Solution Structure of Phenol Hydroxylase Protein Component P2 Determined by NMR Spectroscopy‡

Hong Qian,§ Ulf Edlund,§ Justin Powlowski,|| Victoria Shingler,⊥ and Ingmar Sethson*,§

Department of Organic Chemistry, Umeå University, S-901 87 Umeå, Sweden, Department of Chemistry and Biochemistry, Concordia University, Montreal, Quebec H3G 1M8, Canada, and Department of Cell and Molecular Biology, Umeå University, S-901 87 Umeå, Sweden

Received August 2, 1996; Revised Manuscript Received October 26, 1996⊗

ABSTRACT: Phenol hydroxylase from *Pseudomonas* sp. CF600 is a member of a family of binuclear iron-center-containing multicomponent oxygenases, which catalyzes the conversion of phenol and some of its methyl-substituted derivatives to catechol. In addition to a reductase component which transfers electrons from NADH, optimal turnover of the hydroxylase requires P2, a protein containing 90 amino acids which is readily resolved from the other components. The three-dimensional solution structure of P2 has been solved by 3D heteronuclear NMR spectroscopy. On the basis of 1206 experimental constraints, including 1060 distance constraints obtained from NOEs, 70 ϕ dihedral angle constraints, 42 ψ dihedral angle constraints, and 34 hydrogen bond constraints, a total of 12 converged structures were obtained. The atomic root mean square deviation for the 12 converged structures with respect to the mean coordinates is 2.48 Å for the backbone atoms and 3.85 Å for all the heavy atoms. This relatively large uncertainty can be ascribed to conformational flexibility and exchange. The molecular structure of P2 is composed of three helices, six antiparallel β-strands, one β-hairpin, and some less ordered regions. This is the first structure among the known multicomponent oxygenases. On the basis of the three-dimensional structure of P2, sequence comparisons with similar proteins from other multicomponent oxygenases suggested that all of these proteins may have a conserved structure in the core regions.

The insertion of molecular oxygen into organic molecules is a reaction catalyzed by a diverse array of mono- and dioxygenases. Microorganisms that possess the ability to grow at the expense of aliphatic and aromatic compounds are particularly adept at synthesizing these oxygenases, most of which are involved in the initial metabolic reactions used to degrade the growth substrate. While these enzymes are diverse, without exception they all possess either an organic prosthetic group, such as a flavin or pterin derivative, or a transition metal which is responsible for activation of oxygen before insertion into the substrate. These enzymes are further distinguished according to whether a mono- or dioxygenation is catalyzed and whether or not the enzyme requires more than a single polypeptide component. In addition to the oxygenase itself, which may contain one or more polypeptides, auxiliary polypeptide components typically contain redox centers responsible for the transfer of electrons from the physiological electron donor, NAD(P)H, to the oxygenating component [reviewed in Mason and Cammack (1992) and Powlowski and Shingler (1994)].

Recently, a multicomponent oxygenase from *Pseudomonas* sp. strain CF600, which catalyzes the conversion of phenol and some of its methyl-substituted derivatives to (methyl) catechol, has been described. The enzyme requires a reductase component ("P5"), containing FAD and a ferredoxin-type [2Fe-2S] center, for transfer of electrons from NADH (Powlowski & Shingler, 1990); a hydroxylase component comprised of three polypeptides containing a binuclear iron center; and a small protein ("P2") lacking any redox cofactors and required for optimal turnover of the enzyme (Nordlund et al., 1990; Powlowski & Shingler 1994; Powlowski et al., manuscript in preparation). Amino acid sequences of the polypeptides comprising this enzyme are identical or highly similar to sequences of polypeptides from phenol hydroxylases of two other *Pseudomonas* isolates (Ng et al., 1994; Herrmann et al., 1995) and of *Acinetobacter calcoaceticus* (Ehrt et al., 1995). Nucleotide sequences have also been reported for genes encoding similar polypeptide components of toluene monooxygenases (Yen et al., 1991; Byrne et al., 1995; Johnson & Olsen, 1995), which generally exhibit much lower sequence similarities at the protein level.

Each of the phenol hydroxylase components also shows similarities with corresponding components of methane monooxygenase. Thus, *mmoC* and *dmpP* encode the respective reductases, *mmoXYZ* and *dmpLNO* encode the hydroxylase components, and *mmoB* and *dmpM* encode low molecular weight proteins associated with the hydroxylase. Extensive biophysical studies of the binuclear iron centers of two methane monooxygenases have been carried out, and a crystal structure of the hydroxylase component of methane monooxygenase from *Methanobacterium capsulatus* has been published [reviewed by Rosenzweig and Lippard (1994) and Lipscomb (1994)]. However, structures of the methane monooxygenase reductase and MmoB are currently lacking.

The functions of the low molecular weight proteins of methane monooxygenase (MmoB) and the phenol (P2) and

‡ The coordinates and constraints of the phenol hydroxylase protein component P2 have been deposited in the Brookhaven Protein Data Bank. The PDB ID code is 1HQI and the constraints code is R1HQIMR.

* To whom correspondence should be addressed.
§ Department of Organic Chemistry, Umeå University.
|| Concordia University.
⊥ Department of Cell and Molecular Biology, Umeå University.

⊗ Abstract published in *Advance ACS Abstracts,* December 15, 1996.

S0006-2960(96)01923-X CCC: $14.00

toluene monooxygenases are quite unique among multicomponent oxygenases. None of these proteins appears to be directly involved in redox reactions, but rather each stimulates its respective hydroxylase at optimal concentrations and inhibits at higher levels. Although P2 significantly increases product yield from the phenol hydroxylase in single-turnover experiments, the small proteins associated with the methane and toluene monooxygenases do not (Fox et al., 1989; Newman & Wackett 1995; Powlowski et al., manuscript in preparation). MmoB has also been shown to modify the regioselectivity of methane monooxygenase (Frolund et al., 1992), alter the environment of the binuclear iron center (Fox et al., 1991; Pulver et al., 1993), and shift the redox potential of the iron center in the MmoB-hydroxylase complex (Paulsen et al., 1994). Thus, binding of these small components to their respective hydroxylases at different stages of the catalytic cycle appears to play an important role in regulating the oxygenase activity.

This paper presents the initial structural characterization of one of these hydroxylase-associated proteins, P2 of phenol hydroxylase. The potential biochemical significance of its structural features is discussed.

MATERIALS AND METHODS

Plasmid Construction and P2 Overexpression. In order to express high levels of P2, the *dmpM* gene was placed under the control of the T7 promoter of the pET3a expression vector. All DNA manipulations were done according to standard techniques (Sambrook et al., 1989). The *dmpM* gene was first PCR amplified as an *Nde*I to *Bam*H1 fragment (bp 2075−2352; Nordlund et al., 1990) and cloned into a pBluescript (Stratagene) derivative, in which an *Nde*I site had previously been introduced into the polylinker, to generate pVI206. Both strands of the resulting fragment were sequenced to ensure that no mutations had been introduced. The *Nde*I to *Bam*H1 fragment of pVI206 was subsequently cloned between these sites of pET3a (Rosenburg et al., 1987) to generate pVI207 with *dmpM*.

Recombinant P2 was expressed by introducing pVI207 into *Escherichia coli* BL21 (DE3) (Rosenburg et al., 1987). In a standard expression experiment, fresh transformants were used to inoculate M9-glucose medium (Sambrook et al., 1989) containing ampicillin (50 μg/mL), and the culture was grown at 37 °C with vigorous shaking to $A_{650} = 0.4-0.6$. Cultures were then supplemented with IPTG[1] (0.5 mM) to induce expression of the host-encoded T7 RNA polymerase that is under the control of the P_{tac} promoter. After a further 4 h of growth cells were harvested by centrifugation and washed, and then the cell paste was stored at −80 °C until used. For ^{15}N- and ^{13}C-labeled P2 preparations, a similar procedure was followed except that the M9-glucose medium contained [^{13}C]glucose and/or $^{15}NH_4Cl$.

