. The final stage of refinement was done with the program REFMAC5 (19). The final model consists of all the residues in molecule A except the last four at the C terminus, whereas molecule B is missing residues 11–15 as well as the last four residues (20). The final R-factor is 18.2% (R-free 22.8%), and none of the main-chain torsion angles were in the disallowed regions of a Ramachandran plot. Relevant statistics are found in Table I.

RESULTS AND DISCUSSION

Thermal Stability—Melting curves for the two thermophilic P450s, CYP175A1 and CYP119, and three mesophilic P450s are shown in Fig. 1. The three mesophilic P450s exhibit Tm values ranging from 47 to 61 °C, whereas CYP175A1 and CYP119 have Tm values of 88 and 91 °C, respectively. These values compare favorably for those obtained using the more rigorous technique of differential scanning calorimetry. In these experiments, the Tm for CYP119 was found to be 91 °C, whereas that for P450cam was 54 °C (3, 20).

Structure Solution—Surprisingly, the highest sequence identity to CYP175A1 is P450BM-3 at 26%. Based on this added to the co-polymer protectants. A Matthews coefficient (9) calculation indicated that there were most likely two molecules per asymmetric unit.

Data Collection—All data were collected using crystals flash-frozen in a stream of liquid nitrogen. Initial data sets were collected in-house on an R-Axis IV image plate equipped with a Rigaku rotating copper anode x-ray generator and Osmic focusing optics. Collection of in-house native data was optimized to maximize the anomalous dispersion signal from the heme iron by collection of a highly redundant data set resulting in greater than one million observations of 61,926 unique reflections to 1.95 Å (Table I). A high resolution data set (50–1.80 Å) used for the final refinement was collected at the Advanced Light Source, beamline 5.0.2 with an Area Detector Systems Co. Quantum-4 CCD detector. A model was generated in the initial phases of the molecular replacement guided by the STRATEGY search program of MOSFLM (10). All data collected were reduced using DENZO and SCALEPACK (11).

Phase Combination—Phases were derived from molecular replacement, a platinum heavy atom derivative, and anomalous scattering from the heme iron. Molecular replacement was carried out using a combination of AMoRe (12) and BEAST (13) using data from 20 to 5 Å and a polyalanine search model of P450BM-3 (PDB accession no. 1FAG) with residues 1–15, 223–232, 293–303, and 337–350 deleted from the model. The rotation solution found in BEAST was used in AMoRe to find the translation solution for the first molecule. This provided a slightly higher signal to noise ratio than using the rotation solution generated in AMoRe alone, because the rotation solution generated by finer rotational step size in BEAST was closer to the final orientation. AMoRe was used to find both the rotation and translation solution for the second molecule. The log likelihood gain score for the first rotation in BEAST was 20.07 with a mean of 7.50. The rigid body fitting routine in AMoRe was applied to the final solutions and had a pseudo R-factor score of 0.45 and a crystallographic R-factor of 52.7%. SOLVE (14) was used to interpret both the isomorphous difference Patterson from the platinum derivative and the native iron anomalous difference Patterson. SOLVE was able to locate two platinum sites and two iron sites in the asymmetric unit, in agreement with the predicted two molecules per asymmetric unit. The final Z score for the rotation was 14.08 with an overall figure of merit of 0.41. Data from 12 to 2.8 Å were used for phase combination calculations using SIGMAA (15). The combined phases were further improved by density modification using DM (16), including solvent flattening, histogram matching, and 2-fold noncrystallographic symmetry averaging to provide an interpretable electron density map.

