Crystal Structure and Properties of CYP231A2 from the Thermoacidophilic Archaeon Picrophilus torridus † , ‡

ARTICLE *in* BIOCHEMISTRY · MARCH 2008

Impact Factor: 3.02 · DOI: 10.1021/bi702240k · Source: PubMed

5 AUTHORS, INCLUDING:

Huiying Li
University of California, Irvine
106 PUBLICATIONS **3,671** CITATIONS
SEE PROFILE

Paul R. Ortiz de Montellano
University of California, San Francisco
479 PUBLICATIONS **17,652** CITATIONS
SEE PROFILE

Thomas L Poulos
University of California, Irvine
164 PUBLICATIONS **4,542** CITATIONS
SEE PROFILE

Crystal Structure and Properties of CYP231A2 from the Thermoacidophilic Archaeon *Picrophilus torridus*[†,‡]

Winny W. Ho,[#] Huiying Li,[#] Clinton R. Nishida,[§] Paul R. Ortiz de Montellano,[§] and Thomas L. Poulos*[,#]

Departments of Molecular Biology and Biochemistry, Chemistry, and Pharmaceutical Sciences, University of California, Irvine, California 92697-3900, and Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94158-2517

Received November 9, 2007; Revised Manuscript Received December 14, 2007

ABSTRACT: The crystal structure of a cytochrome P450 from the thermoacidophile *Picrophilus torridus*, CYP231A2 (PTO1399), has been solved. This structure reveals a wide open substrate access channel. To better understand ligand-induced structural transitions in CYP231A2, protein−ligand interactions were investigated using 4-phenylimidazole. Comparison of the ligand-free and -bound CYP231A2 structures shows conformational changes where the F and G helices swing as a single rigid body about a pivot point at the N-terminal end of the F helix, allowing the F helix region to dip toward the heme, resulting in closer contacts with the ligand. Thermal melting data illustrate that the melting temperature for CYP231A2 increases nearly 10 °C upon ligand binding, thus illustrating that the closed conformation is substantially more stable. Furthermore, spectroscopic data indicate that the active site is stable at pH 4.5, although, unusually, the thiolate ligand to the iron can be reversibly protonated. CYP231A2 does not exhibit structural features normally associated with thermophilic proteins such as an increase in salt bridge networks or extensive aromatic clustering. The increase in thermal stability instead is best correlated with the smaller size and shorter loops in CYP231A2 compared to other P450s.

Cytochromes P450[1] are heme-thiolate enzymes that catalyze the following monooxygenation reaction:

$$NAD(P)H + H^+ + R\text{-}H + O_2 \rightarrow NAD(P)^+ + H_2O + R\text{-}OH$$

While P450s are present in many prokaryotes and eukaryotes, not until recently have P450s been found in thermophilic microorganisms. This is of considerable interest for the biotechnological applications of P450s in regio- and stereoselective hydroxylation reactions. If we can understand the structural basis for thermal stability in P450s, it might be possible to engineer such stability into a P450 that catalyzes a medically or industrially important reaction such as steroid hydroxylation or epoxidation reactions. The most straightforward approach in attempting to understand the structural basis of enhanced stability is to compare the same or a very similar enzyme from a mesophile and thermophile. This approach proved useful with the first thermal stable P450 to be discovered, CYP119 (*2*), and led to the conclusion that aromatic clustering is at least partially responsible for the enhanced thermal stability (*6*). CYP119 exhibits a melting temperature of ∼90 °C compared to ∼55 °C for the mesophilic bacterial $P450_{cam}$ (*5*, *6*). The thermal stability of 16 different families of proteins has been analyzed by Gromiha et al. (*7*), leading to the suggestion that the stability of proteins may be a result of better packing and solubility. In addition, other researchers have suggested that a variety of factors such as changes in loop flexibility (*8*), salt bridge interactions (*9*), and amino acid packing (*10*) can also play a role in increasing thermal stability in thermophilic proteins. Since there are several possible factors in increasing the thermal stability in proteins, it is imperative to use comparative analyses to determine specifically which of these factors play a part in the thermal stability of structurally related proteins.

A BLAST analysis of current available genomes of thermophilic organisms reveals the presence of two new thermophilic P450 enzymes in *Picrophilus torridus* (PTO), termed PTO1399 (CYP231A2) and PTO0085 (CYP232A2) (*11*, *12*). *Picrophilus torridus* is a thermoacidophilic euryarchaeon that thrives optimally at 60 °C and pH 0.7, but can grow even at pH 0, and is thus the most acidophilic organism known (*11*). While the intracellular pH of *S. acidocaldarius* from which CYP119 was isolated has been reported to be approximately 5.5 (*13*), the intracellular pH of *P. torridus* at 4.6 is even lower (*11*). Such a low intracellular pH in *P. torridus* provides a rare opportunity to study a P450 that can function *in vivo* at pH values lower than those found in other organisms.

[†] This work was supported by NIH grant GM32688 (T.L.P.) and GM25515 (P.O.M.).

[‡] Coordinates have been deposited in the Protein Data Bank under accession numbers 2RFB and 2RFC .

* Corresponding author. E-mail: poulos@uci.edu; fax: 949-824-3284; phone: 949-824-7020.

[#] University of California, Irvine.

[§] University of California, San Francisco.

[1] Abbreviations: P450, cytochromes P450; PTO1399, CYP231A2, a cytochrome P450 from a *Picrophilus torridus* strain; CYP119, cytochrome P450 from *Sulfolobus acidocaldarius*; P450cam, cytochrome P450 from *Pseudomonas putida*; P450BM3, cytochrome P450 from *Bacillus megaterium*; P450 2B4, xenobiotic metabolizing cytochrome P450; PEG 400, poly(ethylene glycol) 400; RMS, root mean square; PDB, Protein Data Bank www.rcsb.org/pdb.

10.1021/bi702240k CCC: $40.75
Published on Web 01/16/2008

Both CYP231A2 and CYP232A2 are similar in their survival at harsh environments and their sequence homology to CYP119, but they differ in their molecular weights of 39.56 and 44.36 kDa, respectively. Here we report cloning, expression, initial characterization, and crystal structure determination of CYP231A2 from *P. torridus*, and a comparison of the structure with CYP119 from *Sulfolobus acidocaldarius* (previously reported as from *Sulfolobus solfataricus*)[2] that clarifies the structural basis for thermal stability.

MATERIALS AND METHODS

CYP231A2 Cloning. A *P. torridus* freeze-dried culture (DSM 9790) was obtained from the DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH), reactivated, and grown according to the enclosed instructions using DSMZ medium 723 (*Sulfolobus* DSMZ medium 88 plus 2 g/L yeast extract, pH 1.0). Cells were grown in slanted unshaken tubes for 2−5 days at 55 °C. Genomic DNA was isolated (*14*), digested with *BamH*I, and used as template for PCR amplification of the CYP231A2 gene starting at Met10. The PCR product was subcloned into *pCWori* via *Nde*I and *Hind*III sites.

