# Crystal Structures and Functional Studies of T4moD, the Toluene 4-Monooxygenase Catalytic Effector Protein[†,‡]

George T. Lountos,[§] Kevin H. Mitchell,[||] Joey M. Studts,[||] Brian G. Fox,*[,||] and Allen M. Orville*[,§]

*School of Chemistry and Biochemistry and the Parker H. Petit Institute of Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, and the Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin, Madison, Wisconsin 53706-1544*

*Received December 3, 2004; Revised Manuscript Received February 22, 2005*

ABSTRACT: Toluene 4-monooxygenase (T4MO) is a four-component complex that catalyzes the regiospecific, NADH-dependent hydroxylation of toluene to yield *p*-cresol. The catalytic effector (T4moD) of this complex is a 102-residue protein devoid of metals or organic cofactors. It forms a complex with the diiron hydroxylase component (T4moH) that influences both the kinetics and regiospecificity of catalysis. Here, we report crystal structures for native T4moD and two engineered variants with either four (ΔN4-) or 10 (ΔN10-) residues removed from the N-terminal at 2.1-, 1.7-, and 1.9-Å resolution, respectively. The crystal structures have C-alpha root-mean-squared differences of less than 0.8 Å for the central core consisting of residues 11−98, showing that alterations of the N-terminal have little influence on the folded core of the protein. The central core has the same fold topology as observed in the NMR structures of T4moD, the methane monooxygenase effector protein (MmoB) from two methanotrophs, and the phenol hydroxylase effector protein (DmpM). However, the root-mean-squared differences between comparable C-alpha positions in the X-ray structures and the NMR structures vary from ∼1.8 Å to greater than 6 Å. The X-ray structures exhibit an estimated overall coordinate error from 0.095 (0.094) Å based on the *R*-value (*R* free) for the highest resolution ΔN4-T4moD structure to 0.211 (0.196) Å for the native T4moD structure. Catalytic studies of the ΔN4-, ΔN7-, and ΔN10- variants of T4moD show statistically insignificant changes in $k_{cat}$, $K_M$, $k_{cat}/K_M$, and $K_I$ relative to the native protein. Moreover, there was no significant change in the regiospecificity of toluene oxidation with any of the T4moD variants. The relative insensitivity to changes in the N-terminal region distinguishes T4moD from the MmoB homologues, which each require the ∼33 residue N-terminal region for catalytic activity.

T4MO[1] is a member of the diiron monooxygenase family of enzymes (*1*). These evolutionarily related enzyme complexes are classified as either aliphatic or aromatic hydroxylases (*2*). MMO, a member of the aliphatic monooxygenase class, is the most thoroughly studied diiron monooxygenase and serves as the structural and functional paradigm for the family (*3*, *4*). The aromatic hydrocarbon hydroxylases can be grouped into two subclasses based on operon structure, sequence analyses, and the primary substrate that is oxidized (*2*). Subclass I contains the toluene 2-monooxygenase/phenol hydroxylase enzymes (T2MO/phenol), while subclass II contains the toluene 4-monooxygenase/benzene monooxygenase enzymes (T4MO/benzene).

All members of the diiron monooxygenase family are multicomponent enzyme complexes that consist of either a one- or two-protein electron-transfer chain, a catalytic effector protein that contains no metal ions or organic cofactors, and a terminal hydroxylase (*2*, *5*−*7*). Specific

[†] This work was supported by the National Science Foundation (MCB-9733734 and MCB-0316232) to B.G.F. and the Georgia Tech Research Corporation, the Office of the Vice Provost for Research, Georgia Institute of Technology, and an American Chemical Society Petroleum Research Fund type G grant (40310-G4) to A.M.O. G.T.L. was supported in part by a U.S. Department of Education GAANN fellowship. K.H.M. was a trainee of NIH Institutional Biotechnology Pre-Doctoral Training Grant T32 GM08349. J.M.S. was supported in part by a Peterson/Wharton Fellowship from the Department of Biochemistry, University of Wisconsin. Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory (SSRL), operated by the U.S. Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program and the National Institute of General Medical Sciences. Portions of this work were also carried out at beamline 22-ID in the SER-CAT facilities at the Advanced Photon Source (APS). The use of APS is supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science under Contract No. W-31-109-Eng-38.

[‡] The atomic coordinates and structure factors have been deposited to the Protein Data Bank with the corresponding file names: native T4moD, 2bf2 and r2bf2sf; ΔN4-T4moD, 2bf5 and r2bf5sf; ΔN10-T4moD, 2bf3 and r2bf3sf.

* To whom correspondence should be addressed. B.G.F.: e-mail, bgfox@biochem.wisc.edu; telephone, (608) 262-9708; fax, (608) 262-3453. A.M.O.: e-mail, allen.orville@chemistry.gatech.edu; telephone, (404) 385-1154; fax, (404) 894−2295.

[§] Georgia Institute of Technology.

[||] University of Wisconsin.

[1] Abbreviations: T4MO, the four-protein toluene 4-monooxygenase complex from *Pseudomonas mendocina* KR1; T4moH, hydroxylase component of T4MO; T4moC, Rieske (2Fe-2S) ferredoxin component of T4MO; T4moD, catalytic effector component of T4MO; ΔN4, deletion of four residues from the natural N-terminal; ΔN7, deletion of seven residues from the natural N-terminal; ΔN10, deletion of 10 residues from the natural N-terminal; DmpM, catalytic effector protein of phenol hydroxylase from *Pseudomonas* sp. CF600; MMO, the three-component methane monooxygenase; MmoB, methane monooxygenase effector protein; ΔN29, deletion of 29 residues from the natural N-terminal; MmoH, methane monooxygenase hydroxylase component; MAD, multiwavelength anomalous diffraction; rms, root-mean-squared; SeMet, selenomethionine.

10.1021/bi047459g CCC: $30.25 
Published on Web 04/20/2005

interactions between the hydroxylase and reductase/ferredoxin components are required for electron transfer during catalysis (*3*, *5*, *7*). X-ray crystal structures of terminal hydroxylases MmoH from *Methylococcus capsulatus* (Bath) and *Methylosinus trichosporium* OB3b and the toluene/*o*-xylene monooxygenase from *Pseudomonas stuzeri* OX1 have been solved (*8*−*10*). Furthermore, NMR structures of the Rieske ferredoxin T4moC,[2] the ferredoxin domain, and the flavin domain of *M. capsulatus* (Bath) oxidoreductase have also been solved (*11*−*13*).

The effector protein has also been shown to serve an essential role in catalysis through formation of a protein−protein complex with the hydroxylase (*14*−*20*). NMR structures of effector proteins from each catalytic subclass have been solved (*21*−*24*). Although these structures exhibit approximately the same topology of secondary structural elements, there are significant differences in the 3D structures, which may arise in part from the different numbers of NOE restraints used to calculate the solution structures (*21*−*24*). The present crystal structures, determined to 2.1 Å resolution and better, allow comparison with the NMR structure of T4moD and other effector proteins.

In each of the previously determined NMR structures, a portion of the N-terminal region was disordered. For the MMO subclass, the disordered N-terminal region is ∼33 residues long. Deletion of the N-terminal end results in the complete loss of catalytic effector activity (*25*, *26*). More specifically, His33, which is five residues before the start of beta-strand 1 in the folded core of the protein, has a controlling influence on the rate of formation and decay of transient catalytic intermediates (*27*). Intriguingly, mutagenesis of His33 gave no apparent effect on steady-state catalysis, which instead appears to be governed by rate-limiting product release (*27*). Mutagenesis also changed the product specificity (*25*, *26*). As His33 is conserved in all sequenced examples of MmoB, the effector proteins are thought to control internal steps of the multistep catalytic cycle through formation of specific protein−protein interactions (*16*, *18*, *19*, *25*−*31*).

Sequence alignments and structural analyses have shown that there is substantial variation in the length of the disordered N-terminal regions of the effector proteins from the various enzyme subclasses (*21*−*24*, *26*, *32*, *33*). For example, the N-terminal disordered region of the effector proteins from the aromatic ring monooxygenases are only ∼3 (subclass I, T2MO/phenol) or ∼13 (subclass II, T4MO/benzene) residues long, respectively. Furthermore, sequence alignments revealed that there was no sequence conservation among these N-terminal sequences and that only four of the effector proteins from subclass II contained a His residue anywhere in the N-terminal sequence. Consequently, the role or requirement of these shorter N-terminal sequences in effector protein catalysis has not yet been established. In the T4moD solution structure, the disordered region consists of the sequence from Ser1 to Asn12.[3]

Here, we report X-ray crystal structures and functional analyses of natural T4moD and truncated variants with 4, 7, and 10 residues deleted from the N-terminal of the native protein. These latter variants are designated ΔN4-, ΔN7-, and ΔN10-T4moD, respectively. The crystal structures of the natural enzyme and the ΔN4- and ΔN10- variants were refined to resolutions between 2.1 and 1.7 Å in two space groups. Collectively, these structures exhibit low root-mean-squared differences, indicating that the protein core is well-ordered, independent of the presence of the N-terminal region, and not influenced by the different crystal lattice packing environments present in the two space groups. Comparison of the T4moD X-ray crystal structures with the NMR structure reveals substantial agreement with the topology of the fold. However, the high-resolution X-ray structures allow a reassessment of certain structural features proposed from our previous NMR work on T4moD, including hydrogen-bonding interactions, the positions of several side chains, and the nature of an internal cavity. Progressive removal of the N-terminal residues from T4moD is also shown to have only a modest influence on the steady-state kinetic properties and no significant effect on the regiospecificity of product formation in the reconstituted T4MO complex. Thus, specific interactions of residues in the N-terminal end are apparently not required for aromatic hydroxylation in the T4moH, in contrast to the essential role for the N-terminal end in the MmoB homologues during methane hydroxylation.