Purification of P2. The buffer used during the purification was 50 mM Tris-HCl, pH 8.0, containing 10% glycerol and 1 mM DTT (buffer A). Chromatography media were purchased from Pharmacia, and all purification steps were carried out at 4 °C. The presence of P2 in fractions was monitored by SDS−PAGE, and protein concentrations were estimated using the BCA assay (Pierce Chemical Co.) with bovine serum albumin as standard.

Crude extracts of *E. coli* BL21 (DE3) pVI207 (P2) were prepared by sonication of cell suspensions in buffer A supplemented with 1 mM protease inhibitor Pefabloc (Boehringer Mannheim). The supernatant obtained after centrifugation (100000*g* for 1 h) was applied to a DEAE-Sepharose Fast-Flow column (34 × 2.6 cm) equilibrated with buffer A. The column was washed with this buffer (300 mL) and then eluted with a 0−0.3 M NaCl gradient (1.5 L) at a flow rate of 6 mL/min. Fractions containing P2, which eluted near the end of the gradient, were placed in Spectrapor dialysis tubing (6000−8000 molecular weight cutoff) for concentration on a bed of poly(ethylene glycol) (PEG) (average molecular weight 35 000). The concentrated sample was then loaded onto a Sephacryl S-300 HR column (91 × 2.6 cm) equilibrated with buffer A containing 0.5 M NaCl. The column was eluted with the same buffer at a flow rate of 1.5 mL/min: 6 mL fractions were collected. The P2-containing fractions were combined and concentrated using the PEG dialysis procedure. The purified protein was extensively dialyzed against 1 mM sodium−potassium phosphate buffer, pH 7.2, and then lyophilized.

Sample Preparation for NMR. All protein samples were dissolved in 95% H_2O/5% D_2O or alternatively pure D_2O. The pH was adjusted to 7.5 by using 20 mM phosphate buffer, and 0.5 mM NaN_3 was added to inhibit bacterial growth. The concentration of unlabeled and ^{15}N-labeled samples was 2.0 mM while the concentration of the ^{13}C/^{15}N-labeled sample was 0.6 mM.

NMR Spectroscopy. NMR spectra of P2 were acquired either using a Bruker AMX2-500 spectrometer or a Varian UnityPlus 500 spectrometer. All spectra were recorded at 304 K, except where otherwise noted, and a relaxation delay of 1.5 s was used in all experiments. The NMR spectra were processed using CCNMR (written by Christian Cieslar at the Max Planck Institute, Martinsried, Germany) or Bruker UXNMR software.

All of the homonuclear 2D experiments were recorded with a spectral width of 6250 Hz in the acquisition dimension, with the carrier positioned at the water frequency. The solvent signal was suppressed by continuous irradiation during the relaxation delays. Typical data sets for 2D homonuclear experiments were 512 increments of 2K complex data points. The TPPI method (Bodenhausen et al., 1984) was used in the indirectly detected dimension in order to achieve quadrature detection. The data were processed with a Gaussian window function in t_2 and a shifted squared sine bell window function in t_1. Linear prediction was employed to extend the acquisition time in t_1 by 50% prior to Fourier transformation. The final processed data matrix consisted of 2K × 1K real points.

The following 2D homonuclear spectra were recorded in H_2O and ^2H_2O at pH 7.5: COSY, NOESY with a mixing time of 120 ms, and TOCSY using an "MLEV-17 clean" mixing sequence (Griesinger et al., 1988) with a mixing time of 60 and 80 ms. 2D heteronuclear ^{1}H−^{15}N HSQC (Marion et al., 1989b) and ^{1}H−^{15}N HMQC (Kay & Bax, 1990) spectra

[1] Abbreviations: COSY, correlation spectroscopy; HMQC, heteronuclear multiple-quantum coherence; HSQC, heteronuclear single-quantum coherence; IPTG, isopropyl β-D-thiogalactopyranoside; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; TOCSY, total correlation spectroscopy; TPPI, time-proportional phase incrementation; 2D, two dimensional; 3D, three dimensional; $^3J_{HN\alpha}$, three-bond, backbone amide proton−α-proton coupling constant; CSI, chemical shift index; PDB, protein data bank; RMSD, root mean square deviation.

were acquired with 1024 complex data points in t_2 and 128 and 512 complex data points in t_1, respectively. The spectral width was 6250 and 2520 Hz in t_2 and in t_1, respectively. The States−TPPI method (Marion et al., 1989a) was used to achieve quadrature detection in t_1. The ^{15}N carrier frequency was positioned at 115 ppm, and spin-lock pulses were used for water suppression. Prior to Fourier transformation the data were zero-filled once in t_2 and linear predicted to 256 and 1024 complex data points in t_1, respectively. Gaussian and shifted square sine bell window functions were applied in t_2 and t_1, respectively.

The 3D ^{1}H−^{15}N TOCSY-HMQC spectra on a sample of P2 dissolved in H_2O/D_2O (95:5) were acquired with a mixing time of 65 ms at two different temperatures, 304 and 294 K. The TOCSY mixing was achieved with a clean DIPSI-rc sequence (Cavanagh & Rance, 1992). The 3D ^{1}H−^{15}N NOESY-HSQC spectra (Marion et al., 1989b) were recorded with a mixing time of 125 ms at three different temperatures, 304, 294, and 284 K. Water suppression was obtained in the 3D experiments by spin-lock pulses during the heteronuclear magnetization transfer (Messerle et al., 1989). Typical data sizes for the 3D experiments were 90 (t_2) × 32 (t_1) × 1024 (t_3) complex points with spectral widths of 6250, 2525, and 6250 Hz, respectively. The data were acquired using the States−TPPI method (Marion et al., 1989a). Zero-filling was used in the acquisition dimension, and linear prediction was used in all the indirectly detected dimensions. The final processed data matrix consisted of 256 × 128 × 1024 real complex points in the t_2, t_1, and t_3 dimensions, respectively.

The 3D triple resonance CBCA(CO)NH (Grezsiek & Bax, 1992) spectrum was recorded on a Bruker AMX2 500 MHz spectrometer. A total of 80 (t_2) × 14 (t_1) × 1024 (t_3) complex data points were collected, with spectral widths of 8347 Hz in t_2, 3125 Hz in t_1, and 6000 Hz in t_3, respectively. The carrier frequencies used in the three dimensions were 4.7 ppm in the ^{1}H dimension, 46 ppm in the $^{13}C_{\alpha/\beta}$ dimension, and 117 ppm in the ^{15}N dimension. The 3D HNCACB (Wittekind & Mueller, 1993; Muhandiram & Kay, 1994) was recorded using the Varian UnityPlus 500 MHz spectrometer and consisted of 120 × 44 × 1024 complex data points. Pulsed field gradients were used for coherence selection in the ^{15}N dimension. The data set was collected with spectral widths of 8990 Hz in t_2, 2500 Hz in t_1, and 6500 Hz in t_3, with 32 scans per complex data point. The positions of the carrier frequencies were as follows: 47.8 ppm in ^{13}C, 118.0 ppm in ^{15}N, and 4.70 ppm in ^{1}H.