Expression and Purification of P450 (CYP231A2) from P. torridus. Transformation of competent cell XL-1 Blue (Stratagene) was carried out with pCWori/CYP231A2. A single colony was cultured in 50 mL of Luria−Bertani (LB) media containing 100 μg/mL of ampicillin at 37 °C overnight. Next, large cultures in 2xTY medium (1.5 L) were incubated with 5 mL of starter culture. Cells were grown at 37 °C to an OD_{600} of 0.8. The culture was shaken for 30 min at 30 °C and then induced with 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). The culture was grown for 20−40 h at 30 °C and harvested by centrifugation at 6000 rpm for 15 min. Purification steps were carried out at 4 °C. The cell pellet was lysed in lysis buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, and 0.5 mg/mL lysozyme. In addition, protease inhibitors (1 μg/mL antipain, 1 μM leupeptin, 1 μM pepstatin, and 0.1 mM PMSF) were added to the lysis buffer. The protein was purified by chromatography on Ni-NTA by binding and washing in the presence of 25 mM imidazole (50 mM Tris, pH 7.8, 500 mM NaCl) and elution with a linear gradient from 25 to 200 mM imidazole. Following dialysis against 50 mM Tris, pH 7.5, and 20 mM NaCl, the protein solution was passed through a Q Sepharose column, which retained contaminants. Final dialysis was against 50 mM HEPES (pH 7.5) and 200 mM NaCl. The protein purity was analyzed by 12.5% SDS-PAGE.

Anaerobic UV-Visible Spectrophotometry. Spectra were recorded using a Cary 1/E UV−vis spectrophotometer (Varian, Inc., Palo Alto, CA) and an anaerobic cuvette (Adams and Chittenden, Berkeley, CA). Reagents were added via a sidearm after anaerobicity had been achieved using Ridox-treated argon.

Melting Temperatures. The thermal melting temperatures of CYP231A2 and CYP119 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 measurements were carried out in 50 mM Tris pH 7.8 and 20 mM NaCl.

Crystallization. Crystals of CYP231A2 belonging to the monoclinic space group $P2_1$ with unit cell dimensions a = 48.79 Å, b = 168.46 Å, c = 88.23 Å, β = 102.3° were grown by hanging-drop vapor diffusion at room temperature over a reservoir of 0.1 M HEPES pH 7.5, 2.0 M ammonium sulfate, and 2% poly(ethylene glycol) 400. Hanging drops consisted of 2 μL of 15 mg/mL CYP231A2 mixed with an equal volume of reservoir solution. CYP231A2 plate-like crystals with dimensions 325 × 225 μm grew within 1−2 weeks. Crystals of ligand-bound CYP231A2 also belonged to the monoclinic space group $P2_1$ but with different unit cell dimensions a = 48.52 Å, b = 141.01 Å, c = 101.03 Å, β = 90.3°. They were grown with the addition of 20 mM 4-phenylimidazole, by hanging-drop vapor diffusion at room temperature over a reservoir of 0.1 M HEPES pH 7.0, 0.1 M NaCl, and 30% poly(ethylene glycol) 4000. Hanging drops consisted of 2 μL of 10 mg/mL CYP231A2 mixed with an equal volume of reservoir solution. CYP231A2 crystals with dimensions 200−250 × 30−50 μm grew within 2 weeks at room temperature (22 °C). On the basis of the estimation of Matthew's coefficient (*15*), there are three molecules per asymmetric unit for the ligand-free CYP231A2 crystals and four molecules per asymmetric unit for CYP231A2 bound with 4-phenylimidazole. Cryoprotectant soaking consisted of a six-step transfer to artificial precipitant solution (0.1 M HEPES pH 7.5, 2.0 M ammonium sulfate, and 2% poly(ethylene glycol) 400) with increasing concentrations of glycerol from 5% to 30% for the ligand-free form of CYP231A2.

For the ligand-bound form of CYP231A2 with 4-phenylimidazole, cryoprotectant soaking consisted of a four-step transfer to artificial precipitant solution (0.1 M HEPES pH 7.0, 0.1 M NaCl, 30% poly(ethylene glycol) 4000, and 20 mM 4-phenylimidazole) with increasing concentrations of glycerol from 5% to 20%.

Data Collection. All data were collected using crystals flash cooled into liquid nitrogen. Both 2.5 Å data for CYP231A2 and 3.1 Å data for 4-phenylimidazole-bound CYP231A2 were collected on Stanford Synchrotron Radiation Laboratory (SSRL) beam line 1−5 with a Q4 CCD detector. Optimization of data collection was guided by the STRATEGY function of MOSFLM (*16*). All data were reduced using DENZO and SCALEPACK (*17*), and rejections were performed with SCALEPACK. The data collection, process, and refinement statistics are shown in Table 1.

Solution of the CYP231A2 Crystal Structure. Phases were derived from molecular replacement and multiwavelength

[2] In the course of a sequence comparison of the *P. torridus* P450s with other P450s, we discovered that the sequence reported for CYP119, previously identified as originating from *Sulfolobus solfataricus*, was not present in the genome of *S. solfataricus*. The exact sequence was found in the genome of *Sulfolobus acidocaldarius* (*1*). Samples of *solfataricus* were deposited by the group who initially reported on CYP119 (*2*). However, according to the website of the DSMZ (http://www.dsmz.de/microorganisms/html/strains/strain.dsm001616.html), "Comment: cultures of DSM 1616 supplied until April 1989 were contaminated with *Sulfolobus acidocaldarius*". The American Type Culture Collection (ATCC) website reports a history of cultures from Zillig (*3*, *4*) to DSMZ to ATCC. These contaminated samples might have been the source of DNA for the later cloning of CYP119 (*5*), which would explain the not unexpected, but incorrect, "confirmation" that CYP119 is from *S. solfataricus*. In fact, CYP119 is from *S. acidocaldarius*. Hereafter, we refer to CYP119 as being from *S. acidocaldarius*.

Table 1: Data Collection, Phasing, Refinement, and Model Statistics[a]

data set	native[b]	three-wavelength MAD[c]	two-wavelength MAD[c]	ligand-bound[d]
		Data Collection		
ligand	none	none	none	4-phenylimidazole
PDB accession code	2RFB			2RFC
wavelength (Å)	0.999	1.74	1.74	0.999
resolution (Å)	2.5	3.23	3.1	3.1
total observations	169 082	205 536	185 614	70 810
unique reflections	48 305	28 643	25 680	23 827
completeness (%)	98.1 (95.1)	99.4 (99.9)	99.8 (100)	95.4 (98.1)
R_{sym}[e]	0.05 (0.66)	0.12 (0.54)	0.11 (0.75)	0.12 (0.47)
mean $<I/\sigma>$	31.1 (2.0)	16.8 (2.16)	19.9 (2.0)	9.72 (1.93)
		Phasing Statistics		
number of sites	three Fe	three Fe	three Fe	four Fe
figure of merit[e]		0.26	0.22	
		Refinement Statistics		
total reflections	45 855			23 572
total number of atoms	8330			11 211
total number of waters	135			19
R_{factor} (R_{free})[f] (%)	21.1 (26.8)			22.9 (33.0)
		Model Statistics[g]		
RMS deviations				
bonds (Å)	0.011			0.009
angles (deg)	1.5			1.5

[a] The values in parentheses were obtained in the outermost resolution shell. [b] Native data set collected at SSRL beamline 1.5. [c] Anomalous data set collected at SSRL beamline 1.5. [d] Ligand-bound data set collected at SSRL beamline 1.5. [e] Figure of merit (FOM) is the mean of the cosine of error in the phase angle. FOM $= |F(hkl)_{best}|/|F(hkl)|$, where $F(hkl)_{best} = \sum_{hkl} P(\alpha)F_{hkl}(\alpha)/\sum_{hkl} P(\alpha)$. [f] $R_{free} = R$ factor calculated using 5% of the reflection data chosen randomly and set aside from the start of refinement. [g] Statistics for final model.

anomalous diffraction (MAD) from the heme iron. CYP231A2 has a 35% sequence identity to CYP119 (PDB accession number 1F4T (*6*)) from *Sulfolobus acidocaldarius*. Molecular replacement in the $P2_1$ space group was carried out with PHASER (*18*, *19*) using data to 2.5 Å and a polyalanine search model of CYP119 with residues 1−7 deleted. A log likelihood gain (LLG) value of 309.14 and a *Z*-score of 12.69 were obtained in PHASER for three solutions without packing clashes.

