Solution Structure of the Toluene 4-Monooxygenase Effector Protein (T4moD)[†,‡]

Hikaru Hemmi,[§] Joey M. Studts,[||] Young Kee Chae,[§] Jikui Song,[§] John L. Markley,*[,§,||] and Brian G. Fox*[,||]

National Magnetic Resonance Facility at Madison and Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin, Madison Wisconsin 53706-1544

Received June 15, 2000; Revised Manuscript Received October 18, 2000

ABSTRACT: Toluene 4-monooxygenase (T4MO) from *Pseudomonas mendocina* catalyzes the NADH- and O_2-dependent hydroxylation of toluene to form *p*-cresol. The complex consists of an NADH oxidoreductase (T4moF), a Rieske ferredoxin (T4moC), a diiron hydroxylase [T4moH, with $(\alpha\beta\gamma)_2$ quaternary structure], and a catalytic effector protein (T4moD). The solution structure of the 102-amino acid T4moD effector protein has been determined from 2D and 3D ^{1}H, ^{13}C, and ^{15}N NMR spectroscopic data. The structural model was refined through simulated annealing by molecular dynamics in torsion angle space (DYANA software) with input from 1467 experimental constraints, comprising 1259 distance constraints obtained from NOEs, 128 dihedral angle constraints from *J*-couplings, and 80 hydrogen bond constraints. Of 60 conformers that met the acceptance criteria, the 20 that best satisfied the input constraints were selected to represent the solution structure. With exclusion of the ill-defined N- and C-terminal segments (Ser1−Asn11 and Asp99−Met102), the atomic root-mean-square deviation for the 20 conformers with respect to the mean coordinates was 0.71 Å for the backbone and 1.24 Å for all non-hydrogen atoms. The secondary structure of T4moD consists of three α-helices and seven *β*-strands arranged in an N-terminal *βαββ* and a C-terminal *βααβββ* domain topology. Although the published NMR structures of the methane monooxygenase effector proteins from *Methylosinus trichosporium* OB3b and *Methylococcus capsulatus* (Bath) have a similar secondary structure topology, their three-dimensional structures differ from that of T4moD. The major differences in the structures of the three effector proteins are in the relative orientations of the two *β*-sheets and the interactions between the α-helices in the two domains. The structure of T4moD is closer to that of the methane monooxygenase effector protein from *M. capsulatus* (Bath) than that from *M. trichosporium* OB3b. The specificity of T4moD as an effector protein was investigated by replacing it in reconstituted T4MO complexes with effector proteins from monooxygenases from other bacterial species: *Pseudomonas pickettii* PKO1 (TbuV, toluene 3-monooxygenase); *Pseudomonas* species JS150 (TbmC, toluene 2-monooxygenase); and *Burkeholderia cepacia* G4 (S1, toluene 2-monooxygenase). The results showed that the closely related TbuV effector protein (55% sequence identity) provided partial activation of the complex, whereas the more distantly related TbmC (34% sequence identity) and S1 (29% sequence identity) did not. The ^{1}H NMR chemical shifts of the side-chain amide protons of Asn34, a conserved, structurally relevant amino acid, were found to be similar in spectra of effector proteins T4moD and TbuV but not in the spectrum of TbmC. This suggests that the region around Asn34 may be involved in structural aspects contributing to functional specificity.

Toluene-4-monooxygenase (T4MO)[1] is a four-component complex that catalyzes the NADH- and O_2-dependent hydroxylation of toluene to form *p*-cresol (*1*). The complex consists of an NADH oxidoreductase (T4moF, 33 kDa), a Rieske-type ferredoxin (T4moC, 12.5 kDa), a diiron hydroxylase (T4moH, with a $(\alpha\beta\gamma)_2$ quaternary structure, 212 kDa), and an effector protein (T4moD, 11.6 kDa). T4MO belongs to an evolutionarily related family of soluble diiron hydroxylases including other toluene, benzene, and xylene monooxygenases, methane monooxygenase, phenol hydroxylases, and alkene epoxidases (*1*−*3*). In all of these complexes, a small effector protein is required for catalysis (*4*−*8*). Since

[†] This work was supported by the National Science Foundation Early Career Development Program (MCB-9733734) and the Petroleum Research Fund (32955-AC4) to B.G.F. and NIH grants (GM 58667 and RR 02301) to J.L.M. Equipment in the National Magnetic Resonance Facility at Madison (NMRFAM) was purchased with funds from the University of Wisconsin, the NSF Biological Instrumentation Program (DMB-8415048), NIH Biomedical Research Technology Program (RR 02301), NIH Shared Instrumentation Program (RR 02781), and the U.S. Department of Agriculture. J.M.S. was supported in part by a Peterson/Wharton Fellowship from the Department of Biochemistry, University of Wisconsin.

[‡] The atomic coordinates for the minimized average structure and the 20 best conformers have been deposited with the Protein Data Bank under the file names 1G10 and 1G11, respectively.

* To whom correspondence should be addressed. (B.G.F.) E-mail: bgfox@biochem.wisc.edu. Phone: (608) 262-9708. Fax: (608) 265-2904. (J.L.M.) E-mail: markley@nmrfam.wisc.edu. Phone: (608) 263-9349. Fax: (608) 262-3173.

[§] National Magnetic Resonance Facility at Madison.

[||] Department of Biochemistry.

[1] CSI, chemical shift index; DmpM, effector protein of phenol hydroxylase from *Pseudomonas putida* CF600; MmoB, catalytic effector component of soluble methane monooxygenase; RMSD, root-mean-square deviation; S1, catalytic effector component of toluene 2-monooxygenase from *Burkeholderia cepacia* G4; T4MO, four-protein toluene 4-monooxygenase complex from *Pseudomonas mendocina* KR1; T4moD, catalytic effector component of T4MO; TbmC, catalytic effector component of toluene 2-monooxygenase from *Pseudomonas* species JS150; TbuV, catalytic effector component of toluene 3-monooxygenase from *Pseudomonas pickettii* PKO1.

10.1021/bi0013703 CCC: $20.00
Published on Web 03/01/2001

these effector proteins contain neither organic cofactors nor metal ions, it is unlikely that they participate directly in electron transfer or other oxidative aspects of catalysis. Therefore, it has been assumed that the diiron effector proteins form specific protein−protein complexes and produce conformational changes essential for catalysis. This assessment is supported by chemical cross-linking, spectroscopic, and catalytic studies (*9*−*20*).

Structures have been reported for the effector proteins from phenol hydroxylase (DmpM) (*21*) and the methane monooxygenases (MmoBs) from *Methylosinus trichosporium* OB3b (*22*) and *Methylococcus capsulatus* (Bath) (*23*), all determined by NMR spectroscopy. The structure of recombinant DmpM produced in *Escherichia coli* has a $\beta\alpha\beta$ and $\beta\alpha\alpha\beta\beta\beta$ topology with a relatively large RMSD in atomic positions originally ascribed to conformational flexibility and exchange (*21*). However, subsequent studies showed that recombinant DmpM expressed in *E. coli* requires a refolding treatment in order to assume a catalytically active form (*24*). Although the inactive form was deduced to be a dimer, the structural changes introduced by conversion to an active monomer during the refolding step are not known. By contrast, the MmoBs from two species of methanotrophic bacteria were found to share a different secondary structural topology, consisting of $\beta\alpha\beta\beta$ and $\beta\alpha\alpha\beta\beta$ domains (*22*, *23*). The compact tertiary structures of the MmoB proteins differ from one another in the orientations of the β-sheets and the extent of contacts between the α-helices in the two domains.

Here we report the solution structure of a fourth effector protein T4moD, the structure-based phylogenetic analysis of the entire effector protein family, the results of studies designed to investigate the catalytic specificity of effector proteins from the aromatic monooxygenase subfamily of diiron hydroxylases, and preliminary analysis of NMR spectra of two additional effector proteins from this subfamily. The structure of T4moD is compared with those of other members of the effector protein family. The phylogenetic and catalytic studies provide the basis for detailed examination of the protein surface and specific residue contacts required for catalytic protein−protein interactions in the T4MO complex and other diiron hydroxylases.

MATERIALS AND METHODS

Sample Preparation. [*U*-^{13}C, *U*-^{15}N] T4moD was expressed in *E. coli* BL21 (DE3) grown on a minimal medium containing [*U*-^{13}C]-D-glucose (Isotech, Miamisburg, OH) and $^{15}NH_4Cl$ (Cambridge Isotope Labs, Andover, MA). Details of the expression and purification are reported elsewhere (*25*). Similar methods (to be reported elsewhere)[2] were used in the cloning, expression, and purification of TbuV and TbmC.

NMR Spectroscopy and Data Processing. NMR samples used for data collection under steady-state conditions contained 1.1 mM [*U*-^{13}C, *U*-^{15}N] T4moD or 3.0 mM natural abundance T4moD in 50 mM phosphate buffer, pH 7.0 (90:10 v/v H_2O/D_2O or 99.9% D_2O), 0.5 μM sodium azide, and a protease inhibitor cocktail (Product no. P 27124, Sigma Chemical Co., Saint Louis, MO). The sample used in the hydrogen-exchange experiment was a 1.0 mM solution of T4moD prepared in the above buffer adjusted to pH 5.0; this solution was lyophilized and then dissolved in D_2O to initiate the ^{1}H-^{2}H exchange measurement. NMR samples of natural abundance TbuV (3.6 mM) and TbmC (5.4 mM) were prepared in the same buffer used for T4moD studies.