Model Building and Refinement—Electron density map fitting beginning with the truncated P450BM-3 model was carried out with the graphic modeling package O (17). Several sections of the electron density map, especially the core near the heme group, clearly showed the identity of CYP175A1 side chains. Regions of the molecule that lacked clear backbone electron density were deleted prior to refinement in CNS (version 1.1) (18). Several rounds of model building followed by refinements with simulated annealing (at 3000 K) and minimization, continued until the R-factor was ~30%, and then isotropic B-factor refinement was introduced. In addition, simulated annealing omit maps were generated in CNS to assist in model building. When the R-factor reached ~25%, water molecules were automatically picked at the 3.5σ level in an Fo − Fc difference map using CNS and manually added in for the last few rounds of refinements. The final stage of refinement was done with the program REFMAC5 (19). The final model consists of all the residues in molecule A except the last four at the C terminus, whereas molecule B is missing residues 11–15 as well as the last four residues (20). The final R-factor is 18.2% (R-free 22.8%), and none of the main-chain torsion angles were in the disallowed regions of a Ramachandran plot. Relevant statistics are found in Table I.
sequence identity, we chose P450BM-3 as the search model for molecular replacement. The solution to the phase problem for CYP175A1 by molecular replacement represents the only case where a known P450 structure could be used to solve both the rotation and translation functions. Previously, in the CYP119 structure determination (4), the rotation solution for only one of the two molecules in the asymmetric unit was obtained with a modified version of BRUTE (13, 21). The rotation parameters for the second molecule could not be found using routine molecular replacement programs and were generated by application of the self-rotation peak. The present structure determination demonstrates the ability to use preexisting P450 structures as molecular replacement models despite low sequence identity found throughout the P450 superfamily. Although the molecular replacement phases generated an interpretable electron density map, the map was significantly improved by incorporating phases derived from the platinum derivative and iron sites.

### Table I

| Data collection | Native<sup>a</sup> | Native<sup>b</sup> | Pt(NH₂)₂Cl₂ |
|-----------------|-------------------|-------------------|-------------|
| Wavelength (Å)  | 1.54              | 1.10              | 1.54        |
| Resolution range (Å) | 1.95–50.0      | 1.8–50.0          | 2.15–50.0   |
| Reflections Measured/unique | 1,433,974/81,926 | 314,853/84,011   | 690,919/46,677 |
| Completeness (%)/outer shell<sup>c</sup> | 99.9/89.6      | 94.1/72.5         | 98.9/91.9  |
| Mean (I/σI) overall/outer shell<sup>d</sup> | 6.6/59.0      | 5.7/35.4          | 9.5/39.0   |

### Phasing statistics

| Number of sites | 2 Fe | 2 Pt |
|-----------------|------|------|
| Resolution range (Å) | 2.8–20.0 |      |
| Figure of merit combined<sup>e</sup> | 0.41 |      |
| Quality of solution (SOLVE) | 14.03 |      |

### Refinement statistics

| Total no. of reflections | 67,977 |
| Total no. of atoms | 6835 |
| Total no. of waters | 529 |
| R<sub>crys</sub> (R<sub>free</sub>) (%) | 18.2 (22.8) |

### Model statistics<sup>f</sup>

| r.m.s. deviations | Bonds (Å) | 0.028 |
|-------------------|-----------|-------|
| Angles (°) | 2.1 |

<sup>a</sup> Native data set collected in-house on R-axis.

<sup>b</sup> Data collected at Advanced Light Source.

<sup>c</sup> R<sub>sym</sub> = Σ |I<sub>i</sub> – 〈I〉|/Σ I<sub>i</sub>, in which 〈I〉 is the mean intensity of reflection.

<sup>d</sup> Figure of merit (FOM) is mean of the cosine of the error in phase angle.

<sup>e</sup> Statistics for final model.

**Fig. 1.** Thermal melts of P450s. The two thermophile P450s undergo a thermal transition ~40 °C higher than the mesophile P450s. A, P450BM-3 (heme domain), T<sub>m</sub> = 47 °C; B, P450eryF, T<sub>m</sub> = 55 °C; C, P450cam, T<sub>m</sub> = 61 °C; D, CYP175A1, T<sub>m</sub> = 88 °C; E, CYP119, T<sub>m</sub> = 91 °C. The melting temperature for CYP119 agrees with differential scanning calorimetry data presented in previous papers (3, 20).
Overall Structure—CYP175A1 exhibits the typical prism-like P450-fold (Fig. 2) composed of 17 α-helices and 11 β-strands, corresponding to four β segments. At the core of the molecule is a four-helix bundle composed of helices D, E, I, and L and two α-helices (J and K) (22). The heme prosthetic group is embedded between the distal I and proximal L helices, with Cys-336 serving as the fifth axial thiolate ligand to the heme iron. The propionate side chains of the heme are tethered to the protein through interactions with Trp-87, Arg-273, and Arg-334. The I helix spans the length of the entire molecule and contains the highly conserved Thr (Thr-225) residue, which is hydrogen-bonded to the peptide backbone. This region has been postulated to be involved in a proton shuttle network considered to be important for delivering solvent protons required for the activation of O₂ during the catalytic cycle (23, 24). At 389 residues, CYP175A1 is comparable in size to CYP119 (378 residues), but considerably smaller than other known soluble bacterial P450s such as P450BM-3 (heme domain) at 455 residues.