The carbonyl pulses in both experiments were applied as phase-shifted laminar pulses with the power adjusted to cause minimal excitation in the C_α region (Patt et al., 1992). Both spectra were recorded with the States−TPPI method (Marion et al., 1989a) and processed in the same manner. Zero-filling was used in the t_3 and linear prediction was used in the t_1 and t_2 dimensions in order to increase the resolution. The final data matrix consisted of 256 × 128 × 1024 complex data points.

The amide proton−deuterium exchange rates were monitored by using the ^{1}H−^{15}N HSQC spectra. After the protonated protein was dissolved in D_2O, a series of 2D ^{1}H−^{15}N HSQC spectra were acquired at times ranging from 0.5 to 12 h. The acquisition time for each spectrum was 15 min.

RESULTS AND DISCUSSION

Sequential Resonance Assignments. Homonuclear 2D TOCSY, heteronuclear ^{1}H−^{15}N 2D HSQC, 3D ^{1}H−^{15}N TOCSY-HMQC, and 3D ^{1}H−^{15}N NOESY-HSQC spectra of samples dissolved in water, and homonuclear 2D NOESY and 2D COSY spectra of samples dissolved in D_2O were used to obtain the sequential resonance assignments. The total number of amide proton cross peaks in the HSQC spectrum is 132 (the protein contains 88 non-proline residues and the theoretical number of cross peaks is 106). The additional peaks, often of low intensity, may represent the minor conformations of P2.

The standard residue assignment procedures described by Wüthrich (1986) were used. The assignment was initiated by identifying as many spin systems as possible in the 2D and 3D TOCSY spectra. Since the 3D ^{1}H−^{15}N TOCSY spectrum was recorded with a mixing time of 65 ms only, a number of side chain assignments were obtained from the 2D TOCSY spectrum where varying mixing times were used. From the 3D ^{1}H−^{15}N TOCSY spectra, two threonine residues were readily identified and used as the starting point for the sequential assignment. Thus, starting from the HN−H_α cross peak of T52, the amino acid stretch between E51 and W61 could be sequentially assigned via $d_{NN}(i, i + 1)$ and $d_{\alpha N}(i, i + 1)$ NOEs. In a similar manner, the sequential stretches N14−A29 and I71−D76 were identified. The sequential stretches L4−D11, A36−R50, and V77−N90 were identified on the basis of $d_{\alpha N}(i, i + 1)$ and $d_{\beta N}(i, i + 1)$ NOEs. The amino acids R39 and I40 could not be unambiguously assigned due to extensive resonance overlap. This ambiguity was settled later during the assignment procedure when these amino acid residues were found to be included in a β-sheet, whereupon good long-range correlations were found between R39 and F9 and between I40 and A8.

There are some peaks which could not be clearly assigned on the basis of these experiments. These correspond to stretches M1−S3, N12−D13, and Q32−H34. These residues showed very weak peaks in the 2D ^{1}H−^{15}N HSQC spectrum. No sequential assignment could be made on the basis of the 3D ^{1}H−^{15}N TOCSY and 3D ^{1}H−^{15}N NOESY spectra, since only a few intraresidual peaks with low intensity could be observed. At the relatively high pH (7.5) required to keep the protein in solution, rapid exchange with water may result in weak or even missing amide proton resonances.

In order to overcome these difficulties and to confirm the assignments obtained using the ^{15}N-labeled sample, triple resonance CBCA(CO)NH and HNCACB experiments were performed on a ^{13}C/^{15}N double-labeled sample. From these two triple resonance spectra, all the previously made assignments were confirmed. However, for the amino acid stretches M1−S3, N12−D13, and H33−H34, the assignment problem still remained. When the 2D COSY and 2D NOESY spectra recorded in D_2O were inspected, two serine-type spin systems in the aliphatic region which had not previously been assigned were found in the 2D COSY spectra. In the 2D NOESY, these two residues show a weak $d_{\alpha\alpha}(i, i + 1)$ correlation. Furthermore, one of the amino acid residues showed a long-range correlation with R45 and was assigned as S3. At this point there was only one resonance, which showed strong NOESY and TOCSY cross peaks, that was not assigned. It resembles an AX_3 spin system, and it

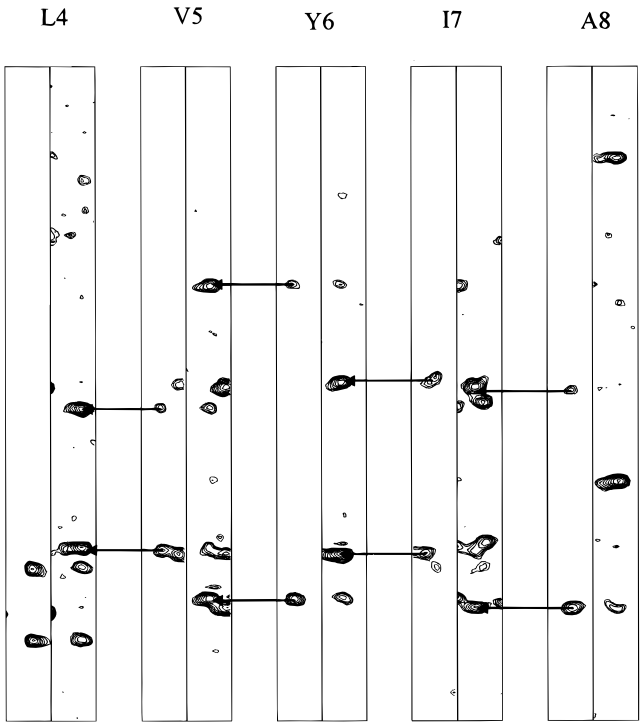

FIGURE 1: Sequential assignment of amino acids 4−8 using CBCA-(CO)NH and HNCACB spectra.

is a possible candidate for one of the histidines, but no sequential correlation could be found for this, neither in 3D NOESY nor in the triple resonance spectra. The C_α and C_β frequencies observed in the CBCA(CO)NH spectrum do not fit with the assigned resonance of Q32, which suggests that this resonance corresponds to H34 rather than H33.

In conclusion, P2 has been sequentially assigned from amino acid residues S2 to N90, excluding the two proline residues, D13 and H33. The chemical shifts (^{1}H, ^{15}N, ^{13}C) of P2 are summarized and supplied as Supporting Information. A figure that displays part of the sequential assignment is shown (Figure 1). The NOEs used in the sequential assignment can be seen in Figure 2.

$^3J_{HN\alpha}$ Coupling Constants and Dihedral Angle Constraints. The coupling constants of $^3J_{HN\alpha}$ were extracted from a ^{1}H−^{15}N HMQC spectrum using the method described by Kay (Kay & Bax, 1990). A total of 70 $^3J_{HN\alpha}$ coupling constants were extracted and converted to angle constraints. Coupling constants were also used in the identification of secondary structure according to the following rules: α-helical segments, $^3J_{HN\alpha} \leq 6$ Hz; β-strand segments, $^3J_{HN\alpha} > 7$ Hz.

Since the $^{13}C_\alpha$ chemical shifts of proteins exhibit a strong correlation with peptide backbone conformation (Spera & Bax, 1991; Wischart & Sykes, 1994; Luginbühl et al., 1995), the empirical $^{13}C_\alpha$ chemical shift surface (Spera & Bax, 1991) can be used to extract dihedral angle constraints of proteins. Recently, Luginbühl et al. (1995) have described a statistical analysis of correlation between experimental dihedral angle constraints derived from $^{13}C_\alpha$ chemical shifts and a group of high-quality protein structures. $\Delta\delta(^{13}C_\alpha)$ values, the differences between the experimental $^{13}C_\alpha$ chemical shifts and the random coil values (Richarz & Wüthrich, 1978), were used to derive dihedral angles and the secondary structures of the protein. It has been shown (Luginbühl et al., 1995) that $\Delta\delta(^{13}C_\alpha) > 2.0$ corresponds to a helix conformation with the φ angle constrained between −100° and 0°, while $\Delta\delta(^{13}C_\alpha) < -1.5$ corresponds to a β-strand conformation with the φ angle constrained between 40° and −140°. On the basis of these observations, 42 φ angle constraints of protein P2 corresponding to α-helices and β-strands were derived from $^{13}C_\alpha$ chemical shifts. Backbone dihedral angles ϕ were restrained with ±40° for upper and lower boundaries, and φ angles were restrained with ±50°.