All NMR spectra were recorded at 25 °C on Bruker DMX600 and DMX500 spectrometers equipped with triple-resonance ^{1}H, ^{13}C, and ^{15}N probes and three-axis pulsed field gradients. Quadrature detection in the indirectly detected dimensions was obtained with the TPPI method (*26*) for homonuclear 2D experiments or with the States-TPPI methods (*27*) for heteronuclear experiments. Gradient pulses as a combination of *x*- and *z*- gradients at the "magic angle" (*28*, *29*) were used for coherence selection and sensitivity enhancement, which resulted in improved suppression of the water solvent signal in all 3D NMR experiments. Experimental details are summarized in Table S1 of the Supporting Information. Solvent exchange of amide protons was monitored by reference to a series of two-dimensional ^{1}H-^{15}N HSQC data sets collected at 25 °C. Each HSQC spectrum took 30 min to collect.

The raw NMR data were processed with FELIX95 (Molecular Simulations Inc., San Diego, CA). In most cases, the indirectly detected dimensions were extended by linear prediction. Typically, a 90°-shifted squared sine-bell window function was applied to each FID prior to being zero-filled, Fourier transformed, and phase corrected.

All ^{1}H dimensions were referenced to internal 4,4-dimethyl-4-silapentane-1-sulfonate (DSS), and ^{13}C and ^{15}N were indirectly referenced to DSS (*30*).

Derivation of Structural Restraints. Interproton distance restraints were derived from NOEs assigned in 2D ^{1}H-^{1}H NOESY (mixing time 100 ms) for aromatic ring protons, 3D ^{15}N-edited NOESY (mixing times 50 and 150 ms) for amide protons, and 3D ^{13}C-edited NOESY (mixing time 100 ms) data for aliphatic protons. NOE restraints were grouped into three distance ranges: strong, 1.8−2.7 Å (1.8−2.9 Å for NOEs involving H^N protons); medium, 1.8−3.3 Å (1.8−3.5 Å for NOEs involving H^N protons); and weak, 1.8−5.0 Å, respectively, according to NOE peaks intensities. An additional 0.5 Å was added to the upper bound of the NOE restraints for methyl protons (*31*). Pseudo-atom corrections were made for nonstereospecifically assigned methylene and methyl resonances (*32*). In addition, dihedral angle restraints were derived from $^3J_{NH\alpha}$ values obtained from a 3D HNHA experiment. The ϕ angles were restrained to $-60 \pm 30°$ for $^3J_{HNH\alpha} < 6.0$ Hz and $d_{NN} > d_{\alpha N}$, $-120 \pm 50°$ for $^3J_{HNH\alpha} = 8.0-9.0$ Hz, and $-120 \pm 30°$ for $^3J_{HNH\alpha} > 9.0$ Hz. An additional ϕ angle restraint of $-100 \pm 80°$ was applied to residues for which the intraresidue $H^\alpha-H^N$ NOE was clearly weaker than the NOE between H^N and H^α of the preceding residue (*33*). A ψ angle restraint was used for residues in α-helix and β-strand structures, as predicted from the chemical shift index (*34*) and from NOE patterns characteristic of secondary structure. The ψ angle was restrained to $-40 \pm 30°$ in α-helix regions and $+120 \pm 60°$ in β-sheet regions. χ^1 angles were determined from $^3J_{H\alpha-H\beta}$ values estimated qualitatively from 3D TOCSY−HSQC with a short-mixing time in combination with N^H-H^β and $H^\alpha-H^\beta$ NOEs from 3D ^{15}N-edited or ^{13}C-edited NOESY-HSQC (*35*). The χ^1 angles were restricted to $+60 \pm 60°$ for a g^+-conformation, $180 \pm 60°$ for a trans-conformation, and $-60 \pm 60°$ for a g^--conformation. Hydrogen bond constraints

[2] Studts, J. M., and Fox, B. G., unpublished results.

were identified from the pattern of sequential and interstrand NOEs involving H^N and H^α protons, CSI analysis of the slowly and very slowly exchanging amide protons, and the first preliminary structures. Peaks designated as "slowly" exchanging remained visible in the ^1H-^{15}N HSQC spectra 2 h after exchange of the protein into D_2O; peaks designated as "very slowly exchanging" remained observable 12 h after exchange. Each deduced hydrogen bond was represented by two distance constraints: 1.7− 2.4 Å for H^N-O and 2.7− 3.4 Å for N−O.

Structure Calculation. Initial structures were calculated by the hybrid distance geometry-simulated annealing protocol in X-PLOR 3.843 (*36*) and refined by the torsion angle molecular dynamics protocol in DYANA 1.5 (*37*). For the calculations using X-PLOR, the following force constants were used: 50 kcal $mol^{-1}Å^{-2}$ for the distance restraints throughout all the calculations; 5 kcal $mol^{-1}rad^{-2}$ for the initial dihedral angle restraints during the high-temperature dynamics, and gradually increasing to 200 kcal $mol^{-1}rad^{-2}$ during the annealing stage. All calculations were performed on a Silicon Graphics (Mountain View, CA) Onyx computer equipped with 10 R10000 processors. The structure calculation proceeded in three stages. In the first stage, a low-resolution structure of T4moD was generated by X-PLOR, using NOE-derived distance restraints and dihedral angle restraints for ϕ and χ^1 angles. In the next stages, hydrogen bond restraints and ψ-angle restraints were added on the basis of the structures generated in the initial stage. During this stage, each round started with 20 structures generated by substructure embedding and regularization in X-PLOR or with 60 randomized conformers in DYANA. The 20 calculated structures from X-PLOR or the 20 structures from DYANA with the lowest target functions were used to analyze restraint violations and to assign additional NOE restraints for the following round. This process was repeated until all peaks in the spectra had been assigned and all violations were eliminated. In the final stage, 60 structures were calculated by DYANA, and of these, the 20 structures with the lowest target functions were used for structural analysis. The mean structure calculated by MOLMOL (*38*) from this ensemble of 20 structures was minimized by DYANA with 4500 steps of conjugated gradient minimization. Analyses such as RMSD, secondary structure and dihedral angle determinations, and preparation of multiple alignments of the T4moD structure were performed with MOLMOL and PROCHECK−NMR (*39*). The ribbon diagrams of T4moD and other effector proteins were prepared with RasMac 2.6b (*40*) and MOLSCRIPT (*41*).

Sequence Analyses. Sequence alignments and inferred phylogenetic trees were produced using Biology Workbench 3.2 (San Diego Supercomputer Center, http://workbench.sdsc.edu). Primary sequences for the effector proteins were obtained from the following GenBank files: TbuV from *Pseudomonas pickettii* PKO1 toluene 3-monooxygenase, U04052; PhlN from *Ralstonia eutropha* JMP134 phenol hydroxylase, AF065891; BmoD1 from *Pseudomonas aeruginosa* JI104 toluene 2-/benzene monooxygenase, D83068; TbhD from *Burkeholderia cepacia* toluene 3-monooxygenase, AF001356; T4moD, from *Pseudomonas mendocina* KR1 toluene 4-monooxygenase, M65106; TouD from *Pseudomonas stutzeri* OX1 toluene 2-/*o*-xylene monooxygenase, AJ005663; XamoD from *Xanthobacter* species Py2 propene epoxidase, AJ012090; AmoB from *Nocardia corallina* B-276 propene epoxidase, D37875; MmoBt from *M. trichosporium* OB3b methane monooxygenase, X55394; MmoBM from *Methylocystis* species M methane monooxygenase, U81596; MmoBc from *M. capsulatus* (Bath) methane monooxygenase, M90050; MmoBK from *Methylomonas* species KSWIII methane monooxygenase, AB025022; PhyB from *Ralstonia* species KN1 phenol hydroxylase, AB031996; PoxC from *Ralstonia* species E2 phenol hydroxylase, AF026065; Orf3 from *Acinetobacter calcoaceticus* NCIB8250 phenol hydroxylase, Z36909; DmpM from *Pseudomonas putida* CF600 phenol hydroxylase, M60276; PheA3 from *P. putida* phenol hydroxylase, D28864; PhhM from *P. putida* P35X phenol hydroxylase, X79063; PhlC from *P. putida* phenol hydroxylase, X80765; TbmC from *Pseudomonas* species JS150 toluene 2-/*o*-cresol hydroxylase, L40033; CrpC from *Ralstonia pickettii* phenol hydroxylase, AF012632; AphM from *Comamonas testosteroni* TA441 phenol hydroxylase, AB006479; and PhcM from *C. testosteroni* R5 phenol hydroxylase, AB024741. Percentage identities between T4moD and other effector proteins were calculated using the program ALIGN and the default parameter set (*42*−*44*). Multiple sequence alignments were carried out using CLUSTALW and the default parameter set (*45*, *46*) and then adjusted by inspection of the available NMR structures. The phylogenetic tree was produced from the aligned sequences using DRAWGRAM and the default parameter set (*47*).

Catalytic Studies. A unit of T4moD activity is defined as the formation of 1 μmol of *p*-cresol/min in air-saturated 50 mM phosphate buffer, pH 7.5, at 298 K in the presence of optimal concentrations of the other T4MO components, toluene (5.8 mM at 298 K) and NADH (0.5 mM). *p*-Cresol was determined using gas chromatography and conditions published elsewhere (*7*, *48*, *49*). For NMR studies, T4moD activity was measured before and after data collection. In catalytic complementation assays, TbuV, TbmC, or S1 were substituted for T4moD and varied relative to fixed concentrations of the other T4MO components to determine the maximal possible stimulation of *p*-cresol formation. For steady-state kinetic analysis, the effector protein was treated as a substrate, and all other protein components of the T4MO complex, O_2, toluene, and NADH were maintained at saturating levels. Rates were determined by linear least-squares fitting of the time dependence of product formation in a manner similar to that described elsewhere (*50*). The data for *p*-cresol formation velocity versus effector protein concentration were fit to an equation accounting for the formation of both activating and inhibiting complexes of the effector protein, $v = V_{max}[S]/[K_M + [S] + ([S]^2/K_I)]$, where V_{max} is the maximal reaction velocity, S is the concentration of effector protein, K_M is the Michaelis constant for formation of the activating complex of effector protein, and K_I is the equilibrium constant for formation the inhibiting complex of effector protein. The kinetic parameters and error estimates were determined using the statistically weighted nonlinear least-squares fitting routine SUBINO (*51*). V_{max}-values are reported as turnover numbers relative to diiron center in the ($\alpha\beta\gamma$) protomer of T4moH.