Crystals used in this study were prepared in the presence of 1-phenylimidazole. As a result, we expected to find this ligand in the active site pocket coordinated to the heme iron. As shown in Fig. 3, there is clearly defined electron density for the inhibitor with continuous density connecting the inhibitor and heme iron atom consistent with an imidazole N atom coordinating to the heme iron. Nevertheless, attempts to accurately model inhibitor to the electron density failed to give good fits. In sharp contrast, the similar inhibitor, 4-phenylimidazole, fit the electron density very well in CYP119. It appears that the inhibitor may either be partially occupied, thus sharing space with solvent, or the inhibitor is partially disordered. Spectral titration data do indicate that the inhibitor does not bind as well to CYP175A1 as to CYP119. The apparent Kᵢ is ~0.3 μM for CYP119 and ~83.7 μM for CYP175A1 in the same buffer used for crystallization. In both cases there are clearly defined spectral shifts characteristic of forming an N–Fe bond. Although the inhibitor binds more weakly to CYP175A1, it remains surprising that the inhibitor structure was not more well ordered and/or fully occupied given the large excess of 1-phenylimidazole (1 mM) used for crystallization.

Comparison with P450BM-3—Because the closest homologue to CYP175A1 is P450BM-3, a bacterial fatty acid monooxygenase, we will next compare these two structures. The r.m.s. backbone deviation between CYP175A1 and palmitoleic acid-bound P450BM-3 (PDB accession no. 1FAG) (25) is 1.64 Å for 304 Ca atoms. Table II provides a comparison of secondary structural elements, and Fig. 2 provides a structural picture of the two P450s. Most of the differences between CYP175A1 and P450BM-3 are located in the loops connecting helices. For example, P450BM-3 has a seven-residue insertion in the loop connecting helices E and F (residues 159–172 in P450BM-3, residues 146–152 in CYP175A1). Another difference includes the β-5 region in P450BM-3 (residues 239–250) compared with the loop found in CYP175A1 (residues 202–205), a difference of eight residues. This region forms a tight loop between the H and I helices, and resembles what was previously found in P450nor, which does not have a redox partner (26), and CYP119. In contrast, all other soluble bacterial P450 structures show this ill defined β-hairpin similar to P450BM-3. Interestingly, the crystal structure of a complex formed between the heme domain of P450BM-3 and FMN domain has shown that this hairpin region makes direct contact with the FMN domain (27).

Although the redox partner for CYP175A1 remains unknown, it generally is thought that electrostatic interactions are important in redox partner recognition. Therefore, it is
instructive to examine the electrostatic surface potential of CYP175A1 compared with P450cam and P450BM-3, where there is a clearly clear picture on the location of the docking site. The electrostatic surface potential of CYP175A1, P40cam, and P450BM-3 are shown in Fig. 4. The arrows indicate the approximate location of the surface above the proximal Cys ligand. This region is known to be the docking site for the P450BM-3 reduct partner (27) and is generally considered the docking site in P450cam (28). The excessive positive charge in this region provides favorable electrostatic interactions with negative charges on the reduct partner. CYP175A1 also has a large positive electrostatic potential in this region, suggesting that this, too, may be a recognition site for reduct partners. The larger positive potential on CYP175A1 on the proximal surface CYP175A1 compared with the other P450s is caused in part by an extra arginine, Arg-319, in CYP175A1 with no homologue in P450cam or P450BM-3. In addition, Arg-319 occupies approximately the same position in three dimensions as Glu-286 in P450cam; however, CYP175A1 has one less negative charge in this region. This might provide an especially strong electrostatic docking site for the CYP175A1 reduct partner.