A three-wavelength MAD data set was collected at different X-ray energies (inflection point, peak, and remote) around the Fe−K absorption edge of 1.74 Å. A total of 180° of data were collected with inverse beam at each wavelength. CYP231A2 crystals decayed during the prolonged data collection, which resulted in a set of low quality 3.23 Å MAD data set. Another two-wavelength MAD data set (inflection and remote) was also collected to 3.1 Å, providing usable phase information. Both HKL2MAP (Shelx) (*20*) and SOLVE (*21*) were used to locate heavy atom position(s) and to refine phases. Three iron sites were located in the asymmetric unit, in agreement with the predicted three molecules per asymmetric unit from Matthew's coefficient calculation. The final *Z*-score reported by SOLVE was 28.2 with an overall figure of merit of 0.218. Electron density calculated from MAD data showed good secondary structures in only one out of three molecules. MAD phasing alone was not enough to produce interpretable electron density maps. The phase combination calculations were carried out using SIGMAA (*22*, *23*), comprising the model phases from the molecular replacement solutions and the two-wavelength MAD phases from the heme iron. The combined phases were further improved by density modification using DM (*24*, *25*), including solvent flattening, histogram matching, and 3-fold NCS averaging, resulting in better quality electron density maps.

Molecular Replacement for 4-Phenylimidazole-Bound CYP231A2. Molecular replacement using the ligand-free form of CYP231A2 as a search model was carried out with PHASER using data up to 3.1 Å. A log likelihood gain (LLG) value of 3871 and a *Z*-score of 46.9 were obtained in PHASER for four solutions without clashes. Maps were calculated using CCP4 (FFT) (*22*).

Model Building and Refinement. Electron density map fitting beginning from the truncated polyalanine CYP119 model was carried out with the graphic modeling package O (*26*). Several sections of the electron density map showed clear identity of CYP231A2 side chains, which were modeled first. Regions of the molecule that lacked clear backbone electron density were deleted prior to further refinement with CNS (*27*). Electron density for these regions were calculated with two different programs, BUSTER (*28*) and CNS. Subsequent rounds of model building and refinement with simulated annealing starting at 3000K were carried out using the maximum likelihood-based approach implemented within CNS on data to the highest resolution. In the first 15 rounds of refinement the entire chain of one molecule was traced while noncrystallographic symmetry (NCS) operators were used to generate the second and third molecules in the asymmetric unit. In the last three rounds of refinement, the NCS restraints were imposed on main chain atoms only and side chains in each subunit were checked separately in O. This process of model building and refinement continued until the crystallographic *R*-factor was approximately 30% after which simulated annealing was replaced by minimization followed by isotropic *B*-factor refinement. When the *R*-factor reached about 25%, water molecules were automatically picked at a level of 3.5σ in an $F_o - F_c$ difference map using CNS, followed by visual inspection and manual adjustment. The final ligand-free model consists of all the residues except for the three missing residues in the B′ region.

A total of 135 water molecules and four sulfate molecules was also included in the final model with a crystallographic *R*-factor of 21.1% and free *R*-factor of 26.8%.

For the 4-phenylimizadole-bound structure, electron density map fitting beginning from the ligand-free CYP231A2 model was carried out with O and refined with CNS. The noncrystallographic symmetry (NCS) operators were used to generate the second, the third, and the fourth molecule in the asymmetric unit from the first molecule built in O. NCS restraints were imposed on main chain atoms throughout refinement. Owing to conformational changes caused by the binding of 4-phenylimidazole to CYP231A2, regions of the protein, composed of the F and G helices, were extensively remodeled. Model building and refinement in these regions continued until the final *R*-factor was 22.9% (*R*-free = 33.0%). The final 4-phenylimidazole-bound model contains 19 water molecules.

For both ligand-free and ligand-bound CYP231A2 structures, none of the main-chain torsion angles were in the disallowed regions of a Ramachandran plot. Relevant statistics for both crystal structures of CYP231A2 are given in Table 1.

RESULTS AND DISCUSSION

Spectroscopic Characterization of the CYP231A2 Active Site. Because of its low intracellular pH, *P. torridus* provides a system to study a P450 that has adapted to tolerate lower pH than mesophilic P450s. One of the most sensitive probes of P450 active site integrity is the absorption maximum of the protein in the ferrous carbonmonoxy (Fe^{II}-CO) state, which appears at approximately 450 nm in the intact, active protein and at 420 nm in the denatured, inactive state. We have therefore investigated the formation of the ferrous CO complex of CYP231A2 at different pH values and compared its behavior to that of $P450_{cam}$, a prototypical mesophilic P450 enzyme.

Ferrous CO formation was monitored at pH of 6.4, 5.4, and 4.5 for both CYP231A2 and $P450_{cam}$. There was little change in the absorbance spectrum of CYP231A2 after the addition of sodium dithionite at pH 6.4 (Figure 1A), but upon introduction of CO, the characteristic P450 species with an absorption maximum at 450 nm was rapidly formed (<2 min) along with approximately 25% P420 species (Figure 1B). Exposure of the sample to air resulted in rapid (<1 min) regeneration of the original ferric spectrum (Figure 1C). Surprisingly, $P450_{cam}$ exhibited higher stability than CYP231A2 not only at pH 6.4 but also at pH 5.4, in that the P450 species dominated the spectrum even after a 60 min incubation (not shown). However, at pH 4.5 $P450_{cam}$ was unstable, and we observed an immediate decrease in the Soret absorbance at 417 nm, an increase in a broad absorbance at 380 nm and precipitation of the protein. Reduction of the $P450_{cam}$ protein at this pH with dithionite yielded exclusively a 420 nm species while precipitation of the protein continued. In contrast, ferric CYP231A2 was spectroscopically stable at pH 4.5, and dithionite addition led to about a 15% decrease in the Soret absorbance, similar to that observed at pH 6.4 (Figure 2A). Nevertheless, addition of CO yielded primarily a 420 nm species (Figure 2B). The 420 nm absorbance is commonly associated with a loss of thiolate coordination due to protonation of the proximal thiolate, a mechanism that might be facilitated at low pH. However, the spectroscopic stability of the ferric form of CYP231A2 at pH 4.5 suggested that the heme environment was not significantly altered, and the heme remained bound in its normal position. Indeed, when sodium hydroxide was added to adjust the pH to the range of 6−7, the spectrum changed within the first scan from one containing almost exclusively P420 to one having a majority P450 component (Figure 2C). The ratio of P450/P420 continued to increase, and within 6−10 min, the spectrum stabilized. The final ratio was only modestly less than that seen when reduction and CO binding were done at pH 6.4 (Figure 1B), approximately 60% P450 at pH 4.5 compared to approximately 75% at pH 6.4. The protein did not detectably precipitate over the course of the experiment.