RESULTS

Status of NMR Samples. In the previous NMR studies of diiron hydroxylase effector proteins, aggregation of MmoB

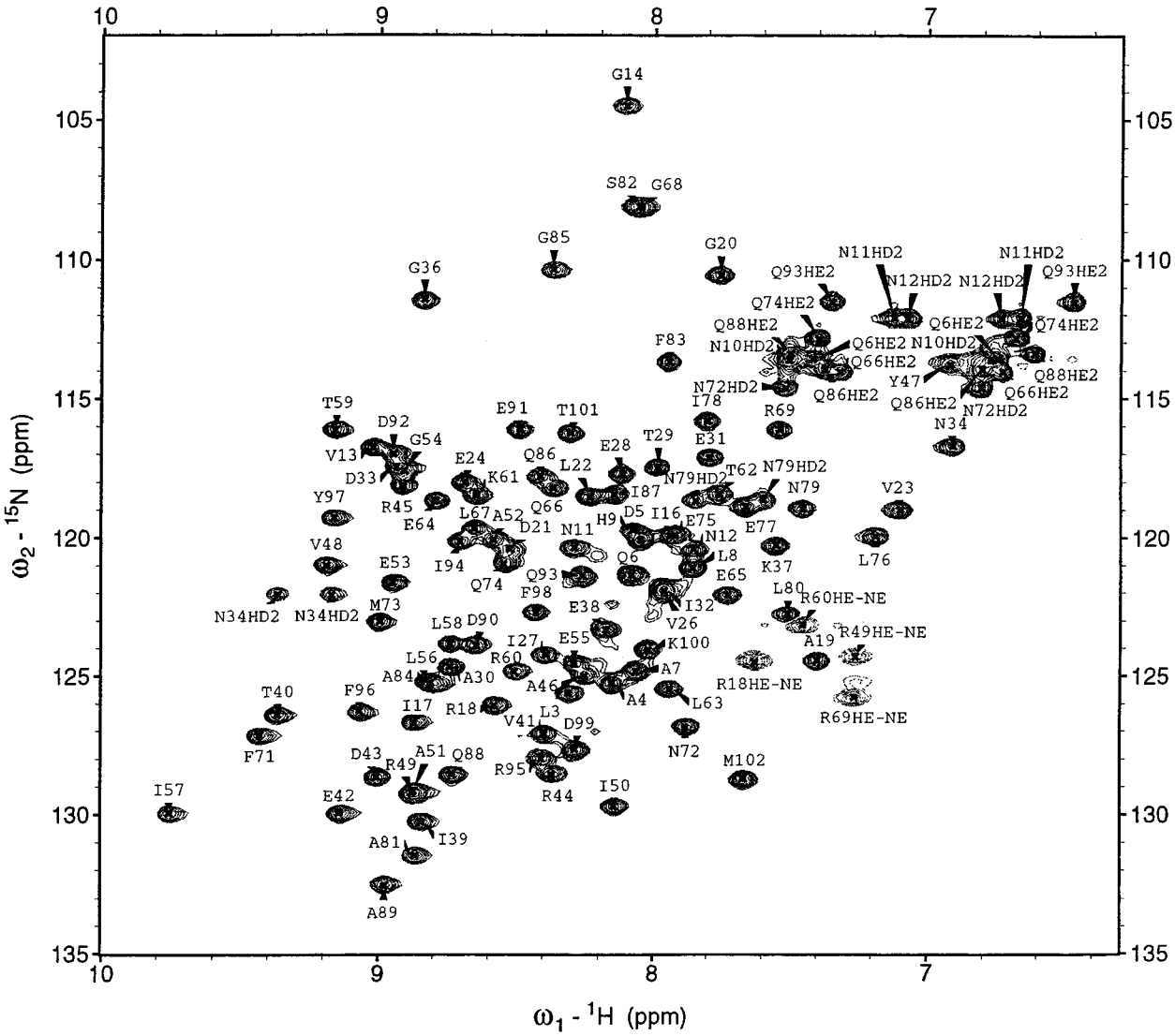

FIGURE 1: 2D ^{1}H-^{15}N HSQC spectrum at 500 MHz of 1.1 mM T4moD obtained at pH 7.0 and 298 K. See Table S1 of the Supporting Information for experimental details. The assignments shown were determined from analysis of through-bond connectivities.

from *M. trichosporium* OB3b was observed at concentrations exceeding 0.2 mM (*22*), while the NMR spectra of DmpM from *P. putida* CF600 exhibited numerous low intensity peaks (*21*). At the concentrations used for the NMR experiments (1.1−5.4 mM), none of the proteins studied here (T4moD, TbuV, or TbmC) showed these complications. NMR samples of T4moD were stable for extended periods at 298 K and exhibited the same ability to stimulate the toluene hydroxylation activity of the reconstituted T4MO complex both before and after NMR data collection.

^{1}H, ^{15}N, ^{13}C Resonance Assignments. Sequence-specific assignments of the polypeptide backbone resonances were determined from ^{1}H-^{15}N HSQC, HNCA, HNCO, HN(CO)-CA, C(CO)NH, and HNCACB spectra by matching intra- and interresidue $^{13}C^{\alpha}$ and/or $^{13}C^{\beta}$ chemical shifts for a given residue (associated with a ^{1}H-^{15}N correlation peak) and the previous residue in the primary sequence. Representative results are shown in Figure S1 of the Supporting Information. Figure 1 shows the ^{1}H-^{15}N HSQC spectrum for T4moD at pH 7.0 and 298 K with 123 assigned ^{1}H-^{15}N cross-peaks (95% complete) out of the expected 129 cross-peaks, which also verifies the overall quality and homogeneity of the spectroscopic sample. The backbone amide protons of Ser1, Thr2, and Asn10, and side-chain amide protons of three arginine residues were not observed in the HSQC spectrum or in the data from 3D triple-resonance experiments. Other backbone atoms, C^{α}, H^{α}, and C, were assigned, except for the C^{α} and H^{α} of Ser1, and the six carbonyl carbons of His9, Gly14, Glu24, Asn34, Arg69, and Met102. Signals were not detected from the N- and C-terminal residues, Ser1 and Met102, probably because the two ends undergo motions leading to exchange broadening of the NMR peaks. The lack of detected signals from the H^N of His9 and Asn10 was also probably the result of local dynamics. The carbonyl carbon resonances of Gly14, Glu24, Asn34, and Arg69 were not determined because the following residue in each case is proline.

Side-chain resonances were determined from analysis of C(CO)NH, HNCACB, HC(CO)NH, HCCH−COSY, and HCCH-TOCSY data. Assignments for the side chains of the six Asn and six Gln residues were confirmed by results from the HNCACB experiment with the delay values optimized for amide NH_2 groups. The ring resonances of the tyrosines, phenylalanines, and histidines were assigned by analyzing 2D ^{1}H-NOESY and TOCSY in D_2O, CT-^{13}C−HSQC, and 3D NOESY-CT-HSQC and TOCSY−CT-HSQC data. About