## MATERIALS AND METHODS

*Construction of T4moD Isoforms*. The ΔN4-, ΔN7-, and ΔN10-T4moD isoforms were created by removal of the indicated number of amino acids from the N-terminal region. These isoforms were constructed by PCR using the expression vector pJDP01 (*1*, *34*, *35*) as the template, Vent DNA polymerase (New England Biolabs, Beverly, MA), and a Perkin-Elmer Model 9600 thermocycler (Perkin-Elmer, Foster City, CA). The following oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA) and used as forward primers to create the deletions: ΔN4- (5′-agcacaCATATGgatcaggctttacataac-3′); ΔN7- (5′-tggctgatCATATGttacataacaataacg-3′); ΔN10- (5′-gctttaCATATGaataacgttggaccgattatccg-3′). Capital letters indicate the position of an *Nde*I restriction site used to place the start codon at the correct location relative to the ribosome-binding site in pJDP01. The reverse primer was 5′-caaggggttatgctagttattgctcagcggt-3′. To generate PCR products, a 50 μL reaction was subjected to 30 cycles of denaturation at 94 °C for 45 s, anneal at 60 °C for 30 s, and extention at 72 °C for 45 s. The PCR products were gel-purified and digested with *Nde*I and *Bam*HI and cloned into similarly digested pET3a (Novagen, Madison WI). The final constructs were sequenced using the Big Dye sequencing kit (Perkin-Elmer) at the University of Wisconsin Biotechnology Center.

*Preparation and Characterization of T4moD.* The T4moD variants and all other T4MO proteins were overexpressed and purified as previously reported (*1*, *36*, *37*). Yields of purified T4moD were ∼10−15 mg/L of culture medium. The SeMet-labeled T4moD was produced by fed-batch fermentation in expression host *Escherichia coli* B834(DE3) as previously described (*36*). Electrospray ionization mass spectrometry was performed at the University of Wisconsin Biotechnology Center using a Perkin-Elmer Sciex API 365 triple quadrupole mass spectrometer. Steady-state kinetic

[2] The crystal structure of T4moC has also been determined to 1.48 Å resolution (PDB code 1VM9).

[3] The residue numbering used in this work is based on counting after removal of the N-terminal Met residue, giving Ser1 as the first residue in the mature protein obtained from bacterial expression.

Table 1: Data Collection and Model Refinement Statistics

| T4moD variant | native- | ΔN4- | ΔN10- |
|---|---|---|---|
| | Data Collection Statistics | | |
| resolution range (Å) | 37−2.05 (2.11−2.05) | 37−1.71 (1.80−1.71) | 16−1.96 (2.07−1.96) |
| unique reflections | 19009 | 23407 | 15889 |
| completeness (%)[a] | 96.5 (78.2) | 99.99 (99.9) | 99.8 (99.8) |
| $R_{sym}$ (%)[a,b] | 6.6 (29.4) | 12.5 (35.1) | 6.4 (31.7) |
| $I/\sigma(I)$[c] | 6 (2.5) | 3.4 (1.9) | 9.6 (2.3) |
| | Refinement Statistics | | |
| resolution range (Å) | 37−2.1 | 37−1.71 | 16−1.96 |
| no. of reflections | 17059 | 21007 | 14262 |
| *R*-factor | 0.243 | 0.154 | 0.192 |
| *R*-free[d] | 0.289 | 0.185 | 0.249 |
| no. of non-H protein atoms | 1579 | 1473 | 1404 |
| no. of water molecules | 58 | 238 | 167 |
| no. of 1,2,3-heptanetriol molecules | 0 | 0 | 1 |
| mean *B*, protein atoms (Å$^2$) | 50.7 | 14.5 | 25.0 |
| mean *B*, water molecules (Å$^2$) | 49.6 | 33.1 | 36.8 |
| mean *B*, 1,2,3-heptanetriol molecule (Å$^2$) | | | 31.8 |
| rms deviations from ideal | | | |
| bond lengths (Å) | 0.018 | 0.016 | 0.019 |
| bond angles (deg) | 1.6 | 1.6 | 1.7 |

[a] Numbers in parentheses are for the highest resolution shell. [b] The average agreement between the independently measured intensities. [c] The root-mean-squared value of the intensity measurements divided by their estimated standard deviation. [d] Calculated with 10% of the data.

assays and determinations of product distributions were as previously reported (*22*, *35*). Velocity data for *p*-cresol formation versus effector protein concentration were fit to $v = k_{cat}[S]/(K_M + [S] + ([S]^2/K_I))$ using the NonLinearRegress routine of Mathematica (v4.0.1.0, Wolfram Research, Inc. Champaign, IL), where $k_{cat}$ is the maximal reaction velocity, [S] is the concentration of effector protein, $K_M$ is the apparent Michaelis constant for formation of the activating complex of effector protein, and $K_I$ is the apparent equilibrium constant for formation of the inhibiting complex of effector protein. The $k_{cat}$-values are reported as turnover numbers relative to diiron center in the ($\alpha\beta\gamma$) protomer of T4moH.

*Crystallization and X-ray Diffraction Data Collection*. Single crystals were obtained for the natural, the SeMet-enriched and the ΔN10- variants as previously described (*38*). Crystals of the ΔN4- variant were obtained from the same conditions used for the ΔN10- variant (*38*). The X-ray diffraction data were collected from cryo-preserved crystals at approximately 100 K and processed as previously reported (*38*) with the diffraction data for the ΔN4- variant collected at the SER-CAT facility of the Advanced Photon Source (APS), Argonne National Laboratory.

*Crystal Structure Determination*. The structure of native T4moD was solved with MAD phasing using a four-wavelength data set collected from a SeMet−T4moD crystal (*38*). Two molecules of T4moD were present in the asymmetric unit. Three Se atoms were located with SOLVE (*39*) using data between 33 and 3.0 Å resolution and refined with SHARP (*40*) as previously reported (*38*). The experimental phases between 33 and 3.0 Å resolution were appended to the 2.09-Å resolution native T4moD data set. Phase extension and solvent-flattening to 2.09 Å was performed with the CCP4 suite of programs (*41*). Model building was performed with the program O (*42*, *43*) and refinements were with CNS (*44*) and/or REFMAC5 (*45*, *46*). Cross-validated $2mF_o - DF_c$ and $mF_o - DF_c$ maps (*47*) were used to evaluate the model and correct errors. Water molecules were located and refined in the final stages of refinement with ARP/Waters (*48*) and REFMAC5.

The ΔN4- and ΔN10-T4moD variants crystallized in space group $P2_13$ with unit cell dimensions of $a = b = c = 86.8$ Å. Two protein molecules were present in the asymmetric unit. The structures for the ΔN4- and ΔN10-T4moD isoforms were solved by molecular replacement in CNS with the native T4moD structure as the search model after appropriate truncation of the N-terminal residues, removal of solvent molecules, and assignment of the *B*-factors to 20 Å$^2$. The final structure for ΔN4-T4moD had an *R*-factor of 0.15 and an *R*-free of 0.19 with data between 37 and 1.71 Å resolution. The refined ΔN10-T4moD structure had an *R*-factor of 0.19 and an *R*-free of 0.25 with data between 16 and 1.96 Å resolution.

Validation of the refined models and Ramachandran analysis were done using SFCHECK (*49*) and PROCHECK (*49*−*51*). Solvent exposed surface areas were calculated with a 1.4 Å probe radius with DMS implemented in the MidasPlus package from the Computer Graphics Laboratory, University of California, San Francisco, CA (supported by NIH P41 RR-01081) (*52*), CNS, CCP4, or Swiss-PdbViewer (v3.7b2) (*53*). Secondary structure assignments were made using KSDSSP (*54*). The rms difference between models was calculated with CCP4 or Swiss-PdbViewer. Structure figures were prepared using Swiss-PdbViewer and PovRay (v3.5) or PyMOL (DeLano Scientific LLC, Castro City, CA).

## RESULTS

*Crystal Structure Determinations.* The X-ray data collection and refinement statistics for the native, ΔN4-, and ΔN10-T4moD are shown in Table 1. Although an NMR structure for T4moD was previously determined (*22*), various search models derived from the solution structure ensemble did not yield a molecular replacement solution for the native crystal structure. Consequently, the SeMet-enriched T4moD was crystallized and used to determine experimental phases via MAD data analysis techniques. Phasing statistics for SeMet−T4moD are reported in Orville et al. (2003) (*38*). The 3.0 Å resolution electron density map resulting from refining three Se atoms[4] in space group $P6_122$ was interpretable and used to build the majority of the initial model. The phase-extended, solvent-flattened map at 2.1-Å resolution for native T4moD was then used to complete the model building. The model

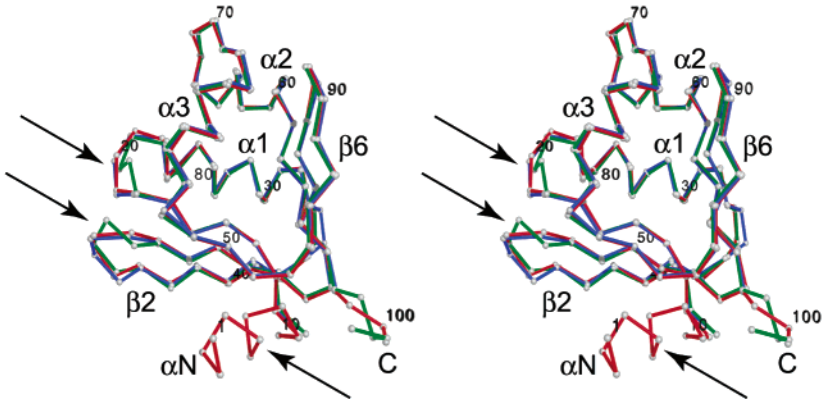


FIGURE 1: Comparison of the X-ray crystal structures of the T4moD variants. A divergent stereo-overlay of the $C_\alpha$ backbone trace for natural- (red), ΔN4- (green), and ΔN10-T4moD (blue), with arrows highlighting the regions that differ slightly in the three structures.

for native T4moD was then refined against the 2.1 Å resolution data obtained from the native protein. The final refined model for native T4moD had an *R*-factor of 0.24 and *R*-free of 0.28 with data between 37 and 2.1-Å resolution.