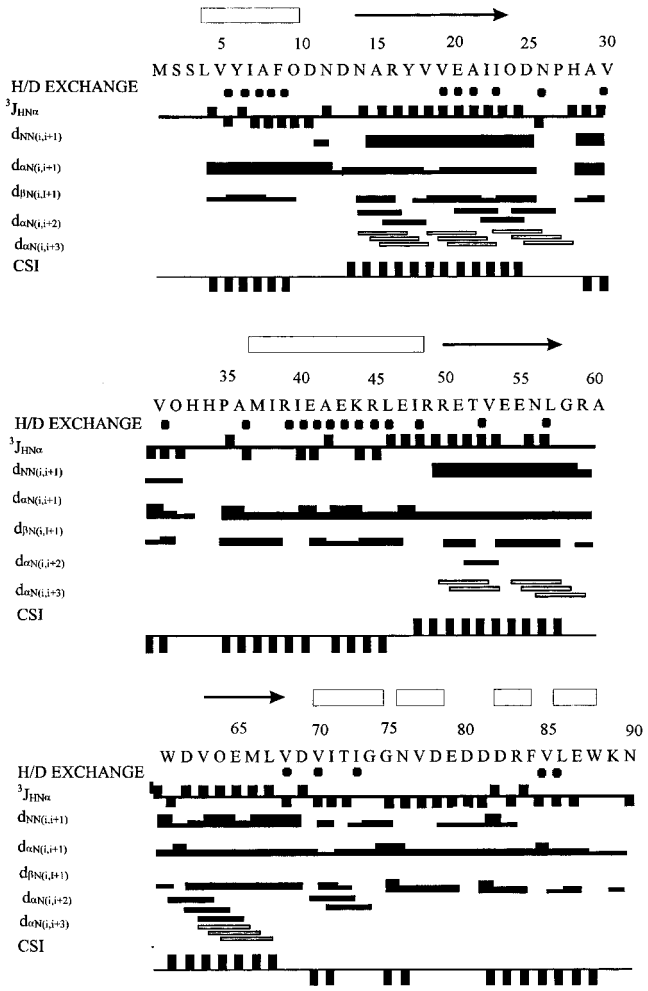

FIGURE 2: Secondary structure analysis of P2. Slowly exchanging backbone amide protons that are still observed in ^{1}H−^{15}N HSQC spectra recorded after 2 h in ^2H_2O solution are indicated by filled circles. The values of the $^3J_{HN\alpha}$ are indicated by an index with values of +1 (≥7 Hz) and −1 (<6 Hz). The sequential NOEs, d_{NN}, $d_{\alpha N}$, and $d_{\beta N}$, are indicated by bars between residues. The NOEs are classified as strong, medium, and weak, according to the heights of the filled bars. The consensus (H_α, C_α, C_β) chemical shift indices are indicated by an index with values of −1, 0, and +1, which correspond to downfield, random coil, and upfield, respectively (Wishart & Sykes, 1994).

Secondary Structure Assignment. The observed NOE pattern was used to extract secondary structure elements of P2. Regions of α-helix and β-sheet were identified according to the criteria established by Wüthrich (1986). The β-sheet assignment was based on strong sequential $d_{\alpha N}(i, i + 1)$, interstrand NH−NH, NH−H_α, and H_α−H_α connectivities, large $^3J_{HN\alpha}$ (>7 Hz) coupling constants, and slow amide proton exchange rates. The identification of α-helical regions was based on sequential $d_{\alpha N}(i, i + 1)$, strong $d_{NH-NH}(i, i + 1)$, medium $d_{\alpha N}(i, i + 3)$, and small coupling constants (≤6 Hz). On the basis of the results of these analyses, we have identified six β-strands consisting of residues 3−10, 37−44, 46−49, 70−75, 77−79, and 84−89 which are arranged

in an antiparallel manner and three α-helices consisting of residues 14−26, 50−58, and 63−67. The location and the extent of the secondary structure elements were further confirmed by the analysis of $^{13}C_\alpha$, $^{13}C_\beta$, and H_α chemical shifts (Wishart & Sykes, 1994). The consensus results of the chemical shift index (H_α, C_α, C_β) analysis (Figure 2) are in excellent agreement with the secondary structure elements derived from the NOE and amide proton−deuteron exchange experiments. However, the chemical shift index analysis (Figure 2) suggests that the amino acids from A29 to Q32 form a β-strand. In the structure analysis process, some long-range interactions ($H_\alpha(i)$−NH(j)) were observed between the amino acid stretches V30−H34 and M37−E41, but it has not been possible to define these regions as a β-sheet due to spectral overlap in these regions. The secondary structural analyses of P2 are summarized in Figure 2.

Structure Calculation. Interproton distance constraints were obtained from either the ^{15}N-edited 3D NOESY-HSQC spectrum recorded with a mixing time of 120 ms or the 2D homonuclear NOESY spectrum recorded in D_2O using the same mixing time. Semiquantitative determination of the cross peak intensity was based on the number of contour levels of the peaks. All the NOE data were categorized as strong, medium, or weak peaks, corresponding to three distance ranges, 1.8−3.0, 1.8−4.0, and 1.8−5.0 Å, respectively. Pseudoatoms were used for all nonstereospecifically assigned protons, and an appropriate distance correction was applied to the distance constraints (Wüthrich et al., 1983). In addition, 0.5 Å was added to the distance constraints involving methyl protons (Clore et al., 1987).

All calculations were carried out using the X-PLOR 3.0 program (Brünger, 1992). In total, 1060 distance constraints [382 intraresidue, 330 sequential, 163 medium range ($|i - j| \leq 5$), and 185 long range ($|i - j| > 5$) NOE distance constraints] and 70 ϕ and 42 φ dihedral angle constraints were used. The calculations were started from a template structure with randomized backbone ϕ and φ torsion angles. The 50 starting structures were embedded using the dg_sub_embed.inp protocol, which was followed by the dgsa.inp protocol (Nilges et al., 1988b; Kuszewski et al., 1992). In the dgsa step, a square well potential function was used for the NOE constraints, and an 18 ps molecular dynamics calculation was performed at 2000 K followed by 6000 steps of simulated annealing. The structures produced using the dgsa.inp protocol were further subjected to a simulated annealing refinement (Nilges et al., 1988ab). The starting temperature of the simulated annealing refinement was 2000 K, and 10 000 annealing steps were used in order to avoid local minima. The final temperature was 100 K (Nilges et al., 1988b), and after that 1000 steps of conjugate gradient minimization were performed. In the structure calculation, force constants for the NOE-derived distance constraints were set to 50 kcal mol^{-1} Å^{-2}, and the dihedral angle constraints were initialized at 5 kcal mol^{-1} rad^{-2} during the high-temperature dynamics and gradually increased to 200 kcal mol^{-1} rad^{-2} during the annealing stage.