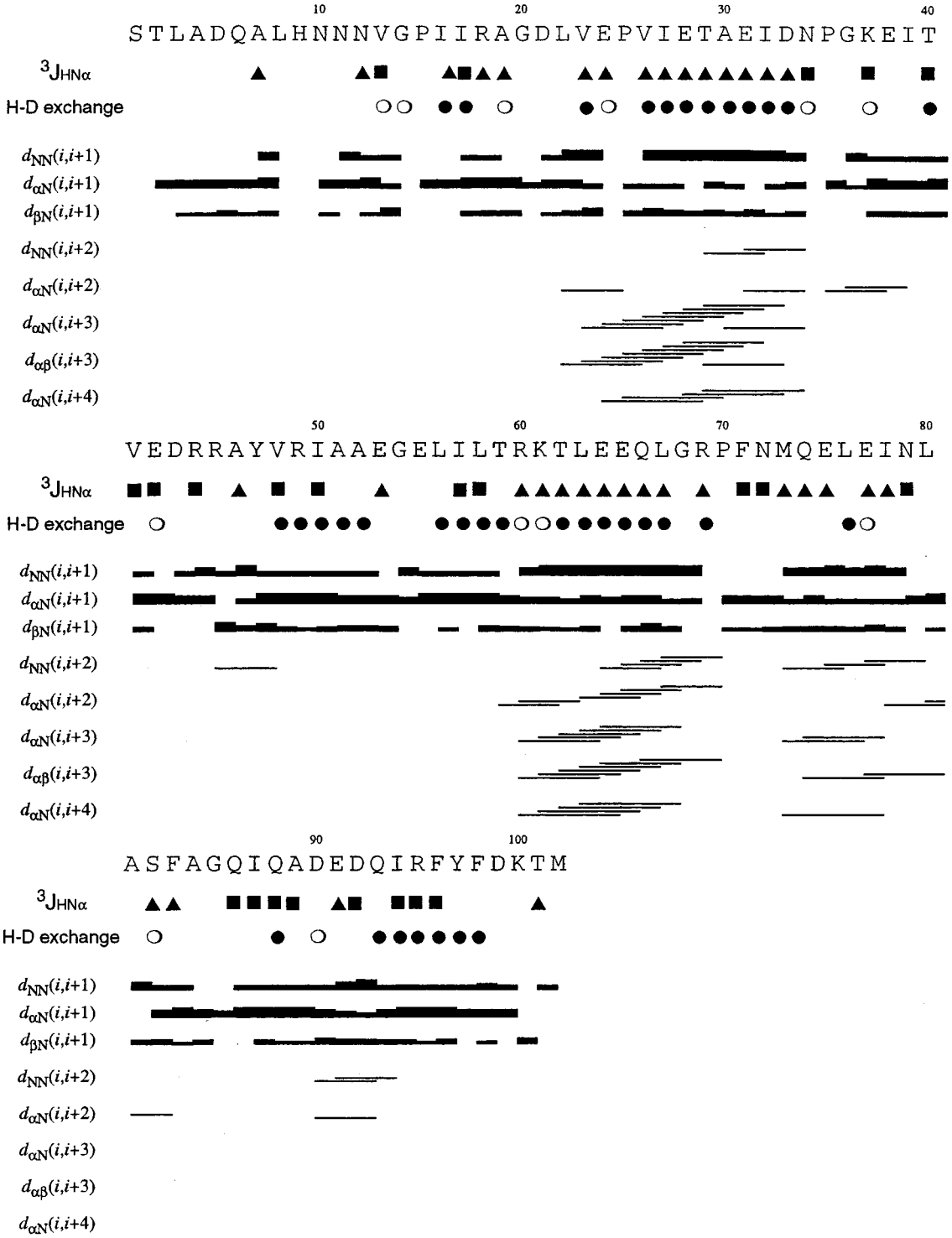

FIGURE 2: Summary of the sequential NOE connectivities observed for T4moD. Bars indicate NOESY cross-peaks observed between two residues. The heights of the bars indicate the strength of the NOE. Also indicated are slowly exchanging (○) and very slowly exchanging amide protons (●) and $^3J_{HN-H\alpha}$ values of <6.0 Hz (□) and >8.0 Hz (■).

88% of the resonances expected from the 901 side-chain atoms were assigned for T4moD. ^{1}H, ^{15}N, and ^{13}C resonance assignments are provided in Table S2 of the Supporting Information; chemical shifts for T4moD have also been deposited with the Biological Magnetic Resonance Bank [accession no. 4560 (*52*)].

Secondary Structure. Qualitative analyses of the short- and medium-range NOEs, $^3J_{HN-H\alpha}$ coupling constants, and slowly exchanging amide protons (Figure 2) served to identify the secondary structural elements of T4moD. The pattern of the $H^{\alpha}_i-H^{N}_{i+2}$ and H^{α}_i-H^{N}_{i+3} NOEs with $^3J_{NH-H\alpha}$ coupling constants smaller than 6 Hz indicated the presence of three α-helices (Val23−Ile32, Arg60−Leu67, and Met73−Glu77). Coupling constants greater than 8.0 Hz, characteristic long-range NOEs ($H^{\alpha}_i-H^{\alpha}_j$; $H^{\alpha}_i-H^{N}_j$), and slowly exchanging amide protons were used to define the β-strands. On the basis of these analyses, one four-stranded antiparallel β-sheet (Asn12−Arg18, Thr40−Glu42, Tyr47−Glu53, Ser82−

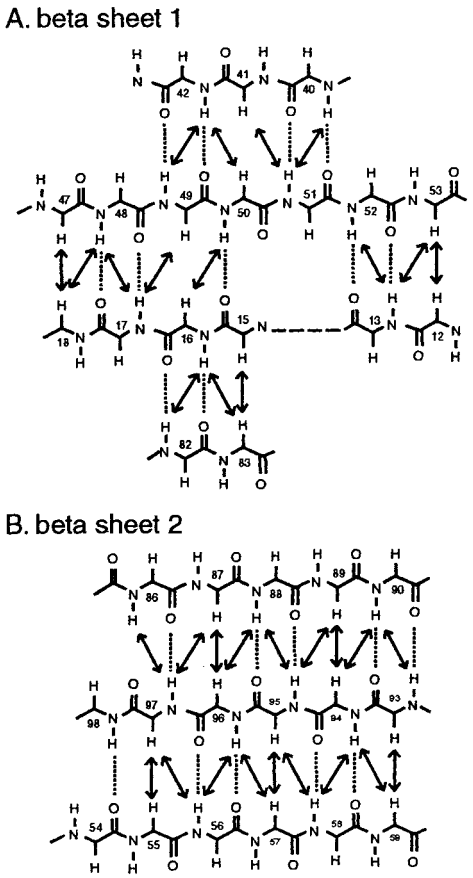

FIGURE 3: Schematic showing the β-sheet structure of T4moD. Solid arrows indicate observed NOEs, and dotted lines indicate potential hydrogen bonds. (A) Four-stranded antiparallel β-sheet consisting of Asn12−Arg18, Thr40−Glu42, Tyr47−Glu53, and Ser82−Phe83. (B) Three-stranded antiparallel β-sheet consisting of Gly54−Thr59, Gln86−Asp90, and Gln93−Tyr97.

Phe83) and one three-stranded antiparallel β-sheet (Gly54−Thr59, Gln86−Asp90, and Gln93−Tyr97) were identified (Figure 3). The location and the extent of the secondary structural elements were further supported by the analysis of $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ chemical shifts (*34*). The consensus results of the chemical shift index (CSI) of $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ in T4moD (Figure 4) were in good agreement with the secondary structural elements derived from other results.

Tertiary Structure of T4moD. A total of 1467 interproton restraints were used for the structure calculation. These comprised 367 intraresidue, 424 sequential, 195 medium-range, and 273 long-range NOE distance restraints and 128 dihedral angle restraints (78 ϕ, 26 ψ, and 24 χ^1). In combination with NOE patterns and the preliminary structures, 80 hydrogen bond restraints (two for each of the 40 hydrogen bonds) identified from the slowly exchanging amide protons were introduced at the second stage. At the final stage, the structure of T4moD was refined by DYANA with input from a total of 1467 experimental restraints (>14 restraints/residue); 60 conformers were calculated. Of these, the 20 conformers with the lowest target functions were selected to represent the structure (Figure 5). The magnitude of the average target function (0.77 ± 0.14 Å^2) indicates good agreement between the selected 20 structures and the input restraints. For the defined region of the structure (Asn12−Phe98), the RMSDs to the averaged structure were 0.71 Å for the backbone atoms and 1.24 Å for all non-

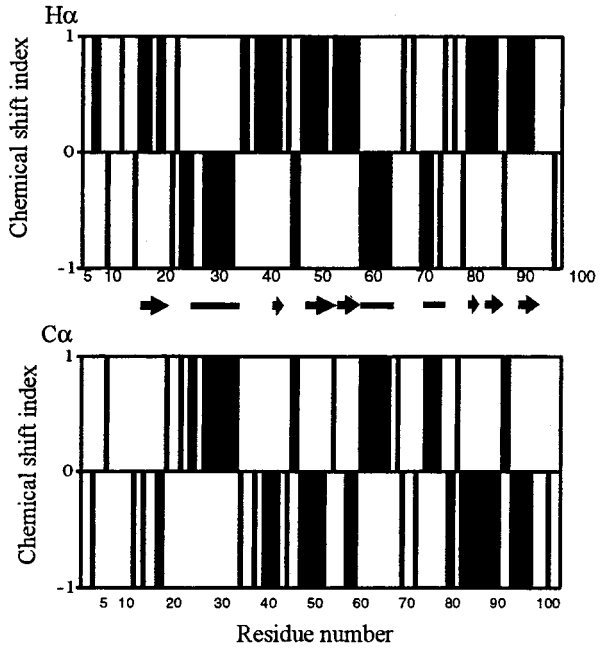

FIGURE 4: Chemical shift index (*45*) derived from $^{1}H^{\alpha}$ and $^{13}C^{\alpha}$ chemical shifts of T4moD. An index of either −1 or +1 indicates a shift deviation from the random coil value of either greater than 0.1 ppm ($^{1}H^{\alpha}$) to lower frequency or 0.7 ppm ($^{13}C^{\alpha}$) to higher frequency. Chemical shifts within the variance of the random coil values defined above are indicated by 0.

FIGURE 5: Superimposition of 20 selected conformers with the lowest target function from the final DYANA calculations that represent the structure of T4moD.

hydrogen atoms. The energy-minimized mean structure had the target function of 0.70 Å^2, and the RMSDs of Asn12−Phe98 were 0.85 Å for the backbone atoms and 1.51 Å for all non-hydrogen atoms. The Ramachandran plot of the ϕ and ψ angles for the 20 conformers showed 65.6% of the ϕ

Table 1: Structural Statistics of the Best 20 DYANA Conformers

(A) NMR-derived restraints	
total interproton restraints	1339
intraresidue	367
sequential ($\|j - i\| < 2$)	424
medium range ($1 < \|j - i\| < 5$)	195
long-range ($\|j - i\| > 4$)	273
hydrogen bonds	80
dihedral angles (ϕ, ψ, χ^1)	128
(B) residual violations[a]	
DYANA target function (Å^2)	0.77 (0.14)
upper limit	
sum (Å)	4.2 (0.5)
maximum (Å)	0.21 (0.06)
lower limits	
sum (Å)	0.1 (0.1)
maximum (Å)	0.05 (0.04)
van der Waals	
sum (Å)	2.8 (0.4)
maximum (Å)	0.21 (0.02)
torsion angles	
sum (radians)	0.0 (0.0)
maximum (radians)	0.02 (0.01)
(C) average RMSD to mean structure (Å)	
N, C^α, CO of residues 12−98	0.71 (0.16)
all heavy atoms of residues 12−98	1.24 (0.17)
(D) Ramachandran plot (% residues)	
residues in most favored regions	65.6
residues in additional allowed regions	29.1
residues in generously allowed regions	4.9
residues in disallowed regions	0.4