Each of the truncated T4moD variants crystallized in space group $P2_13$, and molecular replacement was used to solve these structures using a search molecule derived from the native crystal structure. The crystals of ΔN4-T4moD diffracted to the highest resolution (1.7 Å), and the resulting atomic model was refined to a final *R*-factor of 0.15 and *R*-free of 0.19. Ramachandran analysis of all the crystal structures showed that over 90% of the residues were located in the most favored region. In the natural T4moD, Arg45 is in the disallowed region in both chains A and B. The residue is also in the disallowed region for chain A or B of the ΔN4-T4moD and ΔN10-T4moD structures, respectively. The high-resolution structure of ΔN4-T4moD suggests an alternate conformation for Tyr47 in the B chain. At this position, the $2mF_o - DF_c$ electron density maps are consistent with the two most common conformers found in the O library. The X-ray structure of the ΔN10-T4moD variant contains a single 1,2,3-heptanetriol molecule, an additive in the crystallization solution, located at the interface of the two molecules in the asymmetric unit. Each of the heptanetriol hydroxyl groups hydrogen-bonds with either a solvent molecule or the carbonyl oxygen atoms of residues Leu80-A, Leu80-B, or Glu77-B.

*Overall Structure Description.* The three crystal structures of T4moD are similar, especially for the core region (residues 11−98) as illustrated in Figure 1. Indeed, the rms difference between $C_\alpha$ atoms in the core is less than 0.8 Å (Table 2) despite several differences in secondary structure assignments (discussed below). Figure 2B−F shows the rms differences and *B*-values for individual residue positions. The secondary structure assignments for each structure are illustrated in Figure 3. The structures differ most significantly at the N-terminal region for natural T4moD. In chain A, this region adopts an α-helical structure, designated αN, whereas in the B-chain and for both chains in the ΔN4-T4moD, this region is largely disordered. The other regions exhibiting slight differences are located between the $\beta$1 strand and the α1 helix and the hairpin turn between strands $\beta$2 and $\beta$3. Each of these regions also corresponds to larger thermal factors, suggesting some inherent structural variability to the fold.

Table 2: Root Mean Squared Differences between the T4MoD X-ray Structures and Related NMR Structures[a]

| | ΔN4-A[c] | ΔN10-A[c] | T4moD NMR[d] | MMO (OB3b)[e] | MMO (Bath)[f] | DmpM[g] |
|---|---|---|---|---|---|---|
| native-A[b] | 0.59 | 0.45 | 1.88[h] | 1.74 | 1.88 | 6.96 |
| ΔN4-A[c] | | 0.66 | 1.95 | 1.77 | 1.91 | 6.90 |
| ΔN10-A[c] | | | 1.91 | 1.68 | 2.10 | 6.86 |

[a] Superposition of $C_\alpha$ atoms using SwissPdb-Viewer and the iterative fit function. [b] Residues 5−100 in either chain A or chain B. [c] Residues 11−98. [d] Alignment with the best representative NMR structure [PDB 1G10, (*22*)]. [e] Thirty-four $C_\alpha$ atoms between residues 36−126 in the best representative structure from PDB code 2MOB_10 (*21*). [f] Fifty-nine $C_\alpha$ atoms between residues 36−128 in the best representative structure from PDB code 1CKV_11 (*24*). [g] Fifty-four $C_\alpha$ atoms between residues 31−84 in the best representative structure from PDB code 1HQI_3 (*23*). The best representative structure was selected from analysis of the ensemble NMR structures using OLDERADO (*69*).

From one perspective, the shape of T4moD is roughly triangular with approximately 24 Å × 20 Å × 23 Å edges when measured at $C_\alpha$ atoms of residues 46, 54, and 60. From the orthogonal perspective, the molecule is roughly rectangular with a width and length of approximately 20 Å by 30 Å when measured at $C_\alpha$ atoms of residues 31−89 and 54−69, respectively. Two sides of the triangular shape are comprised of antiparallel $\beta$-sheets, and the backbone trace crosses the corner connecting them three times. A Gly residue is present at each place the backbone chain crosses the vertex of the triangle (Gly36, Gly54, and Gly85 in T4moD). Figure 3 also shows a structure-edited sequence alignment of the effector proteins, and the consensus alignment shows that each of these positions is conserved to a differing extent across the entire family. For example, according to the more extensive alignment shown in Hemmi et al. (*22*), Gly36 is present in all eight of the sequences most closely related to T4moD but not in MMOB and DmpM. In addition, Gly54 is present in 17 of the 24 sequences aligned for the superfamily, while Gly85 is present in all sequences of the superfamily (*22*). Thus, these Gly residues appear to provide the conformational flexibility required for the backbone to radically alter direction. The angles between the $\beta$-sheets range from 74° to 85° in the X-ray structures and from 86° to 113° in the NMR structures (Table 3).

[4] The C-terminal Met residue was not visible in the electron density maps for each structure with the exception of the ΔN4-T4moD chain A.

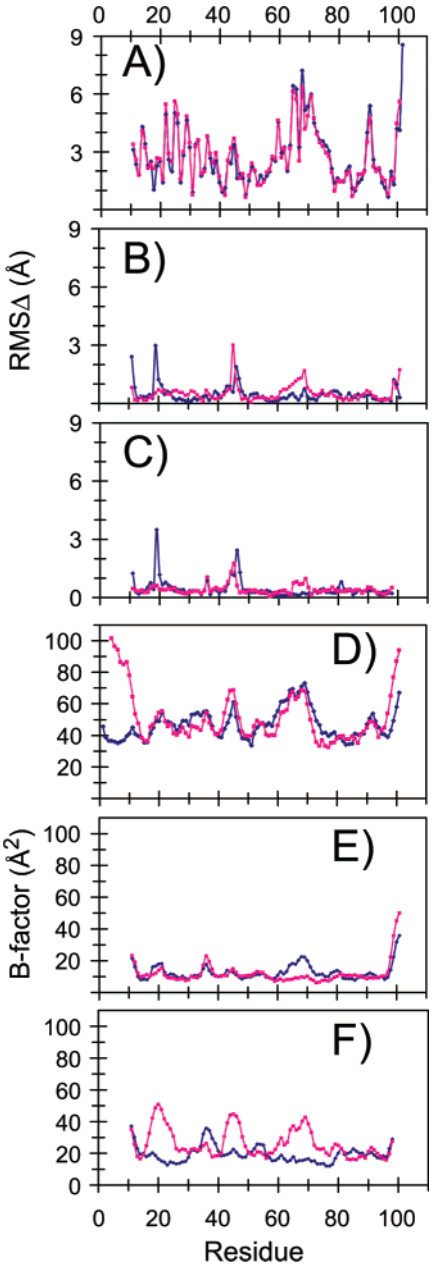


FIGURE 2: The rms differences and *B*-factors for the main-chain atoms in each residue in chain A (blue) and chain B (magenta). The rms differences between: (A) ΔN4-T4moD and the average NMR structure of native T4moD from 1G10 (*22*), (B) ΔN4-T4moD and the X-ray structure of native T4moD, and (C) ΔN4-T4moD and ΔN10-T4moD. The average *B*-factor values for all backbone atoms in each residue for each chain in (D) native T4moD, (E) ΔN4-T4moD, and (F) ΔN10-T4moD.

The interior of T4moD is packed exclusively with hydrophobic residues. Although one small cavity is detected in each chain (Figure 4), the electron density maps do not reveal any ordered solvent molecules within the cavity even at high resolution. The volume of this cavity is somewhat variable depending on the chain and structure, but is typically ~50 Å$^3$, which is approximately comparable to the volume required to accommodate a toluene molecule (*55*−*59*). The borders of the cavity are comprised of Leu22, Pro25, Val26, Leu63, Leu67, Phe71, and Leu76. A few hydrophobic residues are also on the exterior surface of the protein. However, either charged or polar residues typically flank the exterior hydrophobic residues. For example, Glu55, Thr59, Gln86, and Arg95 surround the external residue Ile57. Thus, they provide hydrophobic interactions with Ile57 and polar interactions with solvent molecules. In addition, Ile78, Phe83, and Ile87 are on the exterior surface. These residues are located near the two-fold interface between the two chains in the asymmetric unit. Consequently, their exposure to solvent is limited in the crystal lattice.

All of the charged residues in T4moD are located on the exterior of the protein, and their distribution is not uniform. Consequently, there are regions with distinctive positive or negative electrostatic surface potential (Figure 5). For example, α1 is an amphipathic helix with Asp21, Glu24, Glu28, Glu31, and Asp33 all on the exterior surface, which yields a strong negative electrostatic surface patch (red in Figure 5). The other side of the helix is comprised of Val23, Val26, Ile27, and Ala30, and these residues all project into the hydrophobic core of the protein. In contrast to the negatively charged patch, the region around the hairpin turn between the β2 and β3 strands contains four arginine residues (Arg18, Arg44, Arg45, Arg49), which yields a distinct positive electrostatic patch (blue in Figure 5).