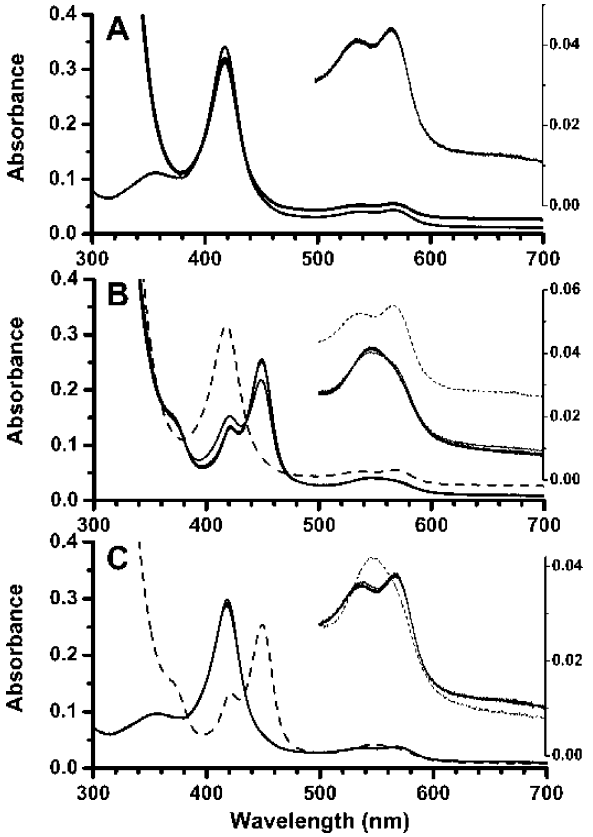

FIGURE 1: Formation of the CYP231A2 ferrous−CO complex at pH 6.4. (A) Addition of dithionite to the ferric enzyme. (B) Exposure of the reduced protein to CO. The dashed spectrum is that of the protein before exposure to CO. (C) Exposure of the ferrous-CO complex to air. The dashed spectrum is that of the protein before exposure to air.

Similar data were obtained in a similar set of experiments carried out at pH 5.4 (not shown). Clearly CYP231A2 not only has much higher stability at pH 4.5 than $P450_{cam}$ but it is able to revert to the P450 (thiolate) form from the P420 (putative thiol) form when the solution is neutralized. These observations are consistent with (1) protonation at pH 4.5 of the proximal thiolate ligand to form P420, (2) resistance of the protein structure to significant pH-dependent deformation, and (3) reversible protonation−deprotonation of the proximal thiol ligand.

These results raise some interesting questions with regard to the function of CYP231A2 at the physiologically relevant intracellular pH of *P. torridus* of 4.6 (*11*). At this pH we would expect the predominant species to be inactive P420.

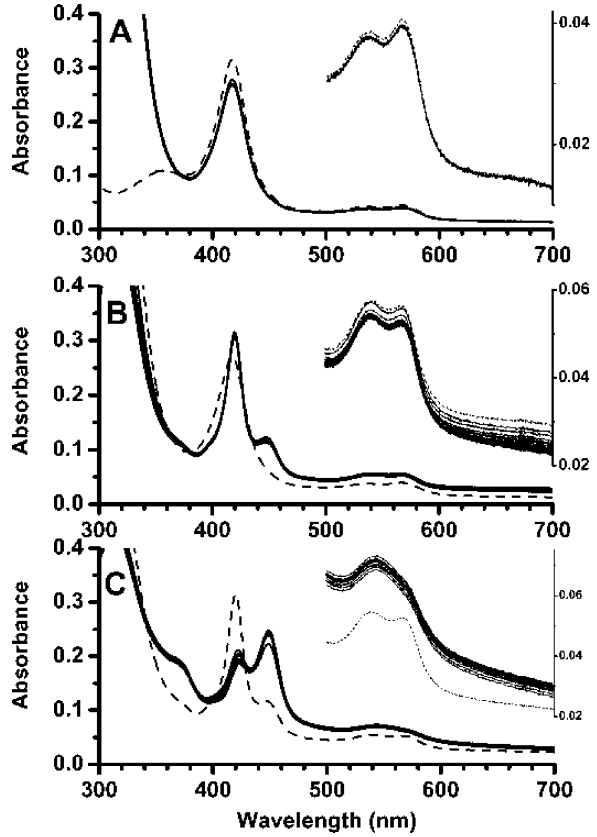

FIGURE 2: Formation of the CYP231A2 ferrous−CO complex at pH 4.5. Spectroscopic change(s) (A) over a period of 10 min after addition of dithionite to the ferric protein, (B) over a period of 30 min after further addition of CO, where the dashed spectrum is that of the protein before exposure to CO, and (C) over a period of 26 min after addition of NaOH, where the dashed spectrum is that of the protein before addition of NaOH.

However, our studies were carried out with substrate-free CYP231A2, and it is well-known that the substrate complex of P450s often stabilize the P450 species at the expense of the inactive P420 form. Indeed, in the case of P450epoK, it has been shown that the P420 form can revert to the active P450 form upon substrate binding (*29*). Unfortunately we do not as yet know the substrate for CYP231A2 so we are unable to directly test if substrate stabilizes the P450 form.

Overall Structure. With 352 residues, CYP231A2 is significantly smaller than CYP119 (368 residues) and other known mesophilic P450s like $P450_{cam}$ (414 residues) and $P450_{eryF}$ (403 residues). This structure retains the typical fold as observed in all P450 enzymes (Figure 3). The overall topology and secondary structure assignments of CYP231A2 were based on $P450_{cam}$ from *Pseudomonas putida* (*30*). A total of 12 helices are present throughout and account for 45% of the residues while five antiparallel β strands comprise about 13% of the structure. P450s normally have an N-terminal A helix, but this helix is absent in CYP231A2 because the terminal nine amino acids were truncated in the cloning procedure. The N-terminus of CYP231A2 therefore starts at β1, although the first visible residue in all three molecules in the asymmetric unit is Leu 12. It is important to note that the B′ helix region has some residues not visible in electron density maps. Residues 50−52 are not defined in molecule A while in molecules B and C density is missing for residues 44−51. One can anticipate that upon substrate binding these regions become more well ordered, although we cannot rule out an effect of the missing A helix on the order of the B′ helix.

The heme prosthetic group is embedded between the I and L helices with Cys 304 as the proximal axial thiolate ligand to the heme iron. The propionate side chains of the heme are H-bonded to the protein through residues Arg 68, Arg 247, and His 302. In the ligand-free state most P450s are low-spin with a water coordinated to the heme iron axial coordination position trans to the Cys ligand. However, in one of the three molecules in the asymmetric unit there is a large lobe of electron density indicating that some unknown ligand is bound to the iron (Figure 3B). Electron density is present in the other two molecules as well but is much weaker. Although imidazole was used in purification and is the most likely candidate for the ligand, the size and shape of the electron density envelope suggests a larger ligand. For the present the identity of this ligand remains unknown. Like other P450s, CYP231A2 has a β-sheet-rich domain on one side and a helix-rich domain on the other side, as seen in Figure 3. Between these domains lies a cleft that provides the access from the protein surface to the distal side of heme. In most of the known P450 structures, this cleft is closed by direct contacts between the two domains often involving contacts between the F/G loop and β regions. In a few P450s such as 2B4 (*31*) and P450nor (*32*) this cleft is wide open. This is the case for CYP231A2 as well. As observed with other P450s, it is very likely that the F/G helices and F/G loop undergo substantial movement when substrates/inhibitors bind resulting in a closing down of the active site entry channel.