[a] In part B, values are mean (standard deviation).

and ψ angles to be in the most favored regions, 29.1% in the additional allowed regions, and 4.9% in generously allowed regions. Only 0.4% were found in the disallowed region. The structural statistics are summarized in Table 1.The structure of T4moD contains three α-helices and seven β-strands (see the ribbon diagram in Figure 6). The first β-sheet consists of four strands: β1 (Asn12−Arg18), β2 (Thr40−Glu42), β3 (Tyr47−Glu53), and β5 (Ser82−Phe83). The sequence between strands β1 and β2 forms a long helix, α1 (Val23−Ile32) and an unstructured loop region. The second sheet consists of strands β4 (Gly54−Thr59), β6 (Gln86−Asp90), and β7 (Gln93−Tyr97). The sequence between strands β4 and β5 forms two helices, α2 (Arg60−Leu67) and α3 (Met73−Glu77). Helix α2 is in close contact with α1 as judged by eight NOEs between them. β5 is the final strand of the first sheet and has two hydrogen bonds to β1. Although residues Arg69−Leu80 including α3 are not well-defined, it appears that the side chain of Leu76 in α3 forms a hydrophobic interaction with that of Ile94 in β6. Strands β6 and β7 are connected by a β-hairpin (Asp90−Gln93). The last strand β7 is hydrogen-bonded to strands β4 and β6. The N-terminal residues Ser1−Asn11 and C-terminal residues Asp99−Met102 of T4moD are unstructured as judged by the lack of NOEs and CSI criteria for regular secondary structure. Figure 6 also shows the location of Asn34, Ile57, and Glu64, which will addressed below.

Sequence Analysis of Effector Proteins in Terms of Structure. Figure 7 shows a primary sequence alignment of diiron hydroxylases effector proteins beginning with the first β-strand revealed by the NMR structures. Figure 8 shows a rooted phylogenetic tree inferred from the alignments of Figure 7. All available sequences of effector proteins, including those from hydroxylase complexes specialized for reaction with benzene, toluene, phenol, methylphenols, methane, and propene have been included in these comparisons. Among the 24 effector proteins analyzed, the percentage of amino acid sequence identity with T4moD ranged from 18 to 58%. In the sequence alignments, a gap insertion of five to six amino acids was required to maximally align the sequences of the larger MmoB proteins with those of the aromatic hydroxylase effector proteins. Furthermore, a gap insertion of two amino acids just prior to β2 was required to maximally align the sequences of the phenol hydroxylase effector proteins and the MmoB proteins with those of the toluene/benzene hydroxylase effector proteins. With few exceptions, Figure 8 shows a remarkable correlation between the sequences of the effector proteins, their inferred phylogenetic relationships, and the chemical nature of the substrates oxidized. The only proteins that do not fit in well are PhlN, which is induced by and reacts with phenol (*53*), and TbmC and S1, which are part of enzyme complexes that react with both toluene and *o*-cresols (*6*, *54*).

Catalytic Specificity of Toluene Monooxygenase Effector Proteins. The following other purified effector protein family members were tested in place of T4moD in reconstituted *P. mendocina* KR1 T4MO complexes: TbuV (recombinant toluene 3-monooxygenase effector protein from *P. pickettii* PKO1); TbmC (toluene 2-monooxygenase effector protein from *Pseudomonas* species JS150); and S1 [toluene 2-monooxygenase/2-methylphenol 6-hydroxylase from *B. cepacia* G4 (*6*)]. The more distantly related TbmC and S1 were unable to replace the catalytic function of T4moD, whereas TbuV was able to substitute for T4moD in the T4MO enzyme complex (Figure 8). Figure 9 shows a comparison of the dependence of reaction velocity on the concentration when either T4moD or TbuV were used in standard T4MO assays. T4moD and TbuV each formed activating and inhibiting complexes of the type regarded as hallmarks of effector protein function (*13*). For T4moD, an apparent $V_{max} = 3.4 \pm 0.5$ s^{-1} and apparent $K_M = 2.0 \pm 0.6$ μM were determined from weighted nonlinear least-squares fitting (solid line of Figure 9), while an apparent $V_{max} = 2.4 \pm 0.2$ s^{-1} and apparent $K_M = 2.5 \pm 0.4$ μM were determined for TbuV (dashed line of Figure 9). *V/K* comparison shows that the T4MO complex has an ∼2-fold preference for T4moD relative to TbuV, with small changes in both catalytic and binding parameters contributing to the preference. T4moD and TbuV also formed similar inhibitory complexes with the T4MO enzyme complex, with apparent K_I-values of 28 ± 10 μM and 32 ± 5 μM, respectively.

Resonances from $^1H^{\delta 21}$ and $^1H^{\delta 22}$ of Asn34. The effector proteins aligned in Figure 7 contain a consensus sequence of Asp-Asn-Pro, with Asp conserved in 18 of 24 sequences and Asn and Pro conserved in all instances except one, where a Gln residue is substituted for Asn. In T4moD, Asp33, Asn34, and Pro35 are located in the loop between helix α1 (the longest helix in the protein) and strand β2. The frequencies of the Asn34 $^1H^{\delta 21}$ and $^1H^{\delta 22}$ resonances (9.29 and 9.49 ppm) are 3σ higher than the average of chemical shifts assigned to amide groups in other proteins (http://www.bmrb.wisc.edu/data_access/outlier_selection_grid.html). These Asn amide protons were found to be unusually resistant to exchange with solvent water; their signals were still visible in a ^{1}H-^{15}N−HSQC spectrum taken 2 h after a sample of T4moD sample lyophilized from H_2O was resuspended in D_2O. This suggests that the side-chain

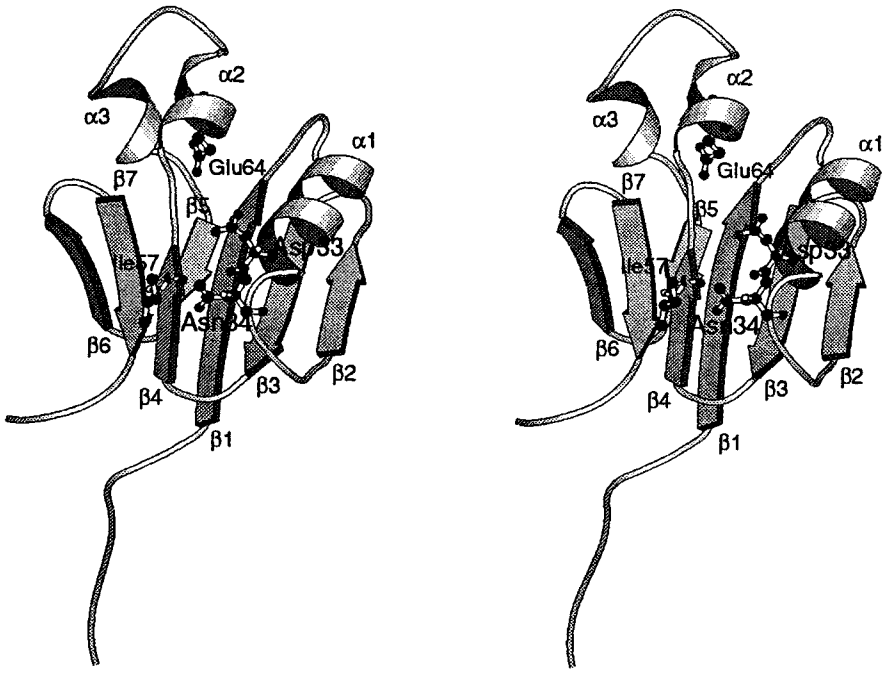

FIGURE 6: Divergent stereoview of the minimized, average structure of T4moD as a ribbon diagram annotated with the identity of assigned α-helices and *β*-strands. The locations of the side chains of Asp33, Asn34, Ile57, and Glu64 are shown. The representation of T4moD shown in Figure 6 can be obtained from 1G10 by successive 148° rotation about the *z*-axis, −30° rotation about the *y*-axis, and 153° rotation about the *x*-axis. This representation can be transformed to the representation shown in Figure 11A by successive −45° rotation about the *x*-axis, 20° rotation about the *y*-axis, and 10° rotation about the *z*-axis.

NH_2 of Asn34 forms structurally relevant hydrogen bonds.

As shown in Figure 10, Asn34 $H^{\delta 21}$ of T4moD is hydrogen bonded to Asp33 O (2.84 Å), while Asn34 $H^{\delta 22}$ is hydrogen bonded to Ile57 O (2.26 Å) of strand *β*4. These interactions likely contribute to the compact structure of T4moD by linking the N-terminal *βαββ* and C-terminal *βααββββ* domains. Neither of the NMR investigations of MmoB reported signals from the side-chain amide protons of the asparagine homologous to Asn34 in T4moD. However, in the structure of MmoB from *M. capsulatus*, the comparable Asn65 $H^{\delta 21}$ is within 2.7 Å of Lys84 O of strand *β*4 (*22*, *23*), which suggests that this protein has an analogous interdomain hydrogen-bonding interaction. In T4moD, the carboxylate side chain of Asp33 is ∼2.9 Å from the guanidinium group of Arg60. Since Arg60 is located at the end of helix α2, this interaction provides another interdomain contact that may contribute to the compact structure of T4moD.