*N-Terminal Region.* The N-terminal regions in the two chains of the native T4moD structure differ. In addition, there is no electron density visible for ΔN4-T4moD prior to Asn11 in both chains. The electron density for the native T4moD suggests that residues Ser1−Ala7 in chain A adopt an α-helical configuration. The electron density is weaker for the analogous residues in chain B, which suggests the presence of a majority-disordered fraction and a minority-ordered fraction for these residues. The *B*-factors for these regions also differ and are consistent with decreased order for the N-terminal region of the B-chain (Figure 2D). Analysis of the crystal lattice packing reveals that these regions are located in different environments. In chain A, the N-terminal residues are located where four chains pack together. Consequently, only approximately 20% of the surface area for the N-terminal α-helix is exposed to bulk solvent. In contrast, nearly the entire N-terminal region in chain B is exposed to solvent because it projects into a large, solvent-filled space in the crystal lattice. The electron density maps for the two chains in ΔN4-T4moD also suggest that the N-terminal regions are disordered, and analysis of lattice packing revealed that these residues are exposed to the bulk solvent between molecules in the crystal lattice.

*Catalytic Activity of N-Terminal Deletions of T4moD.* Table 4 shows the steady-state catalytic parameters and regiospecificity for toluene oxidation measured from the various N-terminal truncations of T4moD. These characterizations were made according to previously described experimental procedures (*34*, *35*) and included control reactions with the natural isoforms (*1*). The results show that the disordered N-terminal region of T4moD is not essential for catalysis by the T4MO enzyme complex. Thus a 10-residue deletion of the N-terminal gave statistically insignificant changes in $k_{cat}$, apparent $K_M$, and $k_{cat}/K_M$. Likewise, the apparent $K_I$ for inhibition of the toluene hydroxylation reaction was also unaffected by removal of the N-terminal from T4moD. Furthermore, the regioselectivity was nearly unchanged, with the percentage of *p*-cresol slightly decreased from 96.2% in native T4moD to 94.5% in ΔN10-T4moD. These observations also suggest that no single residue in the

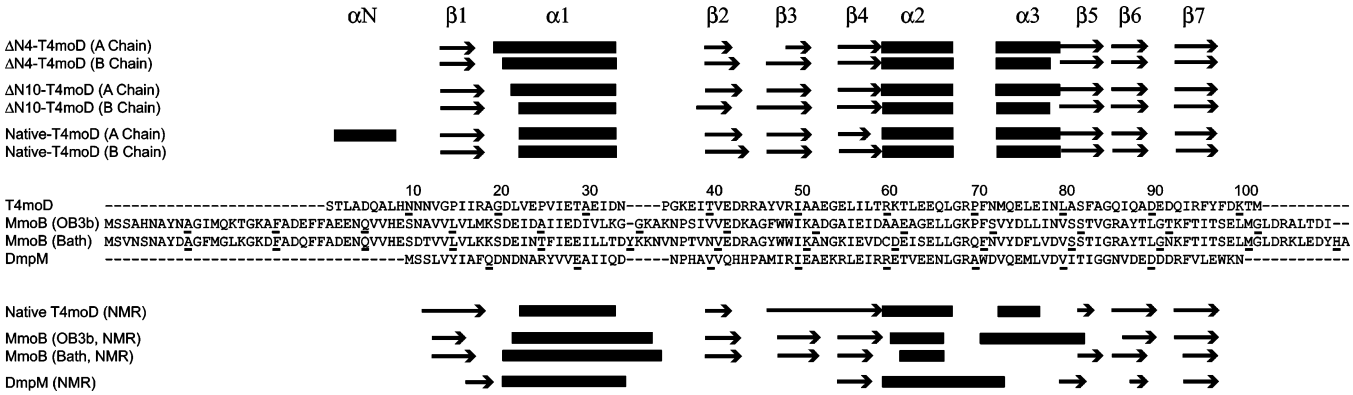


FIGURE 3: Primary sequence and secondary structure alignment for T4moD and other effector proteins. The amino acid sequence alignment was from CLUSTALW using default weights (T4moD gi:45479224; MmoB OB3b, gi:44616 (*21*); MmoB (Bath), gi:127207 (*24*); DmpM, gi:118693 (*23*)). Every 10th residue is underlined with the numbering fixed to that of the native T4moD after post-translational removal of the N-terminal Met. The arrows and rectangles indicate the $\beta$-strands and α-helices as assigned from KSDSSP (*54*). The secondary structure elements from the T4moD X-ray structures are shown on the top of the figure for both A- and B-chains in the asymmetric units. The secondary structure elements assigned from the NMR structures of T4moD and the other effector proteins are shown on the bottom of the figure (*21*−*24*).

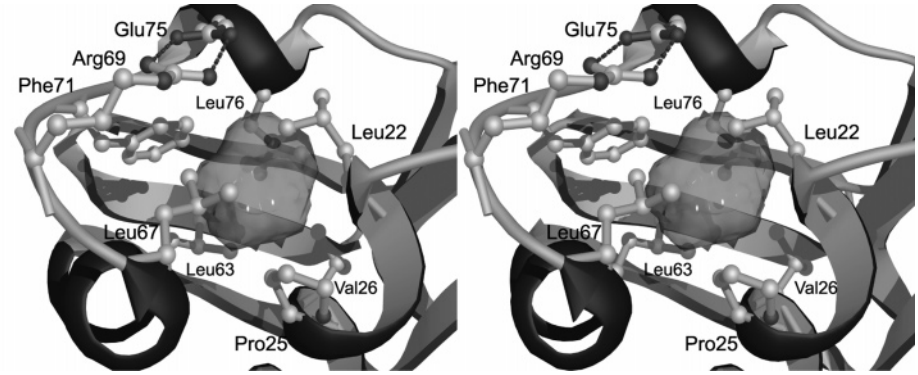


FIGURE 4: A divergent stereoimage of the environment around the hydrophobic cavity illustrated with chain A of the ΔN4-T4moD crystal structure.

Table 3: Hinge Angle Calculated between $\beta$-Sheets of Effector Proteins[a]

| | X-ray[b] | | NMR | | | |
|---|---|---|---|---|---|---|
| T4moD variant | chain A | chain B | T4moD[c] | MMOB (OB3b)[d] | MMOB (Bath)[e] | DmpM[f] |
| Native- | 78° | 82° | 86° | 96° | 88° | 113° |
| ΔN4- | 74° | 85° | | | | |
| ΔN10- | 81° | 83° | | | | |

[a] The angle calculated was between the $\beta$-sheets within the indicated structure. [b] This work, using the $C_\alpha$ of residues 46−52−91 of the given chain. [c] Calculated from 1G10 (*22*) using residues 46−52−91. [d] Calculated from 2MOB_10 (*21*) using residues 76−80−119. [e] Calculated from 1CKV_11 (*24*) using residues 77−81−120. [f] Calculated from 1HQI_3 (*23*) using residues 36−42−81. The best representative structure was selected from the analysis of the ensemble NMR structures using OLDERADO (*69*).

disordered region of T4moD is likely to be responsible for the observed modest changes in catalytic parameters.

Our previous studies on the reconstitution of T4MO with effector proteins from the subclass I (T2MO/phenol) and subclass II (T4MO/benzene) monooxygenases (*22*) showed that TbuV from subclass II was able to effectively complement catalytic activity, while S1 from subclass I was not. To further define the reactions of heterologous effector proteins with T4MO, the natural MmoB effector protein and the catalytically inactive ΔN29-MmoB were also tested as part of this work. As with the S1 effector, neither of the MmoB variants was able to complement the T4MO complex in the standard T4MO in vitro assay (*1*).

## DISCUSSION

*Comparison of the T4moD Crystal Structures and NMR Structures*. Figure 6 shows a ribbon trace of the average NMR structure of native T4moD superimposed on the ΔN4-TmoD X-ray structure using $C_\alpha$ atoms from residues 11−98. The average rms difference from this overlay is 1.95 Å, and the individual rms differences are plotted with respect to residue position in Figure 2A. For ΔN4-T4moD, the structure refined from the higher resolution data gave 93.7% of the residues in the most favored region of the Ramachandran plot. In comparison, only 65.6% of the residues in the NMR structure were in the most favored region. The refined X-ray structures yield an estimated overall coordinate error of 0.095 Å for ΔN4-T4moD, 0.180 Å for ΔN10-T4moD, and 0.211 Å for native T4moD. The correlation coefficients were 0.961 (ΔN14-T4moD), 0.943 (ΔN10-T4moD), and 0.945 (native T4moD). For comparison, the core region of the NMR structure (Asn12−Phe98) exhibited rms differences to the averaged structure of 0.71 Å for the backbone atoms and 1.24 Å for all non-hydrogen atoms (*22*). The crystal structures of the natural and variants of T4moD are more similar to each other than they are to the average NMR structure (Figures 1 and 2 and Table 2). Indeed, the rms differences between $C_\alpha$ atoms in the T4moD crystal struc-