Another important region is the substrate binding environment, the F and G helices, B' helix, and neighboring loops (Fig. 5). The F and G helices and the loop connecting these helices are known to undergo a large open/close motion between the substrate-free and -bound forms of P450BM-3 (25). In addition, CYP119 undergoes a rather large structural change in the F/G region, depending on the type of ligand bound in the active site (4). When a small inhibitor, imidazole, is bound, the C-terminal end of the F helix in CYP119 unwinds, which increases the length of the F/G loop, thus allowing the F/G loop to dip down into the active site and interact with the inhibitor. These two examples illustrate the importance of the F/G region in substrate recognition and binding. It is therefore important to note that the major portions of the F-G region and the B' helix, known to be important in substrate binding (29), are in nearly identical conformations in CYP175A1 and P450BM-3 (Fig. 5). The B' helix forms a portion of the lid to the substrate access channel in P450BM-3 and makes intimate contact with the G helix. The average r.m.s. deviation for the B' helix is 1.23 Å, reflecting the general overlap of this structural element. The F helix in CYP175A1 (residues 150–170) is nearly identical in both size and location to the F helix in P450BM-3 (residues 171–190). The core region of the F helix, residues 157–167 in CYP175A1, has an average r.m.s. deviation of 1.66 Å, despite being shifted by approximately half a turn. The G helix is approximately two turns shorter in CYP175A1 (residues 176–198 in CYP175A1 compared with residues 197–226 in P450BM-3). Although the G helix is shorter, residues 176–195 in CYP175A1 align very well with residues 198–217 in P450BM-3 with an average r.m.s. deviation of 1.14 Å.

Based on the sequence identity and close structural homology in regions associated with the active site, it seemed possible that CYP175A1 may act as a fatty acid hydroxylase similar to P450BM-3. However, we did not observe the characteristic low to high spin shift that accompanies the binding of substrates to P450s using a range of even chain, commercially available fatty acids (C10–C20) either at room temperature or temperatures up to 70 °C. In addition, fatty acid hydroxylation could not be detected using H2O2 or bovine adrenodoxin/adrenodoxin reductase as reduct partners. A superimposition of the P450BM-3 structure with fatty acid bound (1FAG) on to CYP175A1 provides clues on why we failed to observe fatty acid binding (Fig. 6). The main problem is steric crowding owing to the close approach of Glu-67, Trp-269, and Ile-270 to the fatty acid substrate. The corresponding residues in P450BM-3 are, respectively, Ala-74, Pro-329, and Ala-330, which are all smaller side chains, thus providing additional space for the substrate to fit into the pocket. These close contacts involve the polar half of the fatty acid substrate near the substrate access channel, whereas the non-polar half of the substrate closer to the heme fits well into the CYP175A1 active site pocket. In addition, CYP175A1 also is missing Arg-47 and Tyr-51, both of which help to stabilize the fatty acid carboxyl group in P450BM-3. The corresponding residues in CYP175A1 are, respectively, Phe-41 and Leu-45 in CYP175A1. Those observations suggest that CYP175A1 may bind substrates considerably shorter than palmitoleic acid (16 carbons).

Comparison with CYP119.—The only other known thermostable P450 is CYP119 from the acidothermophilic organism S. solfataricus (2) with a Tm of 91 °C (3). At 368 amino acids, CYP119 is slightly smaller than CYP175A1. The r.m.s. backbone deviation is –1.7–1.8 Å. The overall (sequence and structure) similarity between the two thermophilic P450s is less than that between CYP175A1 and P450BM-3. Most of the increase in length of CYP175A1, relative to CYP119, resides at the N-terminal region, which typically has the lowest sequence identity between P450s. CYP175A1 has an extra A' helix, similar in conformation to a short 310 helix found in P450BM-3, which accounts for the most of the increase in length between the two thermophilic P450s. The CYP119 A helix starts at approximately residue 22 in CYP175A1. The largest structural differences between the two P450s reside in regions surround-
ing the active site: the F-G region, including the helices and all connecting loops, and the B′ helix. As noted earlier, the F/G region was found to be highly flexible between two crystal forms of CYP119, where the C terminus of the F helix unwraps by one full turn to dip down into the active site, depending on what ligand is bound at the active site. CYP119 is unusual because it contains a long loop preceding the B′ helix that enables it to adopt a position away from the active site toward the molecular surface. The B′ helix in CYP175A1 is in a similar conformation to P450BM-3 and most likely forms the lid of the substrate access channel.