Preliminary NMR Studies of TbuV and TbmC. Because of the unusual chemical shifts of the side-chain amide signals of Asn34 and the strict conservation of this residue, it was of interest to determine whether similar unusual chemical shifts might be found in ^1H NMR spectra of other aromatic hydroxylase effector proteins. The NOESY spectrum of TbuV (effector protein of the toluene 3-monooxygenase from *P. pickettii* PKO1, which contains residues homologous to Asn33 and Ile57) exhibited cross-peaks at 9.12 and 9.28 ppm comparable to those assigned to the interaction between $^1H^{\delta 21}$ and $^1H^{\delta 22}$ of Asn34 and Ile57 $^1H^N$ in the NOESY spectrum of T4moD.[3] By contrast, the NOESY spectrum of TbmC (the effector protein of the toluene 2-monooxygenase from *Pseudomonas* species JS150, which also contains residues homologous to Asn33 and Ile57) did not exhibit comparable ^1H resonances. These results indicate that the structure at the conserved Asn group is different in TbmC from that in T4moD and TbuV. Since this difference correlates with catalytic complementarity in the T4MO system, it is tempting to speculate that the region around Asn34 is a determinant of this specificity.

DISCUSSION

Secondary and Tertiary Structure. T4moD, which consists of an N-terminal *βαββ* domain and a C-terminal *βααββββ* domain, has a topology similar to that determined for MmoB proteins from two methanotrophs (*22*, *23*). This topology differs from that determined for what appears to be a partially folded form of effector protein DmpM, namely an N-terminal *βαβ* domain and a C-terminal *βααββββ* domain (*21*, *24*). As noted in several contexts (*21*−*23*), the amino acids that define the arrangement of strands and helices in DmpM, MmoBs, and T4moD are conserved nearly universally among the known effector proteins (Figure 7). For T4moD, this includes Val13 in strand *β*1, the sequence of Asp33−Asn34−Pro35 that immediately follows helix α1, Ala46 that immediately precedes strand *β*3, Gly54 that begins strand *β*4, Arg60 that begins helix α2, Gln74 in helix α3, Gly85 that immediately precedes strand *β*6, and Asp90 that ends strand *β*6. Thus, it is likely that the secondary structural topology of the effector proteins will be conserved. However, Figure 11 shows that the average 3D structures of the four effector proteins determined by NMR differ substantially from one another. In the following, we consider how various aspects of effector protein structure may contribute to this difference and discuss how catalytic specialization may be produced.

[3] Spectra of TbuV revealed an addition $^1H^N$ signal at 9.67 ppm with an NOE to the $^1H^N$ of Ile57. The assignment of this 9.67 ppm peak is under investigation; an analogous resonance was not observed in spectra of T4moD.

FIGURE 7: Primary sequence alignment of the effector proteins from diiron hydroxylases. Abbreviations and GenBank accession numbers are listed in the Materials and Methods. The number after each abbreviation indicates the number of amino acids in the N-terminal region of the protein that are not shown in the sequence alignment. Boxed positions have greater than 50% identity among all sequences. Secondary structural elements identified in the NMR structures of T4moD, MmoB from *M. trichosporium* OB3b and *M. capsulatus* (Bath), and DmpM from *P. putida* CF600 are indicated above the relevant sequence. See refs *21* and *23* for related alignments.

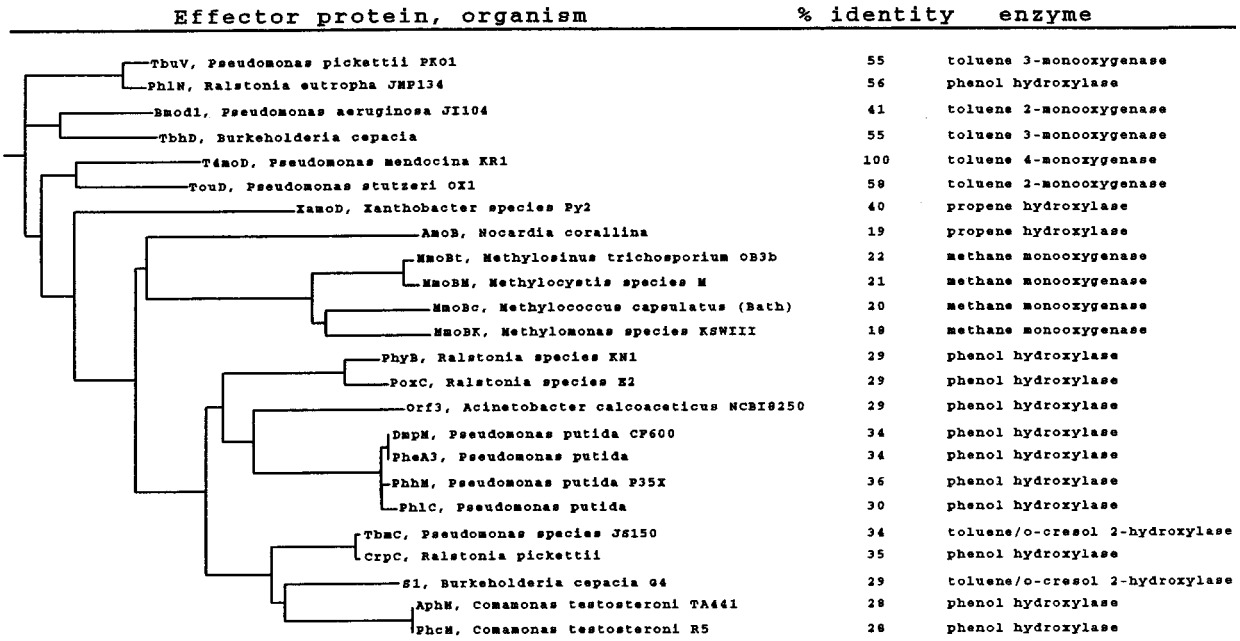

FIGURE 8: Rooted phylogenetic tree inferred from the primary sequence alignment of Figure 7. The organism, percentage of primary sequence identity with T4moD, and the name of the enzyme complex are indicated.

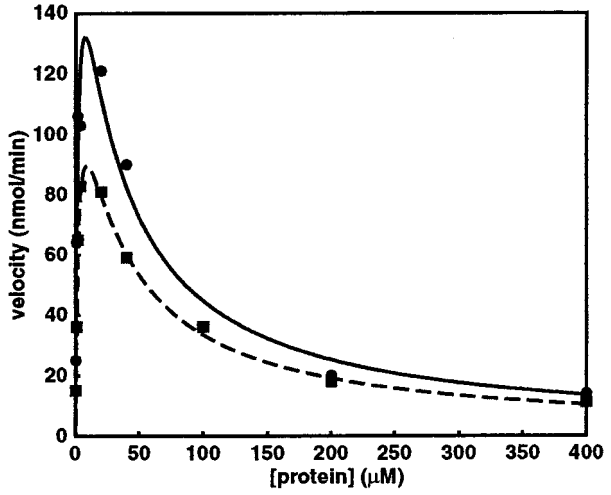

FIGURE 9: Analysis of the velocity of *p*-cresol formation versus effector protein concentration for reconstitution of the T4MO complex with either T4moD (●, solid line) or TbuV (■, dotted line) The lines were generated by nonlinear least-squares fitting as described in Materials and Methods.

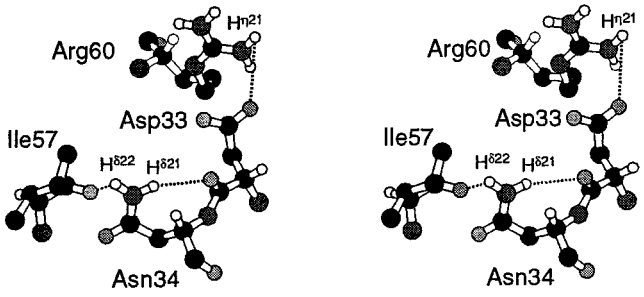

FIGURE 10: Divergent stereoview of the hydrogen-bonding interactions of Asn34 and associated residues in T4moD.

Properties of the N-Terminal Region of T4moD. The diiron hydroxylase effector proteins range in relative mass from ~10 to 17 kDa, with the major differences arising from the variable N-terminal region. The MmoB proteins differ from the other 20 effector proteins by containing an extended N-terminal region (~35 amino acids) containing a Gly triplet (GXXGXXG), whereas most members of the phenol hydroxylase branch apparently lack an extended N-terminal region. The structures of MmoB model this N-terminal extension as either ill-defined or weakly α-helical (*22*, *23*). Proteolytic cleavage within the Gly triplet sequence has been shown to result in decreased catalytic function (*55*). Since the proteolysis apparently does not lead to unfolding or destabilization of the truncated protein (*56*), the catalytic impairment appears to be associated with a loss of ability to effectively bind to the hydroxylase partner. Whether the unstructured (Ser1-Asn11) N-terminal region of T4moD is essential for function has not been addressed in these studies. However, due to the overall lack of sequence conservation, it is likely that this region of the aromatic monooxygenase effector proteins will not be directly required for function.