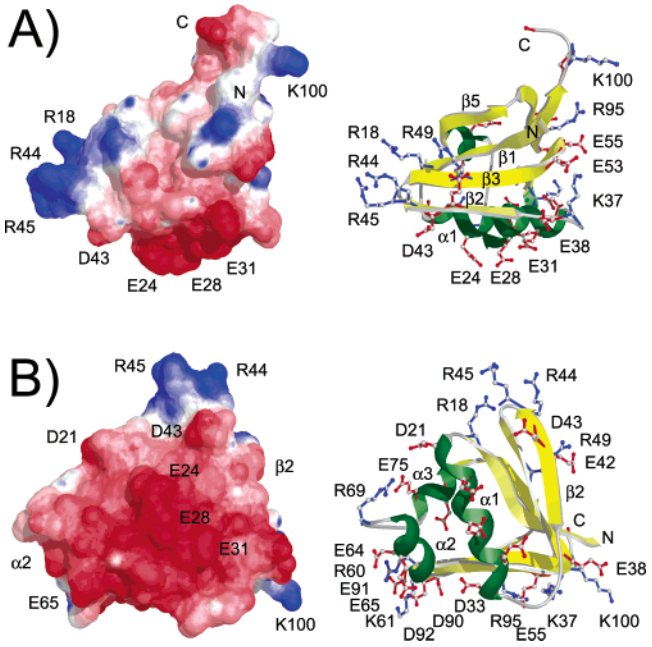


FIGURE 5: The electrostatic potential mapped onto the solvent exposed surface of ΔN4-T4moD. (A) The solvent exposed surface determined with a 1.4 Å probe radius and colored according to the calculated Coulombic electrostatic potentials (−6, −1, and 1.75 are red, white, and blue, respectively). The ribbon drawing of the backbone trace shown on the right is in the identical orientation and scale. The α1 helix is on the bottom, approximately in the plane of the page with its N-terminal end on the left. (B) Another orientation of the molecule in which the α1 helix is in the front.

tures are all less than 0.8 Å. The rms differences between the $C_\alpha$ atoms in the crystal structures and the average T4moD NMR structure are approximately 2.0 Å (Table 2), and increase to ~6 Å when specific secondary structure elements are considered (e.g., the α2 helix in T4moD) or increase beyond 6 Å when the T4moD X-ray structure is compared with the DmpM NMR structure. The high-resolution X-ray data also more accurately describes the positions of Arg69−Leu80, since these residues were not well-defined in the NMR structure (*22*). As illustrated in Figure 2A, this region exhibits the highest rms differences between the X-ray and NMR structure.

Several notable differences between the T4moD NMR and the X-ray structures are readily apparent (Figures 6 and 2A) and will be discussed with respect to the ΔN4-T4moD variant. The three α-helices in the X-ray and NMR structures exhibit significant differences. For example, the α1 helix consists of three turns in the NMR structure, but of four turns in the X-ray structures. Although the α1 helix terminates at analogous locations (Asp33), the helix starts at Gly19 in chain A of the ΔN4-T4moD and at Val23 in the NMR structure (Figure 6a). Consequently, this region yields rms differences of ~5 Å. The axes of the α2 helices intersect with an angle of approximately 45° in the two structures (Figure 6b). Although the α1 and α2 helical axes are nearly parallel in the NMR structure, they are related by ~45° in the X-ray structures. The connecting loop between the α2 and α3 helices and the orientation of the α3 helix are also different in the structures (Figure 6c). Consequently, the region between residues 60 and 80 yields the highest rms differences between the NMR and X-ray structures (see also below). Finally, the hairpin turns connecting the $\beta2-\beta3$ and $\beta6-\beta7$ strands also diverge significantly (Figure 6d and 6e) and yield peaks in the rms difference plot (Figure 2A).

Each T4moD crystal structure exhibits a well-formed hydrophobic core with all charged residues on the exterior surface of the protein. Thus, the crystal structures appear more compact than the NMR structures (Figure 6). Consistent with these differences, the molecular surface area for the core residues (11−98) of the crystal structure (ΔN4-T4moD, 4170 Å$^2$) is smaller than the average NMR structure (5753 Å$^2$). Moreover, the crystal structures have a continuous solvent accessible surface area with only small depressions on the surface, while analysis of the surface area of the NMR structure reveals a large, tubular invagination that extends through the entire protein. Because of this feature, the enclosed solvent excluded volume for the NMR structure (9580 Å$^3$) is smaller than that of the crystal structures (ΔN4-T4moD, 10450 Å$^3$).

The hinge angle between the two β-sheets is listed in Table 3 and has been suggested to be a structural feature that differentiates the effector proteins (*22*). Three conserved glycine residues are located at the vertex of the hinge in T4moD (Gly36, Gly54, and Gly85), and extensive hydrogen bonding across the two β-sheets apparently stabilizes each particular angular relationship between the β-sheets. Thus, differences in hydrogen bonding between the α1, α2, and α3 helices and the two β-sheets apparently stabilize the different hinge angles among the family members. It is also

Table 4: Comparison of the Catalytic Properties and Regiospecificity of the Natural and N-Terminal Deleted Isoforms of Toluene 4-Monooxygenase Effector Protein

| | kinetic parameters[a] | | | | percent of products[b] | | | |
|---|---|---|---|---|---|---|---|---|
| T4moD variant[c] | $k_{cat}$ ($s^{-1}$) | $K_M$ (μM) | $k_{cat}/K_M$ (μM $s^{-1}$) | $K_I$ (μM) | *p*-cresol | *m*-cresol | *o*-cresol | benzyl alcohol |
| native- | 3.3(0.1) | 7.0(0.7) | 0.47(0.1) | 98(10) | 96.2(0.2) | 1.5(0.1) | 0.9(0.0) | 1.4(0.1) |
| ΔN4- | 3.4(0.2) | 2.3(0.3) | 1.47(0.6) | 106(17) | 95.9(0.2) | 1.6(0.1) | 1.0(0.1) | 1.5(0.2) |
| ΔN7- | 2.8(0.1) | 5.3(0.5) | 0.53(0.2) | 110(11) | 95.2(0.2) | 1.8(0.1) | 1.2(0.1) | 1.8(0.2) |
| ΔN10- | 3.1(0.2) | 3.2(0.5) | 0.97(0.4) | 116(19) | 94.5(0.2) | 1.8(0.1) | 1.4(0.1) | 2.2(0.2) |

[a] Apparent kinetic parameters determined with each T4moD variant treated as the variable substrate and all other components of the reconstituted enzyme complex present in optimal amounts and NADH, toluene, and $O_2$ present in saturating concentrations. All kinetic experiments were performed in triplicate with at least seven different concentrations of T4moD. The standard error is shown in parentheses. The $k_{cat}$ values are reported relative to the $\alpha\beta\gamma$ protomer concentration of the T4moH component. [b] Percentage of product distribution observed from toluene oxidation. The standard error is shown in parentheses. [c] N-terminal sequence of the mature native protein obtained after in vivo post-translational removal of the N-terminal Met residue encoded by the expression plasmid, consisting of STLADQALHNNN-β1. β1 refers to the start of β-strand 1, the first identifiable element of secondary structure in T4moD. ΔN4-, deletion of 4 residues from the mature N-terminal to give DQALHNNN-β1; ΔN7-, deletion of 7 residues from the mature N-terminal to give LHNNN-β1; ΔN10-, deletion of 10 residues from the mature N-terminal to give NN-β1.

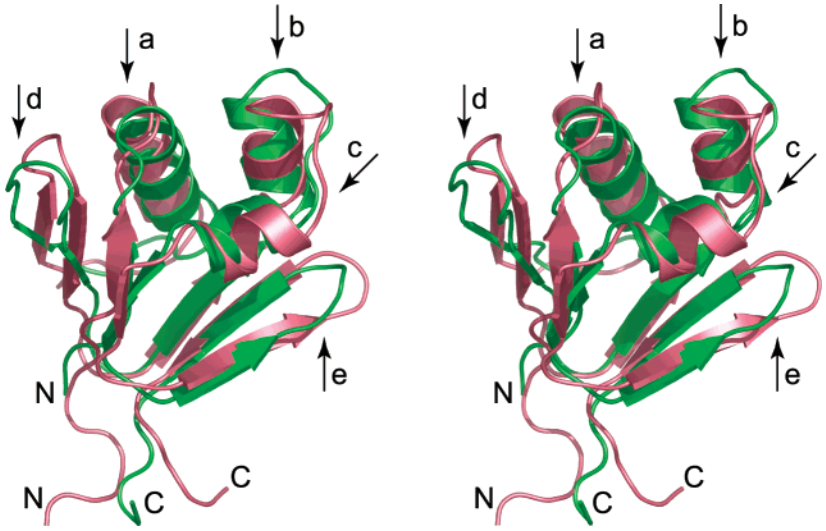


FIGURE 6: Divergent stereoview of the superimposed X-ray and NMR [1G10 (*22*)] structures of T4moD using residues 11−98. The chain A from the crystal structure in ΔN4-T4moD is shown in green, and the average NMR structure is red. The positions where the largest differences in the $C_\alpha$ traces occur are indicated with arrows and described in the text.

worth considering that the hydrophobic cavity is located between these three α-helices and toluene binding (if it does) would potentially influence this hinge angle.