A good correlation was observed for hydrogen bonds between backbone amide protons in the calculated structures and observed slow amide proton exchange. In all cases where slow amide proton−deuteron exchange was observed, hydrogen bonds were also found in the calculated structures. However, for nine residues (A15, R16, Y17, T52, V53, V63, Q64, G74, D78), the amide protons were found to be involved in hydrogen bonds in the calculated structures but were not observed as slowly exchanging in the amide proton−deuteron exchange experiments. Nevertheless, all of these residues exhibit NOEs characteristic of regular secondary structure. The reason why they are not observed as slowly exchanging might be that the dead time (20 min) of the exchange experiment is too long.

At the final stage of the calculations, 34 hydrogen bond distance constraints based on the analysis of amide proton−deuterium exchange experiments and the derived secondary structure were added with target values of 1.8−2.3 Å for NH(i)−O(j) and 2.4−3.3 Å for N(i)−O(j). Fifty structures were calculated out of which 12 structures showed good agreement with the NMR experimental constraints. For these structures, the NOE distances and torsion angle violations were smaller than 0.5 Å and 10°, respectively, and thus they were selected for further analysis.

The mean structure of the ensemble was calculated after superposition of the backbone nuclei (N, C_α, and C') of residues 4−10, 14−26, 37−43, 46−49, 51−58, 63−67, 70−75, and 84−88, which belong to the regular secondary structure elements of P2 and hence were better defined by NMR data. For the 12 final converged structures, an atomic RMSD to the mean coordinate is 2.5 Å for backbone atoms and 3.9 Å for all heavy atoms. A plot of NOE distance constraints vs sequence is presented in Figure 3a. Figure 3b shows the plot of the average ⟨RMSD⟩ to the mean structure. Figure 3c shows secondary structure elements vs sequence number. Finally, Figure 3d,e shows the angular order parameters (Hyberts et al., 1992) for ϕ and ψ angles as a function of the residue number. Figure 3b shows that all the β-strand parts of P2 were well defined compared with the other parts of the structure. When just superposition of the individual secondary structure parts of the protein was considered, the RMSD values decrease substantially. The average RMSD values obtained from this partial superposition of secondary structure elements are summarized in Table 1.

From the comparison of the global RMSD with local RMSD, it appears that the local secondary structure elements of P2 are all well defined, but the positions of the helices relative to other parts of the protein are poorly defined. In fact, no long-range NOEs have been observed between any of the helices and the other parts of the structure.

Another part of the protein which is not well defined is the region from amino acid residues 27 to 35. Contributions to this uncertainty come from the fact that H33 was not assigned and that no interresidue distance constraints were used for the amino acids from H33 to P35 in the structure calculations. Additionally, both V31 and Q32 are amino acids which, according to the NMR spectra, exist in multiple conformations (vide infra). The calculated structure of P2 indicates that this region is solvent exposed, which facilitates the amide proton exchange with water. At pH 7.5, the exchange rates for solvent-exposed amide protons are close to the range where NMR detection is not feasible (Wüthrich, 1986). The same reasoning can also be put forward for the uncertainty in the position of the first three residues located at the amino terminus.

The precision of the P2 structures was also analyzed by examination of the angular order parameter, S, of the ϕ and ψ dihedral angles as a function of residue number (Figure

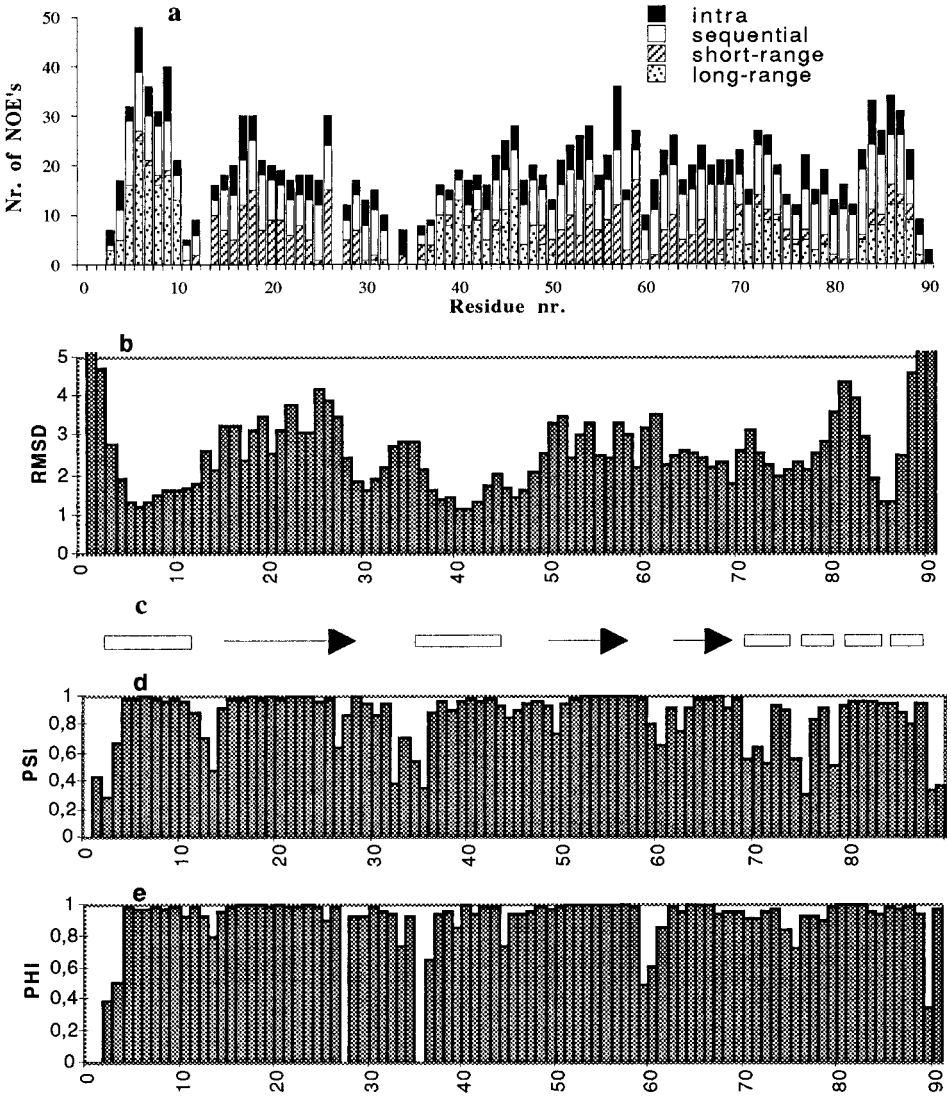

FIGURE 3: (a) Number of NOE constraints vs sequence number. (b) Average backbone atomic (N, Cα, C′) RMSD per residue for the 12 best structures. (c) Secondary structures represented by open boxes for β-strands and arrows for α-helices. (d and e) Plots of angular order parameters for backbone dihedral angles ψ and ϕ calculated from the 12 best structures.