Side-Chain NH_2 Resonances of Asn34. The NMR structures of the four effector proteins are compared in Figure 11. In T4moD (Figure 11A), the four-stranded and three-stranded β-sheets are roughly perpendicular to one another and have a "hinge-like" framework. The axis of helix α1 is parallel to the four-stranded β-sheet. The figure also presents the hydrogen-bonding interaction between Asn34 and Ile57 at the apex of the β-sheets and Asp33 and Arg60 at the upper end of the β-sheets. In MmoB from *M. capsulatus*, (Figure 11B), a putative hydrogen bond between Asn65 $H^{\delta 12}$ and Lys84 O and the hydrophobic interactions between α1 and α2 (described below) contribute to the similarity of this structure with T4moD (Figure 11A). In MmoB from *M. trichosporium*, (Figure 11C), the angle between the β-sheets has collapsed to ~70°, and the strands in the two sheets have twisted ~30° away from the orientation observed for the corresponding strands in T4moD. Consequently, Asn64 is partially exposed to solvent, and there are no interactions between helix α1 and strand β4. In the case of DmpM (Figure 11D), the β-sheets have opened to greater than

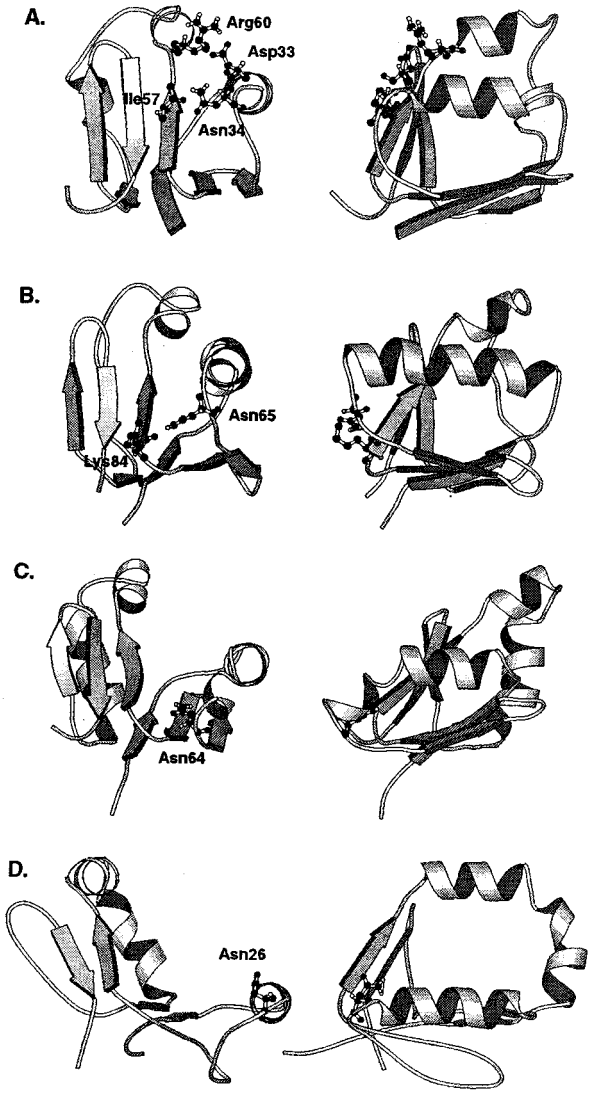

FIGURE 11: Comparison of the structures of four diiron hydroxylase effector proteins, all determined in solution by NMR spectroscopy: (A) T4moD, (B) MmoB from *M. capsulatus* (Bath) (*21*), (C) MmoB from *M. trichosporium* OB3b (*20*), (D) DmpM from *P. putida* CF600 (*19*). (Left) Views down helix α1 of each protein with the N-terminal closest to the viewer. Residues discussed in the text are shown in ball-and-stick representation. (Right) Rotation of the molecule by 90° so as to bring helix α1 to the front. For clarity the following residues have been omitted: 1−11 in Figure 11A; 1−35 in Figure 11B; and 1−34 in Figure 11C.

90° and twisted so that Asn26 is completely exposed to solvent.

Interactions between α1 and α2. T4moD exhibits NOEs from a hydrophobic interaction between the side chains of Val26 (helix α1) and Leu63 and Leu67 (both in helix α2). This interaction has a shape comparable to one "tooth" of a "leucine zipper" structure. These residues are identically conserved in TbuV, whereas other aromatic monooxygenase effector proteins have conserved subsets of two of these three hydrophobic residues (Figure 7). In the four MmoB sequences, the Val-Leu-Leu triad is substituted by (Ile/Val/Phe)−(Ala/Val/Ile)-Leu. In MmoB from *M. capsulatus*, a hydrophobic interaction is apparent between the phenyl ring of Phe51 (helix α1) and a methyl group of Ile92 (helix α2). A comparable interaction between helices α1 and α2 is not present in the structure of MmoB from *M. trichosporium*, owing to rotation of the β-sheets.

Patterns of Amino Acid Substitutions in Effector Proteins. The structures of T4moD and MmoB from *M. capsulatus* contain differences in helices α1, α2, and α3. Helix α1 of T4moD contains three solvent-exposed glutamate residues (Glu24, Glu28, and Glu31). *Glu24* of T4moD is substituted by Ala, Asp, Ile, or Ser in other members of the toluene/benzene monooxygenase branch of the effector protein phylogenetic tree, by either Asp or Glu in the propene epoxidase branch, by either Asp or Asn in the MmoB branch, and universally by Arg in the phenol hydroxylase branch (Figures 8 and 9). *Glu28* of T4moD is conserved identically in all but three of the aromatic monooxygenase effector proteins (each variant contains Asp at this position) and is conserved as Glu in all four MmoB examples (Figure 7). *Glu31* of T4moD is substituted by other polar amino acids in the toluene/benzene monooxygenase branch (Glu, Gln, and Arg) or by small hydrophobics in the propene epoxidase, methane monooxygenase, and phenol hydroxylase branches (Figures 8 and 9).

Functional Contributions of Helices α2 (Residues 60−67) and α3 (Residues 73−77). Helices α2 and α3 of T4moD contain a substantial number of residues that are identical or conservatively substituted in the complete family of effector proteins. In addition to Leu63 and Leu67, which provide a hydrophobic interaction with helix α1, Glu64, and Glu65 are conserved identically in many of the aromatic monooxygenase effector proteins. Glu65 of T4moD occupies a solvent-exposed position, whereas Glu64 of T4moD occupies an unusual internal hydrophobic cavity (Figure 6), with the nearest contacts to the side-chain carboxylate being Leu76 $H^{\delta 2}$ (∼2.4 Å), Ile94 $H^{\delta 2}$ (∼4.3 Å), Phe71 $H^{\epsilon 2}$ (∼4.7 Å), Leu63 $H^{\delta 2}$ (∼5.2 Å), and Leu67 $H^{\delta 1}$ (∼5.2 Å). These latter two residues form part of the conserved hydrophobic linkage between helices α1 and α2. The sequence of MmoB from *M. capsulatus* corresponding to helix α2 through helix α3 of T4moD also defines an internal cavity with Ile92 in the innermost position. This difference may be associated with the orientation of α*2* in the two proteins, as Leu63 of T4moD occupies a solvent exposed position.

In T4moD, the internal cavity containing Glu64 is topped by a long ridge of solvent-exposed amino acids (Glu65−Gly85, Figure 7). The cavity has entries on two sides. The entry on one side is contributed half by portions of helices α2 and α3, along with the intervening sequence between α2 and α3, and half by strands β5 and β6. The innermost surface of this opening is lined with hydrophobic residues, including Gly68, Leu80, Phe83, Ile87, Ala89, and Ile94, whereas Ile78 and polar amino acids Glu65, Glu77, Asn79, Gln86, Gln88, Asp90, Asp92, Arg95, and Tyr97 provide the outermost surface. The entry on the opposite side is defined by portions of strand β1 and helix α1 and by residues of the loop between them. This second opening is outlined by the polar side chains of Arg18, Asp21, Glu75, and Asn79, the carbonyl oxygens of Ala19, Gly20, Leu67, and Pro70, and the hydrophobic residues Leu22, Leu67, Pro70, and Phe71. It should be noted that several of these residues change charge or size in a uniform pattern in the sequences of the effector proteins from different catalytic subclasses of diiron hydroxylases.

Figure 7 shows that a substantial number of the conserved residues in the effector proteins are located in the region of

this cavity and the two openings, including all of the residues from *M. capsulatus* MmoB whose NMR resonances were reported to undergo broadening in the presence of the hydroxylase (*23*). Therefore, we propose that rearrangements of helices α2, α3, their intervening sequence, and perhaps more global shifts upon complex formation, may play an important role in catalysis. It is possible that these putative rearrangements can be readily accommodated by the presence of the internal cavity.

Catalytic Complementarity. Uniform substitution patterns identified at several places in the primary sequence alignments can be correlated provisionally with differences in functional properties (for example, compare the primary sequence patterns in strand *β*1 and in the sequence between *α2* and *β5*). We hypothesize that these changes may be matched by changes in the hydroxylase partners. Thus the protein interaction surfaces of each effector protein and its conjugate hydroxylase may have been optimized within a given organism by compensatory evolutionary changes. This optimization will likely result in minimal catalytic complementarity for effector proteins across the entire family of diiron hydroxylase complexes, despite the likelihood for overall similar protein structures and mechanisms of action. Moreover, since some of the species of bacteria included in Figures 7 and 8 contain more than one diiron hydroxylase with different substrate specificities under the control of a single regulatory system for toluene/benzene and cresol/phenol metabolism (*57*−*60*), the ability of effector proteins to react with only their appropriate hydroxylase partner may have important physiological consequences.

Prior to this work, the addition of purified preparations of both MmoB and reductase from *M. capsulatus* to an ion-exchange fraction of the hydroxylase obtained from *M. trichosporium* cell extracts (*61*) was reported to restore catalytic activity. Although the results of this experiment were not definitive, they suggested that closely related MmoB homologues could complement each other under catalytic conditions. Surprisingly, this complementarity was obtained even with the differences in 3D structure now indicated by Figure 11, panels B and C. The present investigation of catalytic complementarity among the effector proteins from four toluene monooxygenases evolved to have different hydroxylation regioselectivities (Figure 8) shows that the various effector proteins have substantially different abilities to complement T4MO (Figure 9). In agreement with the diversified catalytic function, our preliminary NMR assessments suggest that differences in catalytic complementarity may minimally correspond to a difference in interdomain H-bonding interactions at a critical "hinge-like" region of the three-dimensional structure. Whether these changes propagate to other regions of the protein structure, particularly upon complexation with the hydroxylase, remains to be determined. However, a reasonable hypothesis is that changes in the orientation of the *β*-sheets may correlate with rearrangements of helices α1, α2, α3, the internal cavity, and associated openings.