*Environment of Asn34.* The sequence Asn34-Pro35-Gly36 is highly conserved across the catalytic effector protein superfamily (for example, see Figure 7 in Hemmi et al. (*22*)). Despite this sequence conservation, the X-ray and NMR structures of T4moD do not produce reasonable overlays when only the $C_\alpha$, backbone atoms, or all atoms of these three residues are superimposed. Although the overlay between the selected residues is reasonable, the rest of the structures are completely out of register. Moreover, when the α1 helix and these three conserved residues are used, this does not yield satisfactory superposition. Thus, significant structural differences are present in this region of the structures. During the analysis of the T4moD NMR structure, the Asn34 side-chain amide protons were assigned to resonances that were 3$\sigma$ outliers to the average chemical shifts assigned to amide groups in other proteins (http://www.bmrb.wisc.edu/data_access/outlier_selection_grid.html). These protons were also unusually resistant to exchange with solvent, despite being located near the surface of the protein (*22*), implying a unique chemical environment for this side chain. Figure 7A shows the $2mF_o - DF_c$ electron density for this region of the protein and the quality of the corresponding atomic model. The X-ray structures show that Asn34 and Pro35 cap the C-terminal end of the α1 helix and that no solvent molecules are hydrogen-bonded to $Asn34^N$. In the ΔN4-T4moD crystal structure, nearly every atom of Asn34 that can participate in hydrogen bonding does so. For example, the hydrogen atoms associated with $Asn34^{N\delta2}$ are positioned to donate hydrogen bonds to the carbonyl oxygen atoms of Ala30 and Ile57. $Asn34^{O\delta1}$ is poised to accept hydrogen bonds from $Ile57^N$ and a solvent atom. In addition, $Asn34^N$ donates a hydrogen bond to $Ala30^O$, whereas $Asn34^O$ accepts a hydrogen bond from $Lys37^N$.

The hydrogen-bonding pattern assigned from the NMR structure of T4moD is considerably different, beginning where $Asn34^N$ provides a hydrogen bond to $Glu31^O$ and $Asn34^O$ accepts a hydrogen bond from $Gly36^N$ (see Figure 7C). The differences between the X-ray and NMR structures increase as the hydrogen-bonding patterns propagate further from Asn34. They maximize at Arg60, where the NMR structures indicates that Arg60 forms an ionic interaction with Asp33 (see Figure 7C). In contrast, the X-ray structure shows an ionic interaction between Arg60 and Glu64 (see below and Figure 8). Thus, some of the largest differences between the structures, also indicated in the rms difference plots (Figure 2A), are in the regions near Asp33 and Arg60. Two consequences of the differences in hydrogen bonding detected by the X-ray structures are that the α1 helix is longer and truly amphipathic. Indeed, all five negatively charged residues are on the exterior surface, and all of the hydrophobic residues are buried inside the hydrophobic core of the structure.

*Environment of Glu64.* The crystal structures reveal ionic interactions between Glu64 and Arg60 and between Glu75 and Arg69 (see Figure 8A,B). An interaction between Glu75 and Arg69 partially sequesters the hydrophobic cavity discussed above from bulk solvent and thus resembles an ionic gate for the cavity. The NMR structure of T4moD refined to a consensus family with the side chain of Glu64 projecting into the hydrophobic interior of the protein (Figure 8C), and the lack of a basic residue capable of forming a neutralizing charge pair or other potential hydrogen-bonding interactions was noted (*22*). Moreover, Arg60, Arg69, and Glu75 were not assigned to charge-pairing interactions. The refined NMR structures accommodated these differences through a slightly larger hinge angle between the two $\beta$-sheets (Table 3) as well as allowing the large solvent accessible channel described above. In the alignment of the NMR and X-ray structures, the $C_\alpha$ of Glu64 in the NMR structure nearly superimposes with the $C_\alpha$ of Leu63 in the crystal structures. Consequently, the Glu64 side chain of the NMR structure nearly superimposes with the small hydrophobic cavity observed in the crystal structures (compare parts B and C of Figure 8).

*N-Terminal Region.* The T4moD NMR structure includes a disordered region comprised of the first 10 residues, followed by a well-determined structure for residues 11−98 (*22*). Similarly, disordered and ordered regions are also

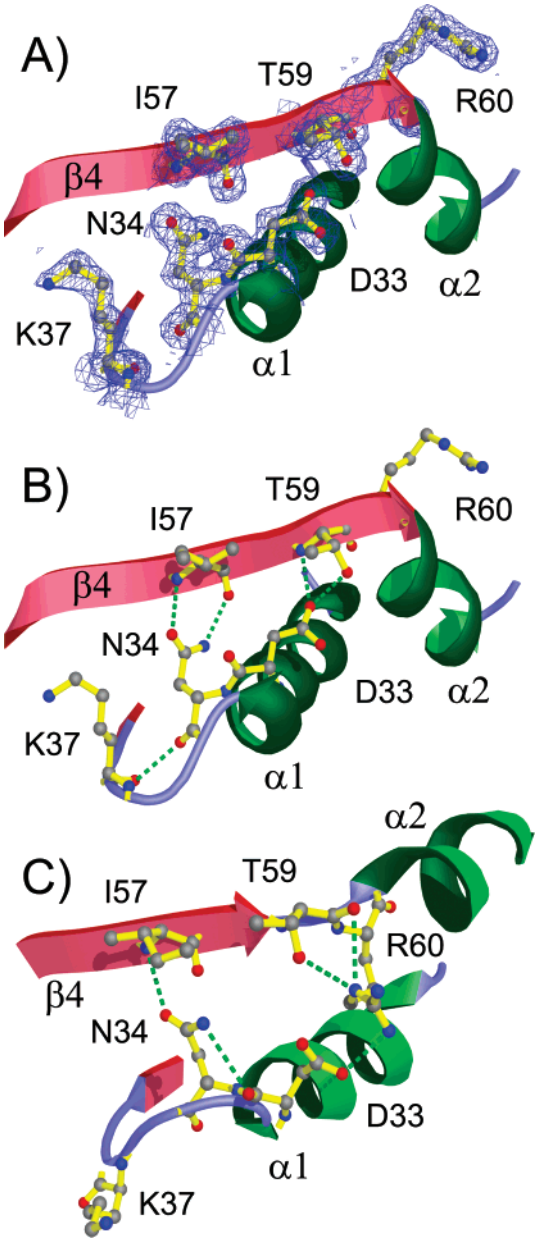


FIGURE 7: Differences in hydrogen-bonding patterns observed in the crystal and NMR [1G10 (*22*)] structures of T4moD surrounding residue Asn34. (A) The $2mF_o - DF_c$ electron density (1$\sigma$, 1.7 Å resolution) superimposed on the refined ΔN4-T4moD model. (B) X-ray structure of ΔN4-T4moD with hydrogen bonding indicated. Note that the position of Arg60 and α2 is different in the X-ray structure relative to the NMR structure. (C) NMR structure (*22*). Note that the hydrogen bonding and electrostatic interactions between Arg60 and Asp33 and Thr59 are not observed in the crystal structures.

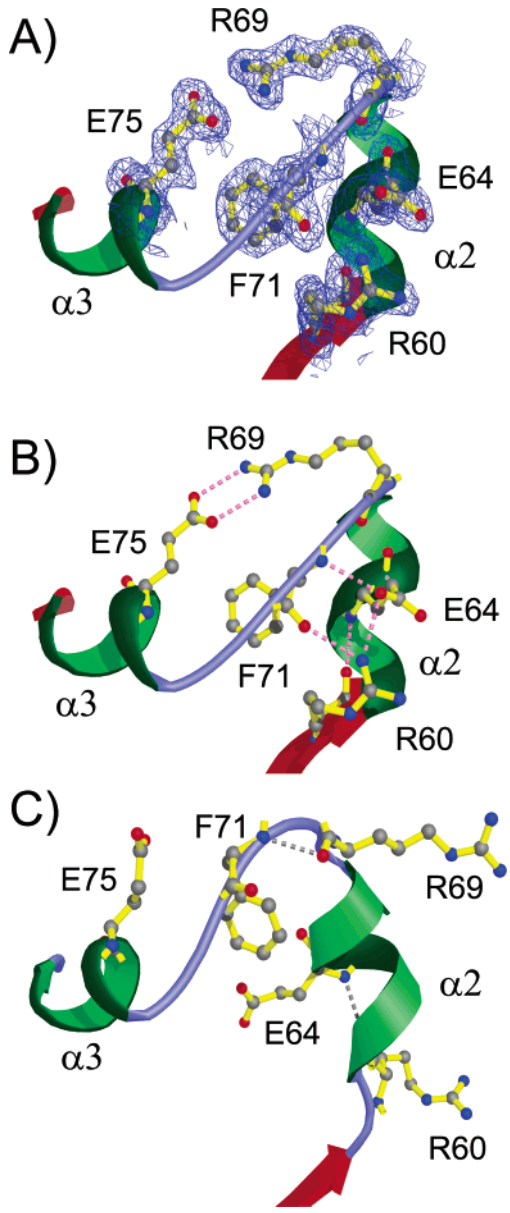


FIGURE 8: Differences in hydrogen-bonding patterns observed in the crystal and NMR (1G10 (*22*)) structures of T4moD surrounding several charged residues. (A) The $2mF_o - DF_c$ electron density (1$\sigma$, 1.7 Å resolution) superimposed on the refined ΔN4-T4moD model. (B) X-ray structure of ΔN4-T4moD showing that Arg60 and Glu64 form a charge pair, Arg69 and Glu75 form a second charge pair, and Phe71 occupies an interior position. (C) NMR structure (*22*) that placed Glu64 in an interior cavity. Interactions between Arg60 and Glu64 and Arg69 and Glu75 observed in the crystal structures were not predicted by the NMR structures.

observed in the crystal structures and appear to correlate with exposure to bulk solvent. Indeed, of the six chains in the crystal structures reported here, only the N-terminal end of chain A in the native T4moD projects into a crowded, lattice-packed environment, and only it exhibits a well-defined secondary structure. Thus, the α-helical structure in the N-terminal region of native T4moD appears to be dependent upon the environment or macromolecular crowding. Although the propensity for secondary structure to be sequence- and context-dependent has been documented, especially within T4 lysozyme (*60*−*62*), lattice interactions or macromolecular crowding that induce secondary structure is more difficult to assess (*63*, *64*).