**Structural Basis for Thermal Stability**—The enhanced thermal stability of proteins from extreme thermophiles has been attributed to a variety of global factors. Early on it was thought that amino acid composition was correlated with an increase in thermal stability (30). Later, general traffic rules of substitutions were loosely established (31–33). For example, an increase in the Arg content at the expense of Lys can aid in thermal stability because of an increase in hydrogen bonding possibilities, a decrease in chemical reactivity of the guanidino group, and resonance stabilization relative to Lys (34). Interestingly, CYP175A1 does show an increase in Arg content at 12.1% and a decrease in Lys (2.8%) relative to some of the mesophilic P450s such as P450BM-3 (Arg 4.4%, Lys 7.9%), whereas CYP119 shows a normal Arg and Lys content (7.6 and 7.1%, respectively). Another general traffic rule is the decrease in the number of polar-uncharged residues (Asn and Gln) and an increase in charged residues (Lys, His, Arg, Glu, and Asp) (34, 35). CYP175A1 shows a decrease in the polar uncharged residues (2.6%), whereas CYP119 shows only a slightly lower than normal content (6.3%) when

---

**Fig. 4.** Electrostatic surface potential of P450cam, CYP175A1, and P450BM-3. All three are in exactly the same orientation with the view along the proximal surface of the heme. The arrows indicate the approximate surface just above the Cys ligand and heme. The region indicated by the arrow is the closest approach of the heme to the molecular surface and, as such, provides the closest approach to the heme by the flavin or iron-sulfur cluster of a redox partner. The electrostatic and surface calculations were carried out with GRASP (transtor.bioc.columbia.edu/grasp/). The positive and negative electrostatic potentials are blue and red, respectively. The potentials are displayed at the ±10-kT levels.

**Fig. 5.** Overlay of substrate binding region in CYP175A1 and P450BM-3. CYP175A1 is depicted in dark gray, whereas P450BM-3 is depicted in light gray. The substrate binding regions of CYP175A1 and P450BM-3 align well, with an r.m.s. deviation less than 1.66 Å. Figure was made with MOLSCRIPT (45) and rendered in Raster3D (46).

**Fig. 6.** Computer graphics modeling of palmitoleic acid in the active site of CYP175A1. Palmitoleic acid was docked into the active site of CYP175A1 by superimposition of the backbones of CYP175A1 and the P450BM-3-palmitoleic acid complex (1FAG). The model indicates that fatty acid binding will most likely be blocked by Gln-67. The potential substrate should be shorter than palmitoleic acid and polar on one end.
compared with a mesophile such as P450BM-3 (9.0%). CYP175A1 also has a very large number of hydrophobic residues such as Ala and Leu (11.5 and 18.0%, respectively), but a normal hydrophobic content of 33.9%. Ala residues were first thought to increase thermal stability (31) because Ala is the best helical forming residue. However, in recent reviews based on genome sequencing projects researchers have noticed a slight decrease in the number of Ala residues (34). Another factor noted was an increase in the number of charged residues in thermophiles (35). However, the percentages of charged residues in both CYP119 and CYP175A1 and the mesophilic P450s remain fairly constant at 30%. As the database grows, a common recurring theme in thermophilic proteins is the potential importance of salt bridges (36), especially salt bridge networks involving more than 2 charged residues. It therefore is not too surprising that both thermophilic P450s have a greater number of such networks compared with their mesophilic counterparts. As shown in Table III, CYP175A1 has eight salt bridge networks whereas CYP119 has 10. By comparison the bacterial P450s listed in