Table 1: Structural Statistics

RMSD of secondary structures	av RMSD vs av coordinates (Å)
strand I (4−10)	0.33
strand II (37−44)	0.41
strand III (46−49)	0.27
strand IV (70−75)	0.83
strand V (76−79)	0.71
strand VI (83−88)	0.97
helix I (14−26)	0.94
helix II (51−58)	0.89
helix III (63−68)	0.97

3d,e). Angular order parameters are another measure of the precision of the ensemble of structures (Hyberts et al., 1992). Precisely determined dihedral angles will give S values close to 1, whereas if the angular order parameter has a value of 0, it means that the angle deviates in a random manner, inferring a disordered structure. In general, the residues with high angular order parameter values exhibit lower RMSD from the mean structure and are well defined by a large number of NOE constraints. A value of $S = 0.99$ corresponds to an angle standard deviation of ±7.5°, a value of $S = 0.95$ corresponds to an angle standard deviation of ±17°, and a value of $S = 0.9$ corresponds to an angle standard deviation of ±24°. The order parameters shown in Figure 3d,e indicate that the overall structures with respect to torsion angles are well defined. The angular order parameters for the ϕ and ψ backbone dihedrals have mean values of 0.95 and 0.91 (for residues 4−88), respectively, which indicate good convergence for these angles. Residues at the C- and N-termini exhibit low angular order parameters, which is expected since these regions tend to be disordered. There are two other regions, 27−35 and 69−75, which also show low angular order parameters. In these cases, disorder may be attributed to the presence of proline (amino acid 27 and 35) and glycine (amino acid 74 and 75) residues.

The Ramachandran plot (Supporting Information) of the ϕ and ψ angles for all 12 converged structures shows that 64% of the ϕ and ψ angles are found within the most favored region, 28.4% are found within the additionally allowed regions, and 6.6% are located within the generously allowed regions. Only 1% of the residues are found in disallowed regions. These residues are marked by residue number.

Description of the Structure. The PDB file of protein P2 was submitted to a protein structure database for searching structure similarity (Holm & Sander, 1994). No similar structure was found in this search.

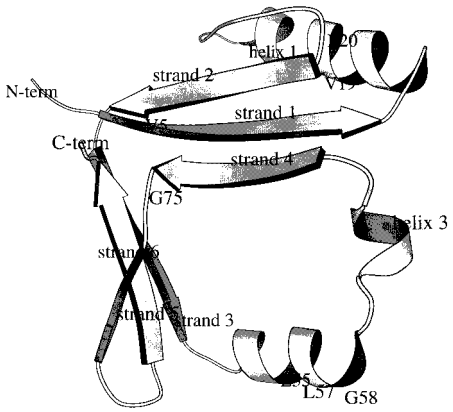

FIGURE 4: Ribbon diagram of the structure of P2. Strictly conserved residues are marked by residue number. This diagram was generated using the program MOLSCRIPT (Kraulis, 1991).

The ribbon diagram representing the structure of P2 is shown in Figure 4. The secondary structure of protein P2 comprises three α-helices and six β-strands. The first β-sheet consists of three β-strands: β1(residues 4−10), β2 (residues 37−43), and β3 (residues 70−75). β1 is flanked by strands β2 and β3 and has long-range NOE interactions with both β2 and β3. Strands β1 and β2 are connected by a bent helix 1 (residues 14−26) and the ill-defined proline-containing loop region (residues 27−35). The second β-sheet consists of strands β4 (residues 46−48), β5 (residues 77−79), and β6 (residues 84−88). β6 is flanked by strands β4 and β5 and has long-range NOE interactions with these strands. Strands β5 and β6 are connected by a β-hairpin region (residues 80−84). Strands β3 and β4 are connected by two helices, helix 2 (residues 50−58) and helix 3 (residues 63−67). A stereoview of 12 ensemble structures of P2 is presented in Figure 5.

The two helices which are connected by a random coil region, helix 2 and helix 3, create an extended region of the structure, which is characterized by a high degree of disorder when the different calculated structures are compared (Figure 5). This is not unexpected since the amino acids in this part of the structure are involved in conformational exchange processes (vide supra). However, the observed disorder could equally well be attributed to the lack of NOEs ranging further than three residues apart within this region. Nevertheless, all the calculated structures have as a common feature that the extended region creates a cavity which is surrounded by hydrophobic residues. The reason why this unusual structural feature, although disordered, is prevalent in all the calculated structures is the experimental constraints that define the β-sheet structures. In the absence of an extended region it would have been impossible to satisfy these constraints.

Although most of the hydrophobic residues of P2 are oriented toward the interior of the protein, there is no obvious hydrophobic core in the structure. Instead, there is an empty region surrounded by hydrophobic residues, located in the middle of the structure. This structural feature is evident in the space-filling model shown in Figure 6.

Conformational Flexibility of P2. In the ^{1}H−^{15}N HSQC spectrum recorded at 30 °C (Figure 7), a number of unassigned resonances are observable. Generally these resonances were of low intensity and appeared in the NMR spectrum at frequencies close to more intense and assigned resonances. When the NOESY patterns for these peaks were compared with the assigned neighboring resonances in the ^{1}H−^{15}N HSQC spectrum (see Figure 7), it was found that they were similar. Accordingly, it is reasonable to conclude that these additional resonances stem from less populated conformations of P2 in solution, coexisting with the major conformation. In resolved regions of the spectra, these resonances could be identified as the following amino acids belonging to additional solution conformations of P2: F9, Q10, V18, V19, E20, I22, A29, V31, Q32, E47, I48, E51, T52, V68, D69, I71, D76, V77, and N90.

Additional evidence for coexisting conformations of P2 is that some of the resonances ascribed to the major conformation changed their frequencies when the temperature was changed. In these experiments, ^{1}H−^{15}N HSQC spectra were recorded at three different temperatures: 30, 20, and 10 °C. In the spectrum recorded at 10 °C, the intensities of those resonances ascribed to amino acid residues of a minor conformation of P2 increased in intensity. Moreover, two intense resonances could be observed at frequencies which are not observable at 30 °C. One of these peaks, located in the upfield region of the ^{15}N dimension, could be assigned to G58 on the basis of the simultaneous disappearance of the resonance assigned to G58 at higher temperatures and on the similar pattern of the corresponding NOESY cross peaks.

Generally, the 3D ^{1}H−^{15}N NOESY-HSQC and 3D ^{1}H−^{15}N TOCSY-HSQC spectra recorded at 10 °C showed that a number of NOESY and TOCSY cross peaks were either missing or had reduced intensity compared with the intensities observed in spectra recorded at higher temperatures. The TOCSY spectrum was especially affected, which can be ascribed to the increased transverse relaxation rates observed at low temperature during the mixing time. This effect on the relaxation rates is indicative of conformational exchange taking place. That these relaxation rates were much slower at 30 °C can be explained by a reduction in the population of the less populated conformations, since this will make the impact of the exchange on the NMR spectra less significant.

Preliminary binding studies using NMR spectroscopy suggest that P2 interacts with both the hydroxylase components and phenol in regions close in space on P2. In an experiment in which P2 was titrated with the hydroxylase components of the oxygenase, the resonances of the minor conformations of P2 disappeared (Qian et al., unpublished). Moreover, the resonances that were affected upon addition of the hydroxylase were all localized to residues surrounding the cavity in the extended region in the solution structure of P2 (Figure 6). Studies of interactions between P2 and phenol have also been initiated (Qian et al., unpublished). When phenol was titrated into a sample of P2, the resonance frequencies of the ξ proton of W61 and the amide proton of D62 changed, which indicates that phenol binds to a region close to helix 3.

Overall, a picture emerges where the conformational flexibility of P2 is localized to the extended region of the structure where a cavity surrounded by hydrophobic residues is located. Interestingly, the same region appears to be responsible for the interaction with phenol and hydroxylase components. It is also of interest to note that the cavity in the structure is large enough to allow passage of a nonhydrated phenol molecule. Therefore, it is conceivable that this structural motif allows entrance of substrates to the active

FIGURE 5: Stereoview of the superposition of backbone atoms of the 12 best P2 structures. The top pair represents the superposition of all regular secondary structure elements of P2. The lower pair represents superposition of strands 1, 2, and 4 of P2.