Structure of Ligand-Bound Complex. Conformational flexibility in cytochromes P450 is commonly seen when substrates enter and leave the active site. Two of the best examples include P450BM3 (*33*) and P4502B4 (*31*, *34*). In both of these P450s the F and G helices shift over the I helix, resulting in a reshaping of the active site. In P4502B4 the changes are quite dramatic. Given that the ligand-free form of CYP231A2 has a wide open entry pocket, it is reasonable to expect a rather large change in structure upon ligand binding. We therefore set out to crystallize CYP231A2 in the presence of various ligands. We first screened a series of known P450 inhibitors and potential substrates for the ability to induce spectral shifts. The substrate(s) for CYP231A2 are not known although the close homology with CYP119 and P450BM3 indicates a potential fatty acid hydroxylase. Several fatty acids were tested for their ability to induce the low-to-high-spin shift in the visible absorption spectrum that often accompanies substrate binding. However, at both room and elevated temperatures we observed no spectral changes. α-Pinene, *trans*,*trans*-farnesol, erythromycin, protocatechuic acid, and benzoic acid produced spectral spin shifts, but the net change after addition was only 1−3% of the overall absorbance of the heme. In contrast, various imidazole-containing ligands, including large azoles such as etaconazole, fenbuconazole, and flusilazole, do bind well, and we were successful in obtaining crystals of the 4-phenylimidazole complex.

Figure 4 shows the composite omit $2F_o - F_c$ map around the ligand and in the vicinity of the F and G helices. Despite the limited resolution of the data, these key regions have well-defined electron density. As expected, the binding of 4-phenylimidazole induces striking conformational changes

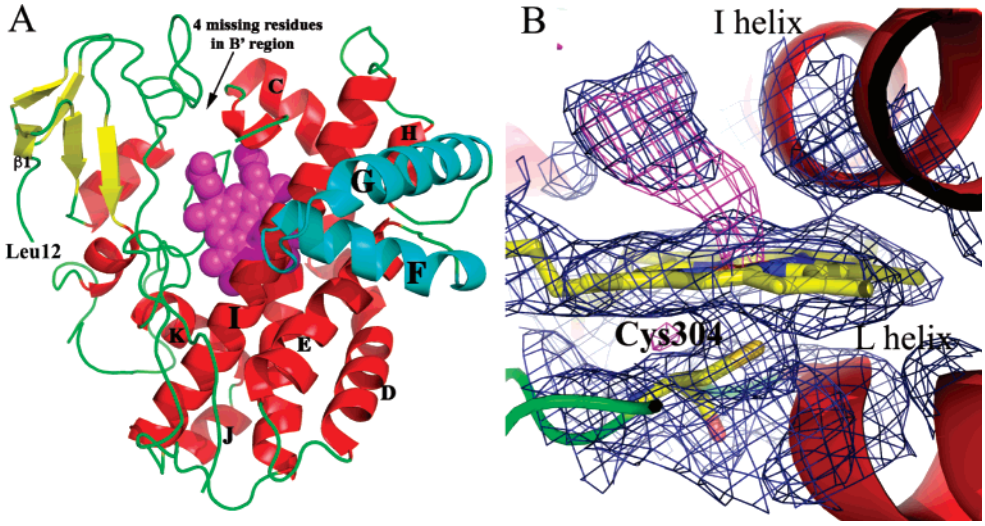

FIGURE 3: (A) Crystal structure of ligand-free CYP231A2. Helices are labeled according to the accepted scheme derived from P450cam. Note that there is no A helix and the N-terminus is just before β1. (B) $2F_o - F_c$ map contoured at 1.0σ (blue) and the $F_o - F_c$ map (magenta) contoured at 3.0σ. In molecule A of the asymmetric unit electron density above the heme iron indicates an unidentified ligand bound in the active site.

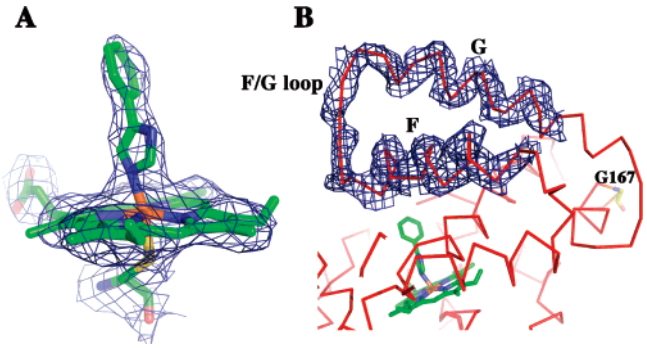

FIGURE 4: $2F_o - F_c$ composite omit maps contoured at 1.0σ for the 4-phenylimidazole complex. (A) Electron density for the 4-phenylimidazole, heme, and thiolate ligand. (B) Electron density for the F/G helical region. For clarity only the Cα model is shown.

in the regions encompassing the F and G helices (Figure 5). Upon ligand binding, both the F and G helices rotate as a single rigid body by ≈26° about a pivot point in the loop region between E and F helices centered on Gly 167 (Figure 4). This swing angle is measured starting from Ser 125 (ligand free) to Leu 120 (ligand free) and then to Ser 125 (ligand bound) at the beginning of the F helix. Even though the F and G helices and the F/G loop move to close the entry channel, these elements of structure all move as a unit with very little change in overall secondary structure. This is very similar to P450BM3 and P4502B4. The conformational change also results in a shift of the F and G helices along the helical axes. The F helix shifts by an entire helical turn such that residue *i* in the ligand-bound form occupies the position of residue *i*+4 in the ligand-free form. The shift in the G helix is smaller and amounts to about half a turn of helix.

Comparison with CYP119. As noted, CYP119 is the closest homologue to CYP231A2. In addition, CYP119 structures have been solved in the ligand-free form (*35*) and with 4-phenylimidazole bound (*6*) exactly as with CYP231A2 and thus is directly comparable with our current CYP231A2 structures. Although both P450s close down around the ligand, the structural changes that result from ligand binding are quite different. As noted above, in CYP231A2 the F and G helices and the F/G loop move as a unit and experience a fairly large displacement similar to other P450s. In sharp contrast, the F and G helices do not move in CYP119, and closure around the ligand involves a large restructuring of the F/G loop. Unwinding in helical turns at the end of the F helix and beginning of the G helix takes place which enables the F/G loop to extend down into the active site toward the ligand yet the position of the helices does not change. Even though the F and G helices do not slide over the I helix as they do in CYP231A2, P450BM3, and P4502B4, restructuring of the F/G loops results in some very large movements. For example, Gly 158 in the F/G loop moves about 18 Å from the surface toward the active site in going from the ligand-free to phenylimidazole complex (Figure 6). The largest displacement in the CYP231A2 F/G loop is much smaller with Arg 133 displaced about 6 Å. However, when motion of the F and G helices is compared, the largest displacement in CYP119 is only about 1.6 Å compared to 4−5 Å in CYP231A2.