*Role of N-Terminal Residues in Effector Proteins.* The N-terminal disordered region of the MmoB effector protein is required for catalysis. This portion of the protein plays a role in the formation of the catalytic intermediates used by MMO for the oxidation of methane (*19*, *26*, *28*), the most difficult hydrocarbon to oxidize. The methane oxidation requires a highly reactive intermediate, possibly a diferryl species (*3*, *5*, *20*, *65*−*67*). In contrast, other diiron enzymes have evolved to catalyze aromatic ring hydroxylations, an energetically less demanding reaction.

Deletion of the N-terminal region from T4moD had only modest influence on the steady-state kinetic properties and the regiospecificity of product formation of the reconstituted T4MO complex (Table 4). Furthermore, catalytic complementation of the T4MO complex with an effector protein from each subclass has now shown that only TbuV, an effector protein from an enzyme complex that is in the same subclass as T4MO, was able to give catalytic activity (*22*, *68*). Specifically, reconstitution studies performed in this work with MmoB and ΔN29-MmoB, which likely have structures closely similar to T4moD (Figure 3 and Table 2), showed no catalytic complementation. These results indicate that, even as the three-dimensional structures of the effector proteins are relatively conserved, highly specific residue-dependent contacts will likely be required for function (Figure 5).

*Fidelity between Catalytic Effector Proteins.* The distribution of charged residues on the surface of the α1 helices is different among the effector proteins (Figures 3 and 5). For example, TbuV, which cross-reacts with T4moH, has similarly charged residues as the T4moD. In contrast, MmoB, which does not cross-react with T4moH, has several residues that are either opposite in charge, polar, or even hydrophobic in places that correspond to Asp or Glu residues (Figure 5, red) in the T4moD crystal structures. Moreover, a positive electrostatic patch observed in T4moD (Figure 5, blue) is derived from Arg18, Arg44, Arg45, and Arg49. Amino acid sequence alignments suggest that similarly charged residues would be located in analogous positions in TbuV. In contrast, the MmoB proteins do not have Arg or Lys residues that correspond with these four arginine residues. Thus, we hypothesize that the α1 helix and electrostatic interactions may be important contributors to specific complex formation between T4moD and T4moH. Changes in these specific interactions may also be responsible for the lack of catalytic complementation of T4moH by the MmoB components. In T4moD, the α1 helix is flanked on one side by the *β*2 strand and on the other by the α2 helix (Figure 5). It is interesting to note in the analysis of the crystal packing that an exposed *β*2 strand forms an antiparallel *β*-sheet across the asymmetric unit interface in both space groups. If the α1 helix and the *β*2-strand are critical features for complex formation, their interaction with T4moH may align the hydrophobic cavity in T4moD and a substrate access tunnel proposed from the recent crystal structure for the hydroxylase component of the toluene/*o*-xylene monooxygenase from *Pseudomonas stutzeri* OX1 (*9*).

*Summary.* The effector proteins have no cofactors or metals yet are able to exert exquisite control over the reactivity of the multiprotein diiron hydroxylase complex. The combination of X-ray and NMR structures now available indicates that the effector proteins have a highly homologous three-dimensional structure, but there are profound differences in ability of the structurally conserved proteins to provide heterologous complementation of catalytic activity. The high-resolution X-ray structures presented here have helped to define the potential differences in electrostatic surfaces that may govern the feasibility of protein−protein interactions. Moreover, the present structures reveal a single, well-defined cavity of size suitable for toluene binding near a region of the protein surface that has substantially different electrostatic properties among the effector protein family members. These results give new information and raise important new questions about these small, enigmatic components of the diiron hydroxylase enzyme complexes.

## ACKNOWLEDGMENT

We thank Dr. B. Wallar and Prof. J. D. Lipscomb for the gift of natural MmoB and the ΔN29-MmoB used in reconstitution studies. We also thank Dr. L. Pedersen for producing the first crystals of native T4moD, Dr. J. Thoden for collecting the native T4moD data set, and Prof. H. Holden and Prof. I. Rayment for generous access to their laboratory and personnel while these initial studies were being undertaken.

## REFERENCES

1. Pikus, J. D., Studts, J. M., Achim, C., Kauffmann, K. E., Münck, E., Steffan, R. J., McClay, K., and Fox, B. G. (1996) Recombinant toluene-4-monooxygenase: catalytic and Mössbauer studies of the purified diiron and Rieske components of a four-protein complex, *Biochemistry 35*, 9106−9119.
2. Leahy, J. G., Batchelor, P. J., and Morcomb, S. M. (2003) Evolution of the soluble diiron monooxygenases, *FEMS Microbiol. Rev. 27*, 449−479.
3. Baik, M. H., Newcomb, M., Friesner, R. A., and Lippard, S. J. (2003) Mechanistic studies on the hydroxylation of methane by methane monooxygenase, *Chem. Rev. 103*, 2385−2419.
4. Wallar, B. J., and Lipscomb, J. D. (1996) Dioxygen activation by enzymes containing binuclear non-heme iron clusters, *Chem. Rev. 96*, 2625−2658.
5. Lipscomb, J. D. (1994) Biochemistry of the soluble methane monooxygenase, *Annu. Rev. Microbiol. 48*, 371−399.
6. Colby, J., Dalton, H., and Whittenbury, R. (1979) Biological and biochemical aspects of microbial growth on C1 compounds, *Annu. Rev. Microbiol. 33*, 481−517.
7. Merkx, M., Kopp, D. A., Sazinsky, M. H., Blazyk, J. L., Müller, J., and Lippard, S. J. (2001) Dioxygen activation and methane hydroxylation by soluble methane monooxygenase: A tale of two irons and three proteins, *Angew. Chem., Int. Ed. 40*, 2782−2807.
8. Elango, N., Radhakrishnan, R., Froland, W. A., Wallar, B. J., Earhart, C. A., Lipscomb, J. D., and Ohlendorf, D. H. (1997) Crystal structure of the hydroxylase component of methane monooxygenase from *Methylosinus trichosporium* OB3b, *Protein Sci. 6*, 556−568.
9. Sazinsky, M. H., Bard, J., Di Donato, A., and Lippard, S. J. (2004) Crystal structure of the toluene/*o*-xylene monooxygenase hydroxylase from *Pseudomonas stutzeri* OX1. Insight into the substrate specificity, substrate channeling, and active site tuning of multicomponent monooxygenases, *J. Biol. Chem. 279*, 30600−30610.
10. Rosenzweig, A. C., Frederick, C. A., Lippard, S. J., and Nordlund, P. (1993) Crystal structure of a bacterial non-haem iron hydroxylase that catalyses the biological oxidation of methane, *Nature 366*, 537−543.
11. Skjeldal, L., Peterson, F. C., Doreleijers, J. F., Moe, L. A., Pikus, J. D., Westler, W. M., Markley, J. L., Volkman, B. F., and Fox, B. G. (2004) Solution structure of T4moC, the Rieske ferredoxin component of the toluene 4-monooxygenase complex, *J. Biol. Inorg. Chem. 9*, 945−953.
12. Müller, J., Lugovskoy, A. A., Wagner, G., and Lippard, S. J. (2002) NMR structure of the [2Fe-2S] ferredoxin domain from soluble methane monooxygenase reductase and interaction with its hydroxylase, *Biochemistry 41*, 42−51.
13. Chatwood, L. L., Müller, J., Gross, J. D., Wagner, G., and Lippard, S. J. (2004) NMR structure of the flavin domain from soluble methane monooxygenase reductase from *Methylococcus capsulatus* (Bath), *Biochemistry 43*, 11983−11991.
14. Fox, B. G., Liu, Y., Dege, J. E., and Lipscomb, J. D. (1991) Complex formation between the protein components of methane monooxygenase from *Methylosinus trichosporium* OB3b. Identification of sites of component interaction, *J. Biol. Chem. 266*, 540−550.
15. Froland, W. A., Andersson, K. K., Lee, S. K., Liu, Y., and Lipscomb, J. D. (1992) Methane monooxygenase component B