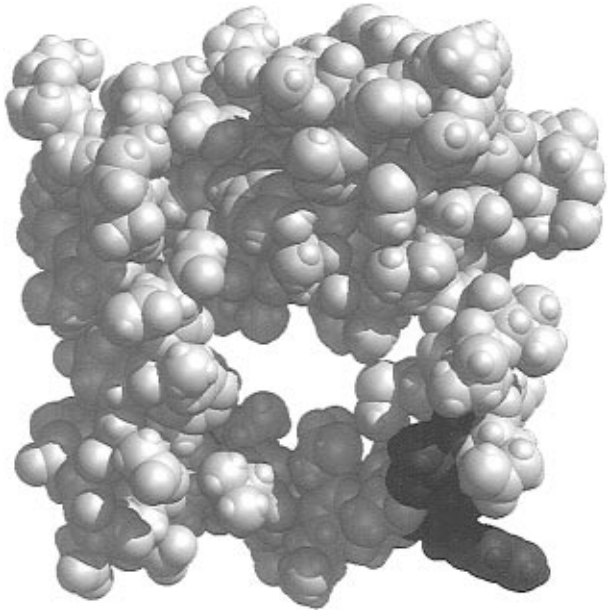

FIGURE 6: Space-filling diagram of P2. The structure and the orientation of the molecule used to generate this plot are the same as in Figure 5. The amino acids that bind phenol are labeled in red. This plot was generated using the program QUANTA.

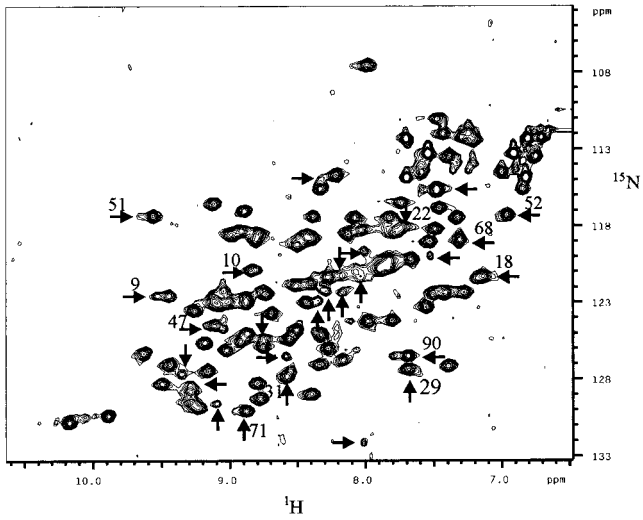

FIGURE 7: ^{1}H−^{15}N HSQC spectrum of ^{15}N-labeled P2 at 30 °C. Arrows and residue numbers indicate peaks appearing as doublets. Some additional peaks which were not assigned are marked by arrows only. The complete chemical shift assignment table is given as Supporting Information.

site, and exit of products, when P2 is complexed with the hydroxylase. A tentative proposal is therefore that this structural motif is essential for the contribution of P2 to the enzymatic activity of phenol hydroxylase.

The flexibility of this region may also be important for function. It is this extended region, particularly the helices and loops located there, that exhibits the largest RMSD values for the calculated structures. The coexistence of minor conformations will affect the observed NOE intensities and thus also contribute to the large observed RMSD. Therefore, it appears that the relatively low resolution of the structure of P2 that we can determine is caused by conformational exchange and structural flexibility in this region. However, it is worth stressing that the small energy barriers for conformational exchange in this region might be of significance for ability of P2 to bind and interact with the other components of phenol hydroxylase and the substrates.

A number of proteins have previously been reported to contain cavities within their structures where major conformational changes can be accommodated. For instance, it has been shown that the cavity found in the myb protein allows conformational changes to occur that are important for DNA binding (Ogata et al., 1995). The large frequency changes of the amide resonances of G58 that occur when the temperature changes indicate that this region experiences substantially altered environments in different conformations of the protein. Another interesting observation is that G58 is conserved in all the P2-like proteins associated with other oxygenases (see below). It may be that the flexibility inherent in this region is important for biological function since glycine residues do not impose severe steric constraints for flexibility. In the case of P2, G58 may form a hinge region between helix 2 and helix 3. In Figure 6, one can see that this structural motif of P2 may permit helix 2 and helix 3, or the peripheral region surrounding this region, to move relative to the core region of P2 without perturbing other parts of the structure substantially.

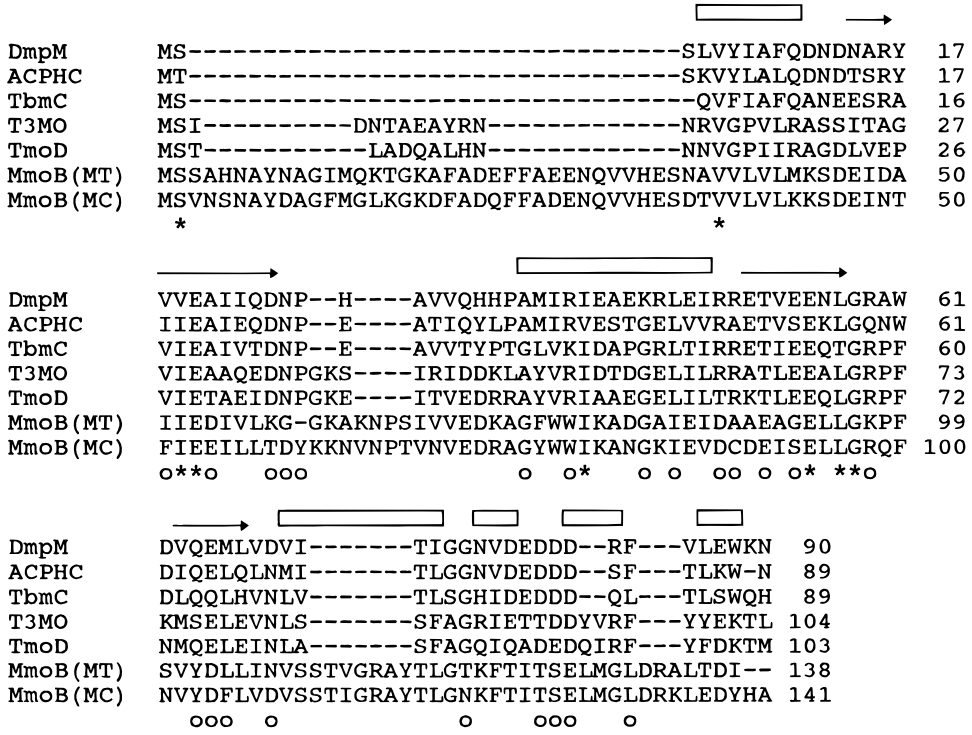

FIGURE 8: Amino acid sequence comparisons between P2 (DmpM) and components of phenol hydroxylase from *A. calcoaceticus* (ACPHC), toluene monooxygenases (tbmC, T3MO, TbmD), and methane monooxygenases (MmoB) from *M. trichosporium* (MT) and *M. capsulatus* (Bath) (MC). At the top of the panel, the secondary structure is represented by open boxes for β-strands and arrows for α-helices. Sequences were aligned using the program Clustal (Higgins & Sharp, 1988, 1989). Symbols represent the following: (*) conserved in at least six of seven sequences; (○) conserved in at least four of five aromatic hydroxylases. Sequences are taken from Nordlund et al. (1990) (DmpM), Ehrt et al. (1995) (ACPHC), Johnson and Olsen (1995) (TbmC), Byrne et al. (1995) (T3MO), Yen et al. (1991) (TmoD), Cardy et al. (1991) [MmoB (MT)], and Stainthorpe et al. (1989) [MmoB (MC)].