### Table III

| Protein          | No. of 2-residue salt bridges | Salt bridge networks                                      | Residues in networks |
|------------------|------------------------------|----------------------------------------------------------|----------------------|
| CYP175A1         | 8                            | R29/R301/E308, R73/R163/E180, R77/R210/D82/E214, R3/R33/D48/D284 | 61.9                 |
| CYP119 imidazole | 12                           | H42/R362/D42, R154/D149/E212, R188/D85/E184, R233/E237/D301, R249/R302/E246, H307/R324/E245, R359/D345/E347, R9/R287/E296/E289/K10 | 59.3                 |
| CYP119 4-Phe-imadazole | 8                          | R9/R287/E296, K10/K11/E7, R49/R65/D61, R233/E237/D301, R249/R302/E246, H307/R324/E245, K91/R188/D85/E184, K100/K336/E96/E332/E333 | 68                   |
| P450BM-3         | 17                           | R56/E38/E344, K129/R147/D144, R203/E200/E207, R239/R375/R378/E329/D370, R290/R342/E287, H347/R364/E286, K266/H391/E262, R178/R186/D182/D251, R67/R330/D52/D328/E329 | 29.2                 |
| P450cam          | 18                           | R239/R340/D326, R290/R342/E287, H347/R364/E286, K266/H391/E262, R178/R186/D182/D251, R67/R330/D52/D328/E329 | 36.8                 |
| P450eryF         | 19                           | K103/D220/D223, R203/D160/D210, R331/D328/D333, R354/H341/K358/E279, K190/E155/D188, H202/E206/E297, R326/E322/E323, R354/H408/E351, R374/E81/E85 | 29.6                 |
| 2C5              | 12                           | K190/E155/D188, H202/E206/E297, R326/E322/E323, R354/H408/E351, R374/E81/E85 | 33.3                 |
| CYP51            | 17                           | R32/E21/E36, R122/E118/E414, R192/E189/E196, K156/R312/D224/D227, R274/H275/H318/E271, R316/R369/R374/E313/E376, R42/2D/D336 | 39.3                 |
| P450terp         | 14                           | R42/D/D336, K56/D448/E350, R210/R211/E214, R309/R362/E306, R424/E150/E402, R127/R134/E130/E392/E393 | 37                   |
It is generally thought that salt bridges should not be stabilizing because the energetic cost of desolvation cannot be fully compensated by electrostatic interactions between side chains and/or between the salt bridge side chains and the rest of the protein. An obvious difference with proteins is that the charged side chains are fixed to a polypeptide backbone. Thus, the folding of the polypeptide chain, driven by hydrophobic factors, "forces" charged groups close to one another on the surface. Nevertheless, such "forced" interactions might well oppose folding because there remains the problem of balancing unfavorable desolvation with favorable electrostatic interactions. One property of salt bridges that, perhaps, has not been fully appreciated is the importance of geometry (37) and electrostatic interactions between the salt bridge residues and the remainder of the protein (38). A survey of 222 salt bridges (37) using electrostatic calculations indicates that good geometry can substantially increase the stability of salt bridges. In this case, the balance between desolvation versus electrostatic interactions tips in favor of electrostatics. In addition, the electrostatic interactions between salt bridge groups and the rest of the protein are significantly greater in thermophilic proteins than in their mesophilic counterparts (37, 38). This suggests that salt bridges alone do not fully account for enhanced thermal stability but that the precise positioning of the charged residues within the electrostatic context of the rest of the protein is critical.

To gain further insights into the energetics of salt bridge networks, we employed the method of Hendesch and Tidor (39) to compute the electrostatic free energy of each salt bridge network in CYP175A1 using DELPHI (40, 41). We found that none of the salt bridges in CYP175A1 are stabilizing (data not shown). One problem with such an analysis, however, is that these comparisons are based on room temperature calculations where the dielectric constant of water is 80. As the temperature increases, the dielectric constant of water decreases and, hence, the energetic cost of desolvation will decrease. The problem, however, is far more complex than a simple change in dielectric constant. Elcock (42) carried out a computational analysis of charge interactions at elevated temperatures and found that the desolvation penalty for charged side chains is substantially decreased at thermophile temperatures. This analysis was extended to a comparison of proteins from thermophiles and mesophiles, and it was found that the effect the electrostatic component of hydration free energy results in greater stability for the thermophiles at elevated temperatures (42). Therefore, the more salt bridges, the more stable the protein is at elevated temperatures. Based on this view, it appears that the primary factors controlling thermostability in CYP175A1 and also CYP119 are salt bridge networks and shorter loops between major elements of secondary structure.