These differences in conformational dynamics are no doubt tied to how the F and G helices interact with other regions of the protein. In CYP231A2 the F helix shifts by one full helical turn, which requires changes in interhelical contacts. The primary contacts are with the I helix. The I helix runs along the distal surface of the heme and provides key groups required for O_2 activation. The I helix remains in place upon ligand binding, but since the F and G helices do move then inter-helical contacts between the I helix and F/G helices must change. A good example is Ile 138 in the F helix. In the ligand-free form, Ile 138 forms good inter-helical contacts with Ile 195 in the I helix. However, in the ligand-bound form this interaction is disrupted owing to a shift along the helical axis which positions Ile 138 closer to the active site where it contacts the ligand. Thus, a favorable inter-helical contact lost when ligand binds is compensated for by forming a protein−ligand contact.

This type of motion does not occur in CYP119 and instead the primary inter-helical contacts are retained in the ligand-free and -bound structures. A possible reason for this is more efficient packing between the I helix and F/G helices in

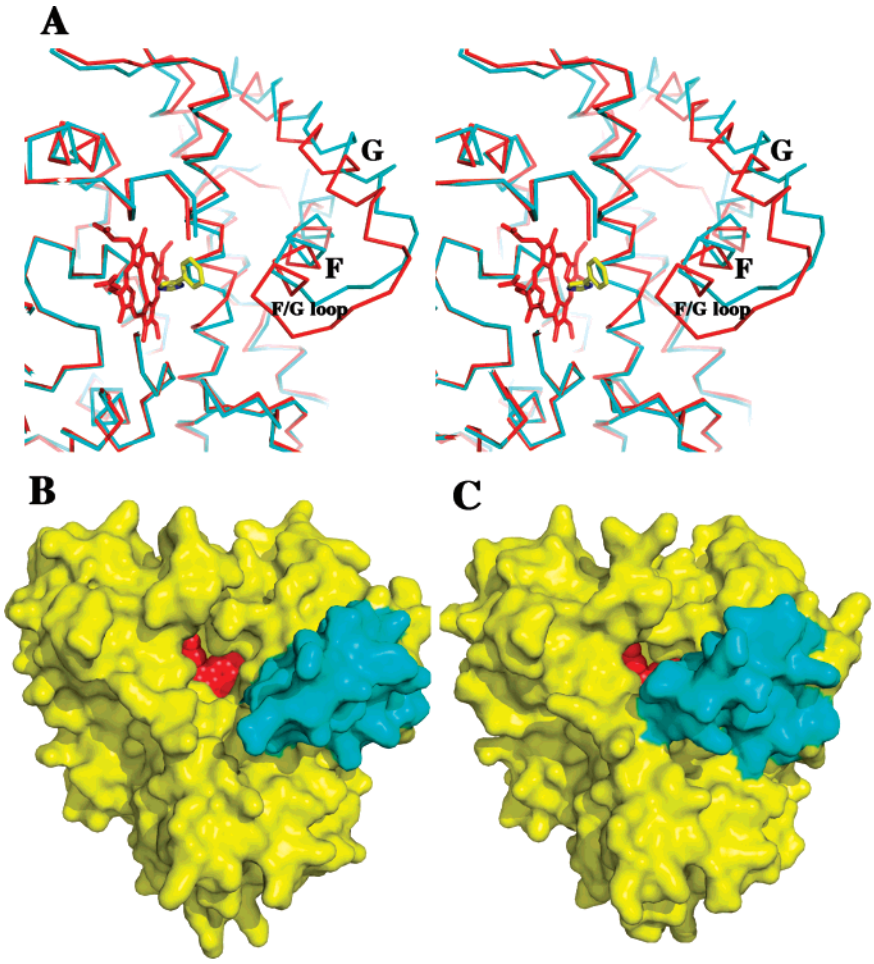

FIGURE 5: Models showing differences between the ligand-bound and -free forms in CYP231A2. (A) A backbone trace in the F/G helical region with the ligand-bound form in red and ligand-free in cyan. In the ligand-bound form the F and G helices slide over the I-helix resulting in movement of the F and G helices including the F/G loop closer to the ligand. Panels B and C are surface diagrams showing the ligand-free and -bound forms, respectively.

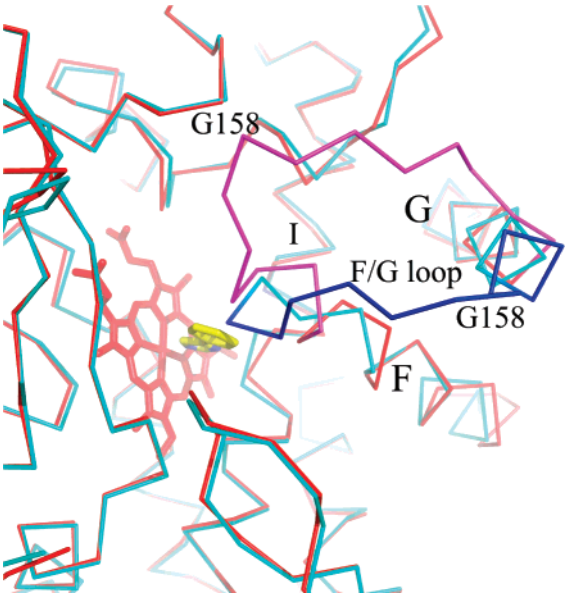

FIGURE 6: Superpostion of ligand-free (1IO8, cyan) and 4-phenylimidazole-bound (1F4T, red) CYP119. The F/G loop is in blue for the ligand-free structure and magenta for ligand-bound structure.

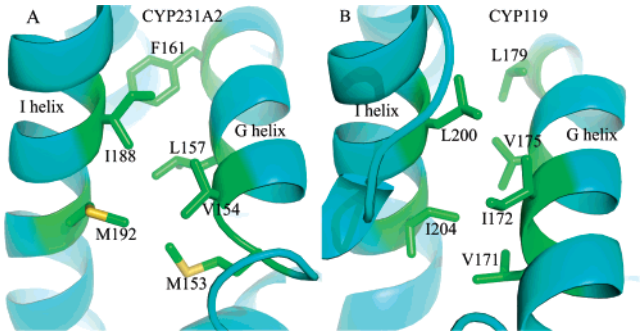

FIGURE 7: A comparison between (A) CYP231A2 and (B) CYP119 at the interface between the G and I helices. CYP119 has a more tightly packed interface composed of branched chain amino acids while in CYP231A2 the interface contains more flexible less tightly packed side chains.