Comparison with Similar Proteins of Other Multicomponent Oxygenases. Both methane and toluene multicomponent monooxygenases include protein components which show sequence similarities with the phenol hydroxylase-associated P2 (Figure 8). It is therefore of interest to compare the positions of conserved residues in six of these sequences with the structure of P2 (Figure 8). Six residues (V5, E20, E55, L57, G58, G75) are strictly conserved in all of these proteins, and conservative substitutions are found for a number of the hydrophobic amino acid residues; in addition, 26 residues are strictly conserved in all, or all but one, aromatic hydroxylase-associated protein. Although the lengths of the polypeptide chains of the different proteins vary from 89 amino to 141 amino acids, all of the highly conserved residues in the group of proteins shown in Figure 6 are found in regions of regular secondary structure elements found in P2. There are three conserved amino acid residues in helix 2 (E54, L57, G58), which implies that this helical part is important for the biological function of P2. As was discussed above, some interesting phenomena were observed in this region in the presence of substrate and the hydroxylase partner. The extra amino acid residues inserted in the methane monooxygenase MmoB component at the amino-terminus, at the carboxy-terminus, and between P27 and A29 are in regions corresponding to the ill-defined local structures of P2. In summary, the correspondence between regions of secondary structure in P2 and conserved amino residues in the other sequences suggests that all of these proteins have very similar secondary structure elements, at least in the core regions.

ACKNOWLEDGMENT

The authors gratefully acknowledge the Swedish NMR Center for allowing us to use the NMR instruments and Dr. Lotta Johansson for assistance in recording the experiments there.

SUPPORTING INFORMATION AVAILABLE

One table giving ^{1}H, ^{15}N, ^{13}C chemical shifts of P2 and two figures showing a plot of Ramachandran ϕ and ψ backbone dihedral angles for the 12 best structures and a plot for the distribution of the ϕ and ψ angles of P2 (4 pages). Ordering information is given on any current masthead page.

REFERENCES

Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370.

Brünger, A. T. (1992) *X-PLOR Version 3.1*, Yale University Press, New Haven, CT.

Byrne, A. M., Kukor, J. J., & Olsen, R. H. (1995) *Gene 154*, 65−70.

Cardy, D. L., Laidler, V., Salmond, G. P., & Murrell, J. C. (1991) *Mol. Microbiol. 5*, 335−342.

Cavanagh, J., & Rance, M. (1992) *J. Magn. Reson. 96*, 670−678.

Clore, G. M., Gronenborn, A. M., Nilges, M., & Ryan, C. A. (1987) *Biochemistry 26*, 8012−8023.

Ehrt, S., Schirmer, F., & Hillen, W. (1995) *Mol. Microbiol. 18*, 13−20.

Fox, B. G., Frolund, W. A., Dege, J., & Lipscomb, J. (1989) *J. Biol. Chem. 264*, 10023−10033.

Fox, B. G., Liu, Y., Dege, J., & Lipscomb, J. D. (1991) *J. Biol.Chem. 266*, 540−550.

Frolund, W. A., Andersson, K. K., Lee, S.-K., Liu, Y., & Lipscomb, J. D. (1992) *J. Biol. Chem. 267*, 17588−17597.
Grezsiek, S., & Bax, A. (1992) *J. Am. Chem. Soc. 114*, 6291−6293.
Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870−7872.
Herrmann, H., Muller, C., Schmidt, I., Mahnke, J., Petruschka, L., & Hanke, K. (1995) *Mol. Gen. Genet. 247*, 240−246.
Higgins, D. G., & Sharp, P. M. (1988) *Gene 73*, 237−244.
Higgins, D. G., & Sharp, P. M. (1989) *CABIOS 5*, 151−153.
Holm, L., & Sander, C. (1994) *Proteins 19*, 165−173.
Hyberts, S. G., Goldberg, M. S., Havel, T. F., & Wagner, G. (1992) *Protein Sci. 1*, 736−751.
Johnson, G. R., & Olsen, R. H. (1995) *Appl. Environ. Microbiol. 61*, 3336−3346.
Lipscomb, J. D. (1994) *Annu. Rev. Microbiol. 48*, 371−399.
Kay, L. E., & Bax, A. (1990) *J. Magn. Reson 86*, 110−126.
Kraulis, P. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kuszewski, J., Nilges, M., & Brünger, A. T. (1992) *J. Biomol. NMR 2*, 33.
Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
Luginbühl, P., Szyperski, T., & Wüthrich, K. (1995) *J. Magn. Reson., Ser. B 109*, 229−233.
Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989a) *J. Magn. Reson. 85*, 393−399.
Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., & Bax, Ad. (1989b) *J. Am. Chem. Soc. 111*, 1515−1517.
Mason, J. R., & Cammack, R. (1992) *Annu. Rev. Microbiol. 46*, *277*−305.
Messerle, B. A., Wider, G., Otting, G., Weber, C., & Wüthrich, K. (1989) *J. Magn. Reson. 85*, 608−613.
Muhandiram, D. R., & Kay, L. E. (1994) *J. Magn. Reson., Ser. B 103*, 203−216.
Newman, L., & Wackett, L. (1995) *Biochemistry 34*, 14066−14076.
Ng, L. C., Shingler, V., Sze, C. C., & Poh, C. L. (1994) *Gene 151*, 29−36.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988a) *FEBS Lett. 239*, 129−136.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988b) *FEBS Lett. 229*, 317−324.
Nordlund, I., Powlowski, J., & Shingler, V. (1990) *J. Bacteriol. 172*, 6826−6833.
Ogata, K., Kanei-lshii, C., Sasaki, M., Hatanaka, H., Nagadoi, A., Enari, M., Nakamura, H., Nishimura, Y., Ishii, S., & Sarai, A. (1996) *Nat. Struct. Biol. 3*, 178−187.
Patt, S. L. (1992) *J. Magn. Reson. 96*, 94−102.
Paulsen, K. E., Liu, Y., Fox, B. G., Lipscomb, J. D., Munck, E., & Stankovich, M. (1994) *Biochemistry 33*, 713−722.
Powlowski, J., & Shingler, V. (1990) *J. Bacteriol. 172*, 6834−6840.
Powlowski, J., & Shingler, V. (1994) *Biodegradation 5*, 219−236.
Pulver, S., Frolund, W., Fox, B. G., Lipscomb, J. D., & Solomon, E. I. (1993) *J. Am. Chem. Soc. 115*, 12409−12422.
Richarz, R., & Wüthrich, K. (1978) *Biopolymers 17*, 2133.
Rosenburg, A. H., Lade, B. N., Chui, D.-S., Lin, S. W., Dunn, J. J., & Studier, F.W. (1987) *Gene 56*, 125−135.
Rosenzweig, A. C., & Lippard, S. J. (1994) *Acc. Chem. Res. 27*, 229−236.
Sambrook, J., Fritch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Spera, S., & Bax, A. (1991) *J. Am. Chem. Soc. 113*, 5490−5492.
Stainthorpe, A. C., Murrell, J. C., Salmond, G. P., Dalton, H., & Lees, V. (1989) *Arch. Microbiol. 152*, 154−159.
Wishart, D. S., & Sykes, B. D. (1994) *J. Biomol. NMR 4*, 171−180.
Wittekind, M., & Mueller, L. (1993) *J. Magn. Reson., Ser. B 101*, 201−205.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949−961.
Yen, K.-M., Karl, M. R., Blatt, L. M., Simon, M. J., Winter, R. B., Fausset, P. R., Lu, H. S., Harcourt, A., & Chen, K. (1991) *J. Bacteriol. 173*, 5315−5327.

BI9619233