An additional structural feature possibly relevant to enhanced stability unique to CYP119 is the presence of a large aromatic network extending -39 Å (C-α–C-α) down one side of the molecule that was correctly predicted in the homology models (3, 6) and further confirmed in the crystal structure (4). Mutagenesis work by Maves et al. (20) showed that disruption of this aromatic network by an F24S mutation lowered the thermal melting transition by -10.2 °C. Enhanced thermal stability owing to aromatic interactions also has been found in other thermophiles (34). Surprisingly, there was no large aromatic networks present in CYP175A1 despite the fact that the aromatic content was nearly the same at 10% in CYP175A1 and 9.8% in CYP119. The largest network found was -13 Å and was composed of residues 117, 118, 121, 234, 353, 354, and 355.

In summary, it appears that the two thermophilic P450s share some common structural features expected to contribute to enhanced thermal stability. Both are smaller than mesophilic P450s, which is especially notable in CYP175A1. CYP175A1 closely resembles P450BM-3 without the loops. In addition, both contain a significant increase in salt bridge networks compared with mesophilic P450s. Where the two thermostable P450s sharply differ is the extensive aromatic network found in CYP119, which is not present in CYP175A1 or any other P450 for which a structure is available. Although the function of both CYP119 and CYP175A1 remain unknown, there have been recent advances in using an engineered CYP119 to catalyze fatty acid hydroxylation (43). Thus, it may be possible to engineer thermostable P450s to catalyze reactions of interest using both directed evolution (44) and structure-based design.

Acknowledgments—We thank Drs. Rebecca Wade (EMBL, Heidelberg, Germany), Bruce Tidor (Massachusetts Institute of Technology, Cambridge, MA), and Ray Luo (University of California, Irvine, CA) for advice and help with electrostatic calculations.
REFERENCES