CYP119 (Figure 7). Leu 200 and Ile 204 in the I helix cluster together with G helix residues Leu 171, Ile 172, Val 175, and Leu 179 to form a tight branched chain amino acid cluster. The corresponding I helix residues in CYP231A2 are Ile 188 and Met 192 in the I helix and Met 153, Val 154, Leu 157, and Phe 161 in the G helix. Particularly important are the two Met and Phe in CYP231A2 which do not pack as efficiently with branched chain amino acids (Figure 7). Thus, the G−I helical interface is more "fluid" in CYP231A2. The corresponding G−I helical interface also is not packed as tightly in P450BM3. In fact, there are more polar contacts. For example Ile 188 and Val 154 are replaced by Arg 255 and Asn 213, respectively, in P450BM3. It thus

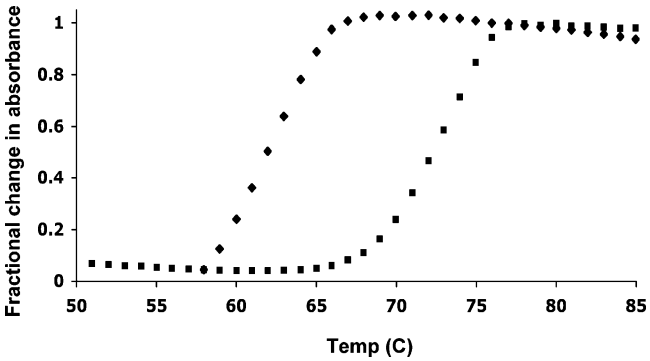

FIGURE 8: Thermal melting curves for ligand-free and ligand-bound CYP231A2. All measurements were carried out in 50 mM Tris pH 7.8 and 20 mM NaCl. There is no significant change in T_m at pH 4.0.

appears that compared to other P450s the F and G helices in CYP119 are locked in the closed form and the required motion for substrate entry and product release is controlled by movement of the long F/G loop. It is interesting to consider the possibility that the lack of relative helical motion in CYP119 as the result of more efficient and tighter packing is related to enhanced thermal stability. If the requirements of thermal stability necessitated a tightly folded and less conformationally dynamic P450, then this was compensated for by having a longer more mobile F/G loop that controls substrate entry and product release.

Structural Basis for Thermal Stability. Melting curves for the ligand-free and -bound forms of CYP231A2 are shown in Figure 8. CYP231A2 has a T_m value of 65 °C whereas the T_m value for CYP119 is 91 °C and that for $P450_{cam}$ is 55 °C. A T_m of 65 °C is inconsistent with the favored growth temperature of the organism, 60 °C. However, with 4-phenylimidazole bound the T_m goes up to 73 °C, demonstrating that the closed ligand-bound form is more thermally stable. Since the true substrate complex of CYP231A2 should also adopt the closed conformation, we can anticipate that the substrate complex will also be stable at the optimal growth temperature of the organism. Interestingly, the T_m does not change significantly at pH 4.0 which is closer to the reported internal cellular pH of 4.6 for *P. torridus* (*11*).

Several factors have been proposed to contribute to thermal stability, but for every "rule" there is often more than one exception. The one structural feature that does appear to correlate well with thermal stability is salt-bridged networks that contain more than two residues (*36*, *37*). CYP119 with 4-phenylimidazole bound has eight such networks, CYP175A1 from the thermophillic organism *Thermus thermophilus* also has eight, and several mesophillic P450s have four to six (*38*). CYP231A2 has only one such network. However, the use of salt bridges may not be a strategy for stabilization adopted by acidophiles since the internal pH of *P. torridus* is about 4.6 (*11*). At this pH there should be a significant fraction of carboxylate groups protonated. In addition we found that T_m is about the same at both pH 4 and 7, indicating that the degree of protonation of carboxylates is probably not important for thermal stability.

An additional structural feature that may be relevant to enhanced stability is the presence of an aromatic network extending down one region of the molecule, as in CYP119 (*5*, *6*, *38*). A disruption of this aromatic network has been shown to lower the stability of a protein by decreasing its thermal melting temperature (*39*, *40*) CYP231A2 has more aromatic networks but none as large as in CYP119. Overall, it appears that CYP231A2 does not have any of the characteristic structural features often associated with enhanced thermal stability other than being the smallest P450. Thus, the lack of hyperthermal stability as in CYP119 may be due to the lack of salt-bridged networks as well as extensive aromatic networks while the increase in thermal stability over mesophilic P450s like P450cam is due to its smaller size.

CONCLUSION

Similar to P450BM3 and P4502B4, a new thermophilic cytochrome P450 from *P. torridus* named CYP231A2 exhibits changes in structure upon ligand binding. Comparison of both ligand-free and -bound structures of CYP231A2 show a rearrangement of the F/G helical region as in other P450s. However, there is a clear difference between ligand-induced conformational changes in CYP119 and CYP231A2. In CYP231A2 as in other P450s the F and G helices and F/G loop move as a unit to close the active site entry channel. This movement requires large changes in inter-helical contacts between the F/G helices and the I-helix. In sharp contrast, closure of the active site around the ligand in CYP119 involves a dramatic restructuring of the F/G loop yet the F and G helices do not move. We hypothesize that this difference is due to a much more efficient and tighter packing between the F/G helices and other regions of the protein, especially the I-helix, in CYP119. This structural feature that limits motion may also be associated with the enhanced thermal stability of CYP119. A $T_m = 73$ °C for ligand-bound CYP231A2 is between $P450_{cam}$, $T_m = 55$ °C, and CYP119, $T_m = 93$ °C so CYP231A2 does not quite qualify as a hyperstable protein. CYP231A2 shares none of the expected characteristics associated with enhanced thermal stability such as salt-bridged networks or extensive aromatic interactions. CYP231A2 did exhibit stability against irreversible P420 formation at pH 4.6, at which $P450_{cam}$ precipitates. The requirement for acid stability may have given rise in CYP231A2 to a set of structural stabilizing forces that mostly exclude ion pairing. In fact, retention of ion pairs would be expected to destabilize the enzyme under conditions of low pH. The one feature unique to CYP231A2 that possibly contributes to moderate thermal stability is its smaller size. CYP231A2 is the smallest P450 whose structure has been solved including other thermophilic P450s. Thus, CYP231A2 looks to follow at least one property associated with enhanced stability: a decrease in surface to volume ratio owing to shorter loops (*36*, *41*) and the discarding of excess secondary structure that are not critical in forming and stabilizing the core P450 structure.

ACKNOWLEDGMENT

We thank Dr. Melanie Cocco for pointing out the effect of pH on salt bridges.

REFERENCES

1. Chen, L., Brügger, K., Skovgaard, M., Redder, P., She, Q., Torarinsson, E., Greve, B., Awayez, M., Zibat, A., Klenk, H.-P., and Garrett, R. A. (2005) The genome of *Sulfolobus acidocaldarius*, a model organism of the crenarchaeota, *J. Bacteriol. 187*, 4992−2999.