CONCLUSIONS

This work has provided a highly refined solution structure of T4moD, an effector protein from the aromatic monooxygenase subfamily of diiron hydroxylases. Evaluation of this structure and phylogenetic relationships has given new insight into subtle differences that may contribute to function and to the variable catalytic complementarity of the closely related homologues in this enzyme family. We have noted a correlation between unique H-bonding interactions at an interdomain "hinge" point and the ability of different effector proteins to complement the catalytic activity of the purified T4MO complex. Furthermore, the spatial juxtaposition of a well-defined internal cavity and numerous conserved surface residues suggests a possible mechanism to accommodate conformational changes upon binding of the effector protein to the appropriate hydroxylase partner. The combination of catalytic and spectroscopic approaches developed here for the aromatic monooxygenase effector proteins should prove useful in further investigations of these important questions in diiron hydroxylase structure and function.

ACKNOWLEDGMENT

We thank Dr. W. M. Westler and Dr. B. F. Volkman for advice on structure calculations and the use of DYANA, Dr. F. Abildgaard for assistance in acquiring and processing the NMR data, and Dr. E. S. Mooberry and other members of the National Magnetic Resonance Facility at Madison for assistance. We thank Prof. L. Wackett (University of Minnesota) for the generous gift of the purified S1 protein from toluene 2-monooxygenase from *B. cepacia* G4 used in the catalytic complementation experiments and for sharing the unpublished amino acid sequence of this protein. In addition, we thank Dr. A. M. Orville (University of Oregon) for many useful discussions on the relationship between the NMR and X-ray structures of T4moD as this manuscript was being prepared.

SUPPORTING INFORMATION AVAILABLE

Acquisition parameters for the NMR experiments (Table S1), chemical shift assignments (Table S2), and a representative through-bond sequential assignment of Ala81−Gly85 in T4moD from C(CO)NH and HNCACB data (Figure S1) are provided. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Fox, B. G. (1998) in *Comprehensive Biological Catalysis* (Sinnott, M., Ed.) pp 261−348, Academic Press, London.
2. Fox, B. G., Shanklin, J., Ai, J., Loehr, T. M., and Sanders-Loehr, J. (1994) *Biochemistry 33*, 12776−12786.
3. Coufal, D. E., Blazyk, J., Whittington, D. A., Wu, W. W., Rosenzweig, A. C., and Lippard, S. J. (2000) *Eur. J. Biochem. 267*, 2174−2185.
4. Green, J., and Dalton, H. (1985) *J. Biol. Chem. 260*, 15795−15801.
5. Fox, B. G., Froland, W. A., Dege, J. E., and Lipscomb, J. D. (1989) *J. Biol. Chem. 264*, 10023−10033.
6. Newman, L. M., and Wackett, L. P. (1995) *Biochemistry 34*, 14066−14076.
7. Pikus, J. D., Studts, J. M., Achim, C., Kauffmann, K. E., Münck, E., Steffan, R. J., McClay, K., and Fox, B. G. (1996) *Biochemistry 35*, 9106−9119.
8. Small, F. J., and Ensign, S. A. (1997) *J. Biol. Chem. 272*, 24913−24920.
9. Lipscomb, J. D. (1994) *Annu. Rev. Microbiol. 48*, 371−399.
10. Wallar, B. J., and Lipscomb, J. D. (1996) *Chem. Rev. 96*, 2625−2657.
11. Froland, W. A., Andersson, K. K., Lee, S.-K., Liu, Y., and Lipscomb, J. D. (1992) *J. Biol. Chem. 267*, 17588−17597.

12. Froland, W. A., Andersson, K. K., Lee, S.-K., Liu, Y., and Lipscomb, J. D. (1993) in *Microbial Growth on C_1 Compounds* (Murrell, J. C., and Kelly, D. P., Eds.) pp 81−92, Intercept, Ltd, Andover, U.K.
13. Fox, B. G., Liu, Y., Dege, J. E., and Lipscomb, J. D. (1991) *J. Biol. Chem. 266*, 540−550.
14. Liu, Y., Nesheim, J. C., Lee, S.-K., and Lipscomb, J. D. (1995) *J. Biol. Chem. 270*, 24662−24665.
15. Paulsen, K. E., Liu, Y., Fox, B. G., Lipscomb, J. D., Münck, E., and Stankovich, M. T. (1994) *Biochemistry 33*, 713−722.
16. Davydov, R., Valentine, A. M., Komar-Panicucci, S., Hoffman, B. M., and Lippard, S. J. (1999) *Biochemistry 38*, 4188−4197.
17. Matsuo, H., Walters, K. J., Teruya, K., Tanaka, T., Gassner, G. T., Lippard, S. J., Kyogoku, Y., and Wagner, G. (1999) *J. Am. Chem. Soc. 121*, 9903−9904.
18. Gassner, G. T., and Lippard, S. J. (1999) *Biochemistry 38*, 12768−12785.
19. Liu, K. E., Valentine, A. M., Wang, D., Huynh, B. H., Edmondson, D. E., Salifoglou, A., and Lippard, S. J. (1995) *J. Am. Chem. Soc. 117*, 10174−10185.
20. Valentine, A. M., Stahl, S. S., and Lippard, S. J. (1999) *J. Am. Chem. Soc. 121*, 3867−3887.
21. Qian, H., Edlund, U., Powlowski, J., Shingler, V., and Sethson, I. (1997) *Biochemistry 36*, 495−504.
22. Chang, S.-L., Wallar, B. J., Lipscomb, J. D., and Mayo, K. H. (1999) *Biochemistry 38*, 5799−5812.
23. Walters, K., Gassner, G. T., Lippard, S. J., and Wagner, G. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 7877−7882.
24. Cadieux, E., and Powlowski, J. (1999) *Biochemistry 38*, 10714−10722.
25. Studts, J. M., and Fox, B. G. (1999) *Protein Expression Purif. 16*, 109−119.
26. Bodenhausen, G., Kogler, H., and Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370−388.
27. Marion, D., and Wuthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.
28. Zijl, P. C. M. V., Johnson, M. O. N., Mori, S., and Hurd, R. E. (1995) *J. Magn. Reson., Ser. A 113*, 256−270.
29. Warren, W. S., Richter, W., Andreotti, A. H., and Farmer, B. T. (1993) *Science 262*, 2005−2009.
30. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) *J. Biomol. NMR 6*, 135−140.
31. Clore, G. M., Gronenborn, A. M., Nilges, M., and Ryan, C. A. (1987) *Biochemistry 26*, 8012−8023.
32. Wuthrich, K., Billetter, M., and Braun, W. (1983) *J. Mol. Biol. 169*, 949−961.
33. Clubb, R. T., Ferguson, S. B., Walsh, C. T., and Wagner, G. (1994) *Biochemistry 33*, 2761−2772.
34. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) *Biochemistry 31*, 1647−1651.
35. Wagner, G., Braun, W., Havel, T. F., Schaumman, T., Go, N., and Wuthrich, K. (1987) *J. Mol. Biol. 196*, 611−639.
36. Bruenger, A. T. (1993) *X-PLOR Version 3.1 Manual*, Yale University, New Haven, CT.
37. Gunter, P., Mumenthaler, C., and Wuthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
38. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graph. 14*, 51−55.
39. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
40. Sayle, R., and Milner-White, E. J. (1995) *Trends. Biochem. Sci. 20*, 374.
41. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
42. Pearson, W. R., and Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2444−2448.
43. Myers, E. W., and Miller, W. (1989) *CABIOS, Comput. Appl. Biosci. 4*, 11−17.
44. Pearson, W. R. (1990) *Methods Enzymol. 183*, 63−98.
45. Higgins, D. G., Bleasby, A. J., and Fuchs, R. (1992) *CABIOS, Comput. Appl. Biosci. 8*, 189−191.
46. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673−4680.
47. Felsenstein, J. (1989) *Cladistics 5*, 164−166.
48. Pikus, J. D., Studts, J. M., McClay, K., Steffan, R. J., and Fox, B. G. (1997) *Biochemistry 36*, 9283−9289.
49. Pikus, J. D., Mitchell, K. H., Studts, J. M., McClay, K., Steffan, R. J., and Fox, B. G. (2000) *Biochemistry 39*, 791−799.
50. Haas, J. A., and Fox, B. G. (1999) *Biochemistry 38*, 12833−12840.
51. Cleland, W. W. (1979) *Methods Enzymol. 63*, 103−139.
52. Hemmi, H., Studts, J. M., Chae, Y. K., Fox, B. G., and Markley, J. L. (2000) *J. Biomol. NMR 16*, 359−360.
53. Kim, Y., Ayoubi, P., and Harker, A. R. (1996) *Appl. Environ. Microbiol. 62*, 3227−3233.
54. Johnson, G. R., and Olsen, R. H. (1995) *Appl. Environ. Microbiol. 61*, 3336−3346.
55. Lloyd, J. S., Bhambra, A., Murrell, J. C., and Dalton, H. (1997) *Eur. J. Biochem. 248*, 72−79.
56. Branstetter, H., Whittington, D. A., Lippard, S. J., and Frederick, C. A. (1999) *Chem. Biol. 6*, 441−449.
57. Harker, A. R., and Kim, Y. (1990) *Appl. Environ. Microbiol. 56*, 1179−1181.
58. Matrubutham, U., and Harker, A. R. (1994) *J. Bacteriol. 176*, 2348−2353.
59. Leahy, J. G., Johnson, G. R., and Olsen, R. H. (1997) *Appl. Environ. Microbiol. 63*, 3736−3739.
60. Johnson, G. R., and Olsen, R. H. (1997) *Appl. Environ. Microbiol. 63*, 4047−4052.
61. Stirling, D. I., and Dalton, H. (1979) *Eur. J. Biochem. 96*, 205−212.

BI0013703