1. Tijet, N., Helvig, C., and Feyerisen, R. (2001) Gene (Amst.) 262, 189–198
2. Wright, R. L., Harris, K., Solow, R., White, R. H., and Kennelly, P. J. (1996) FEBS Lett. 394, 235–239
3. McLean, M. A., Maves, S. A., Weiss, K. E., Krepich, S., and Sligar, S. G. (1998) Biochem. Biophys. Res. Commun. 252, 166–172
4. Yano, J. K., Kao, L. S., Schaller, D. J., Li, H., Ortiz de Montellano, P. R., and Poulos, T. L. (2000) J. Biol. Chem. 275, 31086–31092
5. Park, S. Y., Yamane, K., Adachi, S., Shiryo, Y., Maves, S. A., Weiss, K. E., and Sligar, S. G. (2002) J. Inorg. Biochem. 91, 491–501
6. Chang, Y. T., and Loew, G. (2000) Biochemistry 39, 2484–2498
7. Zeikus, J. G., Vieille, C., and Savchenko, A. (1998) Extremophiles 2, 179–183
8. Onuma, T., and Sato, R. (1984) J. Biol. Chem. 259, 2370–2378
9. Matthews, B. W. (1968) J. Mol. Biol. 33, 491–497
10. Leslie, A. G. W. (1992) Joint CCP4 ESF–EAMCB Newsletter Protein Crystallogr. 26
11. Zbyszek Otwinowski, W. M. (1997) Methods Enzymol. 276, 307–326
12. Jorge Navanza, P. S. (1997) Methods Enzymol. 276, 581–594
13. Read, R. J. (2001) Acta Crystallogr. Sect. D Biol. Crystallogr. 57, 1373–1382
14. Terwilliger, T. C., and Berendzen, J. (1999) Acta Crystallogr. Sect. D Biol. Crystallogr. 55, 849–861
15. Fujinaga, M., and Read, R. J. (1987) J. Appl. Crystallogr. 20, 517–521
16. Cowsert, K. D. (1994) Joint CCP4 ESF–EAMCB Newsletter Protein Crystallogr. 34, 34–38
17. Jones, T., and Kjeldgaard, M. (1997) Methods Enzymol. 277, 173–208
18. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grzesiek, S., Argos, P., Read, R. J., Rice, L. M., Simonsen, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D Biol. Crystallogr. 54, 905–921
19. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1999) Acta Crystallogr. Sect. D Biol. Crystallogr. 55, 247–255
20. Maves, S. A., and Sligar, S. G. (2001) Protein Sci. 10, 161–168
21. Fujinaga, M., and Read, R. J. (1987) J. Appl. Crystallogr. 20, 517–521
22. Graham, S. E., and Peterson, J. A. (1999) Arch. Biochem. Biophys. 369, 24–29
23. Raag, R., and Poulos, T. L. (1989) Biochemistry 28, 7586–7592
24. Schlichting, I., Berendzen, J., Chu, K., Stock, A. M., Maves, S. A., Benson, D. E., Sweet, R. M., Ringo, D., Petsko, G. A., and Sligar, S. G. (2000) Science 287, 1615–1622
25. Li, H., and Poulos, T. L. (1997) Nat. Struct. Biol. 4, 140–146
26. Park, S. Y., Shimizu, H., Adachi, S., Nakagawa, A., Tanaka, I., Nakahara, K., Shoun, H., Hayashi, E., Nakamura, H., Izuka, T., and Shire, Y. (1997) Nat. Struct. Biol. 4, 827–832
27. Sverniukova, I. F., Li, H., Zhang, H., Peterson, J. A., and Poulos, T. L. (1999) Proc. Natl. Acad. Sci. U.S.A. 96, 1863–1868
28. Pechapsky, T., Lyons, T. A., Kaxanis, S., Arakaki, T., and Ratnaswamy, G. (1996) Biochimie (Paris) 78, 723–733
29. Gotz, O. (1992) J. Biol. Chem. 267, 83–90
30. Perutz, M. F., and Reidt, H. (1975) Nature 255, 256–259
31. Argos, P., Rosenman, M. G., Grau, U. M., Zuber, H., Frank, G., and Trattnig, J. D. (1979) Biochemistry 18, 5698–5703
32. Vogt, G., Woell, S., and Argos, P. (1997) J. Mol. Biol. 269, 631–643
33. Menendez-Arias, L., and Argos, P. (1989) J. Mol. Biol. 206, 397–406
34. Vieille, C., and Zeikus, G. J. (2001) Microbiol. Mol. Biol. Rev. 65, 1–43
35. Cambillau, C., and Claverie, J. M. (2000) J. Biol. Chem. 275, 32383–32386
36. Kumar, S., and Nussinov, R. (2001) Cell. Mol. Life. Sci. 58, 1216–1233
37. Kumar, S., and Nussinov, R. (1999) J. Mol. Biol. 293, 1241–1255
38. Xiao, L., and Honig, B. (1999) J. Mol. Biol. 289, 1435–1444
39. Hendsch, Z. S., and Tidor, B. (1994) Protein Sci. 3, 211–226
40. Gilsen, M. K., and Honig, B. H. (1988) Proteins 4, 7–18
41. Gilsen, M. K., and Honig, B. H. (1988) J. Comput. Chem. 9, 327–335
42. Elcock, A. H. (1988) J. Mol. Biol. 204, 489–502
43. Koo, L. S., Immoos, C. E., Cohen, M. S., Farmer, P. J., and Ortiz de Montellano, P. R. (2002) J. Am. Chem. Soc. 124, 5684–5691
44. Joh, H., Lin, Z., and Arnold, F. H. (1999) Nature 399, 670–673
45. Kraulis, P. J. (1991) J. Appl. Crystallogr. 24, 946–950
46. Merritt, E. A., and Bacon, D. J. (1997) Methods Enzymol. 277, 505–524
47. Delano, W. L. (2002) PyMOL. 0.78 Ed., Delano Scientific, San Carlos, CA (www.delanoscientific.com)
48. McDonald, I. K., and Thornton, J. M. (1994) J. Mol. Biol. 238, 777–793
Preliminary Characterization and Crystal Structure of a Thermostable Cytochrome P450 from Thermus thermophilus
Jason K. Yano, Francesca Blasco, Huiying Li, Rolf D. Schmid, Anke Henne and Thomas L. Poulos

J. Biol. Chem. 2003, 278:608-616.
doi: 10.1074/jbc.M206568200 originally published online October 24, 2002

Access the most updated version of this article at doi: 10.1074/jbc.M206568200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 46 references, 7 of which can be accessed free at http://www.jbc.org/content/278/1/608.full.html#ref-list-1