2. Wright, R. L., Harris, K., Solow, B., White, R. H., and Kennelly, P. J. (1996) Cloning of a potential cytochrome P450 from the archaeon *Sulfolobus solfataricus*, *FEBS Lett. 384*, 235−239.
3. Schleper, C., Pühler, G., Holz, I., Gambacorta, A., Janekovic, D., Santarius, U., Klenk, H.-P., and Zillig, W. (1996) *Picrophilus* gen. nov., fam. nov.: a novel aerobic, heterotrophic, thermoacidophilic genus and family comprising archaea capable of growth around pH 0, *J. Bacteriol. 177*, 7050−7059.
4. Schleper, C., Pühler, G., Klenk, H.-P., and Zillig, W. (1996) Picrophilus oshimae and *Picrophilus torridus* fam. nov., gen. nov., sp. nov., two species of hyperacidophilic, thermophilic, heterotrophic, aerobic archaea, *Int. J. Syst. Bacteriol. 46*, 814−816.
5. McLean, M. A., Maves, S. A., Weiss, K. E., Krepich, S., and Sligar, S. G. (1998) Characterization of a cytochrome P450 from the acidothermophilic archaea *Sulfolobus solfataricus*, *Biochem. Biophys. Res. Commun. 252*, 166−172.
6. Yano, J. K., Koo, L. S., Schuller, D. J., Li, H., Ortiz de Montellano, P. R., and Poulos, T. L. (2000) Crystal structure of a thermophilic cytochrome P450 from the archaeon *Sulfolobus solfataricus*, *J. Biol. Chem. 275*, 31086−31092.
7. Gromiha, M. M., Oobatake, M., and Sarai, A. (1999) Important amino acid properties for enhanced thermostability from mesophilic to thermophilic proteins, *Biophys. Chem. 82*, 51−67.
8. Russell, R. J., Hough, D. W., Danson, M. J., and Taylor, G. L. (1994) The crystal structure of citrate synthase from the thermophilic archaeon, Thermoplasma acidophilum, *Structure 2*, 1157−1167.
9. Perutz, M. F., and Raidt, H. (1975) Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2, *Nature 255*, 256−259.
10. Zeikus, J. G., Vieille, C., and Savchenko, A. (1998) Thermozymes: biotechnology and structure-function relationships, *Extremophiles 2*, 179−183.
11. Futterer, O., Angelov, A., Liesegang, H., Gottschalk, G., Schleper, C., Schepers, B., Dock, C., Antranikian, G., and Liebl, W. (2004) Genome sequence of *Picrophilus torridus* and its implications for life around pH 0, *Proc. Natl. Acad. Sci. U.S.A. 101*, 9091−9096.
12. Nishida, C. R., and Ortiz de Montellano, P. R. (2005) Thermophilic cytochrome P450 enzymes, *Biochem. Biophys. Res. Commun. 338*, 437−445.
13. Lübben, M., and Schäfer, G. (1989) Chemiosmotic energy conversion of the archaebacterial thermoacidophile *Sulfolobus acidocaldarius*: oxidative phosphorylation and the presence of an F0-related *N*,*N*′-dicyclohexylcarbodiimide-binding proteolipid, *J. Bacteriol. 171*, 6106−6116.
14. Wilson, K. (1997) Preparation of genomic DNA from bacteria, *Curr. Prot. Mol. Biol.*, 2.4.1−2.4.5.
15. Matthews, B. W. (1968) Solvent content of protein crystals, *J. Mol. Biol. 33*, 491−497.
16. Leslie, A. G. W. (1992) Recent changes to the MOSFLM package for processing film and image plate data, *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography 26*; CCP4, Daresbury, UK.
17. Zybszek Otwinowski, W. M. (1997) Processing of X-Ray Diffraction Data Collected in Oscillation Mode, *Methods Enzymol. 276*.
18. McCoy, A. J. (2007) Solving structures of protein complexes by molecular replacement with Phaser, *Acta Crystallogr., Sect. D.: Biol. Crystallogr. 63*, 32−41.
19. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) Likelihood-enhanced fast translation functions, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 61*, 458−464.
20. Pape, T., and Schneider, T. (2004) HKL2MAP: a general user interface fo phasing with SHELX program, *J. Appl. Crystallogr. 37*, 843−844.
21. Terwiliger, T. G., and Breendenzen, J. (1999) Automated MAD and MIR structure solution, *Acta Crystallogr. 55*, 849−861.
22. Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 59*, 1131−1137.
23. Read, R. J. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors, *Acta Crystallogr., Sect. A: Found. Crystallogr.* 140−149.
24. Cowtan, K. D., and Main, P. (1993) Improvement of macromolecular electron-density maps by the simultaneous application of real and reciprocal space constraints, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 49*, 148−157.
25. Cowtan, K. D., and Zhang, K. Y. (1999) Density modification for macromolecular phase improvement, *Prog. Biophys. Mol. Biol. 72*, 245−270.
26. Jones, T., and Kjeldgaard, M. (1997) Electron-density map interpretation, *Methods Enzymol. 277*, 173−208.
27. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 54*, 905−921.
28. Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S. M., and Bricogne, G. (2004) Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2210−2221.
29. Ogura, H., Nishida, C. R., Hoch, U. R., Perera, R., Dawson, J. H., and Ortiz de Montellano, P. R. (2004) EpoK, a Cytochrome P450 involved in biosynthesis of the anticancer agents epothilones A and B. substrate-mediated rescue of a P450 enzyme, *Biochemistry 43*, 14712−14721.
30. Poulos, T. L., Finzel, B. C., Gunsalus, I. C., Wagner, G. C., and Kraut, J. (1985) The 2.6-A crystal structure of Pseudomonas putida cytochrome P-450, *J. Biol. Chem. 260*, 16122−16130.
31. Scott, E. E., He, Y. A., Wester, M. R., White, M. A., Chin, C. C., Halpert, J. R., Johnson, E. F., and Stout, C. D. (2003) An open conformation of mammalian cytochrome P450 2B4 at 1.6-A resolution, *Proc. Natl. Acad. Sci. U.S.A. 100*, 13196−13201.
32. Park, S. Y., Shimizu, H., Adachi, S., Nakagawa, A., Tanaka, I., Nakahara, K., Shoun, H., Obayashi, E., Nakamura, H., Iizuka, T., and Shiro, Y. (1997) Crystal structure of nitric oxide reductase from denitrifying fungus Fusarium oxysporum, *Nat. Struct. Biol. 4*, 827−832.
33. Li, H., and Poulos, T. L. (1997) The structure of the cytochrome P450BM-3 haem domain complexed with the fatty acid, palmitoleic acid, *Nature Struct. Biol.*, 140−146.
34. Scott, E. E., White, M. M., He, Y. A., Johnson, E. F., Stout, C. D., and Halpert, J. R. (2004) Structure of mammalian cytochrome P450 2B4 complexed with 4-(4-chlorophenyl)imidazole at 1.9-Å resolution, *J. Biol. Chem. 279*, 27294−27301.
35. Park, S.-Y., Yamane, K., Adachia, S.-i., Shiro, Y., Weissb, K. E., Mavesb, S. A., and Sligar, S. G. (2002) Thermophilic cytochrome P450 (CYP119) from Sulfolobus solfataricus: high resolution structure and functional properties, *J. Inorg. Biochem. 91*, 491−501.
36. Kumar, S., Tsai, C. J., and Nussinov, R. (2000) Factors enhancing protein thermostability, *Protein Eng. 13*, 179−191.
37. Kumar, S., and Nussinov, R. (2001) How do thermophilic proteins deal with heat? *Cell Mol. Life. Sci. 58*, 1216−1233.
38. Yano, J. K., Blasco, F., Li, H., Schmid, R. D., Henne, A., and Poulos, T. L. (2003) Preliminary characterization and crystal structure of a thermostable cytochrome P450 from *Thermus thermophilus*, *J. Biol. Chem. 278*, 608−616.
39. Maves, S. A., and Sligar, S. G. (2001) Understanding thermostability in cytochrome P450 by combinatorial mutagenesis, *Protein Sci. 10*, 161−168.
40. Puchkaev, A. V., Koo, L. S., and Ortiz de Montellano, P. R. (2003) Aromatic stacking as a determinant of the thermal stability of CYP119 from *Sulfolobus solfataricus*, *Arch. Biochem. Biophys. 409*, 52−58.
41. Russell, R. J., Fergusom. J. M., Hough, D. W., Danson, M. J., and Taylor, G. L. (1997) The crystal structure of citrate synthase from the hyperthermophilic archaeon *Pyrococcus furiosus* at 1.9 Å resolution. *Biochemistry 36*, 9983−9994.

BI702240K