and reductase alter the regioselectivity of the hydroxylase component-catalyzed reactions. A novel role for protein−protein interactions in an oxygenase mechanism, *J. Biol. Chem. 267*, 17588−17597.
16. Green, J., and Dalton, H. (1985) Protein B of soluble methane monooxygenase from *Methylococcus capsulatus* (Bath). A novel regulatory protein of enzyme activity, *J. Biol. Chem. 260*, 15795−15801.
17. Liu, Y., Nesheim, J. C., Lee, S. K., and Lipscomb, J. D. (1995) Gating effects of component B on oxygen activation by the methane monooxygenase hydroxylase component, *J. Biol. Chem. 270*, 24662−24665.
18. Lloyd, J. S., Bhambra, A., Murrell, J. C., and Dalton, H. (1997) Inactivation of the regulatory protein B of soluble methane monooxygenase from *Methylococcus capsulatus* (Bath) by proteolysis can be overcome by a Gly to Gln modification, *Eur. J. Biochem. 248*, 72−79.
19. MacArthur, R., Sazinsky, M. H., Kuhne, H., Whittington, D. A., Lippard, S. J., and Brudvig, G. W. (2002) Component B binding to the soluble methane monooxygenase hydroxylase by saturation-recovery EPR spectroscopy of spin-labeled MMOB, *J. Am. Chem. Soc. 124*, 13392−13393.
20. Shu, L., Nesheim, J. C., Kauffmann, K., Münck, E., Lipscomb, J. D., and Que, L., Jr. (1997) An $Fe^{IV}_2O_2$ diamond core structure for the key intermediate Q of methane monooxygenase, *Science 275*, 515−518.
21. Chang, S. L., Wallar, B. J., Lipscomb, J. D., and Mayo, K. H. (1999) Solution structure of component B from methane monooxygenase derived through heteronuclear NMR and molecular modeling, *Biochemistry 38*, 5799−5812.
22. Hemmi, H., Studts, J. M., Chae, Y. K., Song, J., Markley, J. L., and Fox, B. G. (2001) Solution structure of the toluene 4-monooxygenase effector protein (T4moD), *Biochemistry 40*, 3512−3524.
23. Qian, H., Edlund, U., Powlowski, J., Shingler, V., and Sethson, I. (1997) Solution structure of phenol hydroxylase protein component P2 determined by NMR spectroscopy, *Biochemistry 36*, 495−504.
24. Walters, K. J., Gassner, G. T., Lippard, S. J., and Wagner, G. (1999) Structure of the soluble methane monooxygenase regulatory protein B, *Proc. Natl. Acad. Sci. U.S.A. 96*, 7877−7882.
25. Brandstetter, H., Whittington, D. A., Lippard, S. J., and Frederick, C. A. (1999) Mutational and structural analyses of the regulatory protein B of soluble methane monooxygenase from *Methylococcus capsulatus* (Bath), *Chem. Biol. 6*, 441−449.
26. Chang, S. L., Wallar, B. J., Lipscomb, J. D., and Mayo, K. H. (2001) Residues in *Methylosinus trichosporium* OB3b methane monooxygenase component B involved in molecular interactions with reduced- and oxidized-hydroxylase component: a role for the N-terminus, *Biochemistry 40*, 9539−9551.
27. Wallar, B. J., and Lipscomb, J. D. (2001) Methane monooxygenase component B mutants alter the kinetics of steps throughout the catalytic cycle, *Biochemistry 40*, 2220−2233.
28. Brazeau, B. J., and Lipscomb, J. D. (2003) Key amino acid residues in the regulation of soluble methane monooxygenase catalysis by component B, *Biochemistry 42*, 5618−5631.
29. Brazeau, B. J., Wallar, B. J., and Lipscomb, J. D. (2003) Effector proteins from P450(cam) and methane monooxygenase: lessons in tuning nature's powerful reagents, *Biochem. Biophys. Res. Commun. 312*, 143−148.
30. Callaghan, A. J., Smith, T. J., Slade, S. E., and Dalton, H. (2002) Residues near the N-terminus of protein B control autocatalytic proteolysis and the activity of soluble methane mono-oxygenase, *Eur. J. Biochem. 269*, 1835−1843.
31. Kopp, D. A., Berg, E. A., Costello, C. E., and Lippard, S. J. (2003) Structural features of covalently cross-linked hydroxylase and reductase proteins of soluble methane monooxygenase as revealed by mass spectrometric analysis, *J. Biol. Chem. 278*, 20939−20945.
32. Brandstetter, H., Whittington, D. A., Lippard, S. J., and Frederick, C. A. (1999) Mutational and structural analyses of the regulatory protein B of soluble methane monooxygenase from *Methylococcus capsulatus* (Bath), *Chem. Biol. 6*, 441−449.
33. Coufal, D. E., Blazyk, J. L., Whittington, D. A., Wu, W. W., Rosenzweig, A. C., and Lippard, S. J. (2000) Sequencing and analysis of the *Methylococcus capsulatus* (Bath) soluble methane monooxygenase genes, *Eur. J. Biochem. 267*, 2174−2185.
34. Pikus, J. D., Mitchell, K. H., Studts, J. M., McClay, K., Steffan, R. J., and Fox, B. G. (2000) Threonine 201 in the diiron enzyme toluene 4-monooxygenase is not required for catalysis, *Biochemistry 39*, 791−799.
35. Pikus, J. D., Studts, J. M., McClay, K., Steffan, R. J., and Fox, B. G. (1997) Changes in the regiospecificity of aromatic hydroxylation produced by active site engineering in the diiron enzyme toluene 4-monooxygenase, *Biochemistry 36*, 9283−9289.
36. Studts, J. M., and Fox, B. G. (1999) Application of fed-batch fermentation to the preparation of isotopically labeled or selenomethionyl-labeled proteins, *Protein Expression Purif. 16*, 109−119.
37. Studts, J. M., Mitchell, K. H., Pikus, J. D., McClay, K., Steffan, R. J., and Fox, B. G. (2000) Optimized expression and purification of toluene 4-monooxygenase hydroxylase, *Protein Expression Purif. 20*, 58−65.
38. Orville, A. M., Studts, J. M., Lountos, G. T., Mitchell, K. H., and Fox, B. G. (2003) Crystallization and preliminary analysis of native and N-terminal truncated isoforms of toluene-4-monooxygenase catalytic effector protein, *Acta Crystallogr. D59*, 572−575.
39. Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution, *Acta Crystallogr. D55*, 849−861.
40. de la Fortelle, E., and Bricogne, G. (1997) Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods, *Methods Enzymol. 276*, 472−494.
41. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D50*, 760−763.
42. Jones, T. A., and Kjeldgaard, M. (1997) Electron-density map interpretation, *Methods Enzymol. 277*, 173−208.
43. Kleywegt, G. J., and Jones, T. A. (1997) Model building and refinement practice, *Methods Enzymol. 277*, 208−230.
44. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
45. Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S., and Dodson, E. J. (1999) Efficient anisotropic refinement of macromolecular structures using FFT, *Acta Crystallogr. D55*, 247−255.
46. Pannu, N. S., Murshudov, G. N., Dodson, E. J., and Read, R. J. (1998) Incorporation of prior phase information strengthens maximum-likelihood structure refinement, *Acta Crystallogr. D54*, 1285−1294.
47. Read, R. J. (1997) Model phases: Probabilities and bias, *Methods Enzymol. 277*, 110−128.
48. Lamzin, V. S., and Wilson, K. S. (1993) Automated refinement of protein models, *Acta Crystallogr. D49*, 129−147.
49. Vaguine, A. A., Richelle, J., and Wodak, S. J. (1999) SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model, *Acta Crystallogr. D55*, 191−205.
50. Morris, A. L., MacArthur, M. W., Hutchinson, E. G., and Thornton, J. M. (1992) Stereochemical quality of protein structure coordinates, *Proteins 12*, 345−364.
51. Laskowski, R. A., Moss, D. S., and Thornton, J. M. (1993) Main-chain bond lengths and bond angles in protein structures, *J. Mol. Biol. 231*, 1049−1067.
52. Ferrin, T. E., Huang, C. C., Jarvis, L. E., and Langridge, R. (1988) The midas database system, *J. Mol. Graphics 6*, 2−12.
53. Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling, *Electrophoresis 18*, 2714−2723.
54. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers 22*, 2577−2637.
55. Moe, L. A., Hu, Z., Deng, D., Austin, R. N., Groves, J. T., and Fox, B. G. (2004) Remarkable aliphatic hydroxylation by the diiron enzyme toluene 4-monooxygenase in reactions with radical or cation diagnostic probes norcarane, 1,1-dimethylcyclopropane, and 1,1-diethylcyclopropane, *Biochemistry 43*, 15688−15701.
56. Eriksson, A. E., Baase, W. A., Wozniak, J. A., and Matthews, B. W. (1992) A cavity-containing mutant of T4 lysozyme is stabilized by buried benzene, *Nature 355*, 371−373.
57. Morton, A., Baase, W. A., and Matthews, B. W. (1995) Energetic origins of specificity of ligand binding in an interior nonpolar cavity of T4 lysozyme, *Biochemistry 34*, 8564−8575.

58. Morton, A., and Matthews, B. W. (1995) Specificity of ligand binding in a buried nonpolar cavity of T4 lysozyme: linkage of dynamics and structural plasticity, *Biochemistry 34*, 8576−8588.
59. Xu, J., Baase, W. A., Quillin, M. L., Baldwin, E. P., and Matthews, B. W. (2001) Structural and thermodynamic analysis of the binding of solvent at internal sites in T4 lysozyme, *Protein Sci. 10*, 1067−1078.
60. Sagermann, M., Baase, W. A., and Matthews, B. W. (1999) Structural characterization of an engineered tandem repeat contrasts the importance of context and sequence in protein folding, *Proc. Natl. Acad. Sci. U.S.A. 96*, 6078−6083.
61. Sagermann, M., Baase, W. A., Mooers, B. H., Gay, L., and Matthews, B. W. (2004) Relocation or duplication of the helix A sequence of T4 lysozyme causes only modest changes in structure but can increase or decrease the rate of folding, *Biochemistry 43*, 1296−1301.
62. Sagermann, M., and Matthews, B. W. (2002) Crystal structures of a T4-lysozyme duplication-extension mutant demonstrate that the highly conserved beta-sheet region has low intrinsic folding propensity, *J. Mol. Biol. 316*, 931−940.
63. Juers, D. H., and Matthews, B. W. (2001) Reversible lattice repacking illustrates the temperature dependence of macromolecular interactions, *J. Mol. Biol. 311*, 851−862.
64. Zhang, X. J., Wozniak, J. A., and Matthews, B. W. (1995) Protein flexibility and adaptability seen in 25 crystal forms of T4 lysozyme, *J. Mol. Biol. 250*, 527−552.
65. Gherman, B. F., Baik, M. H., Lippard, S. J., and Friesner, R. A. (2004) Dioxygen activation in methane monooxygenase: a theoretical study, *J. Am. Chem. Soc. 126*, 2978−2990.
66. Brazeau, B. J., Austin, R. N., Tarr, C., Groves, J. T., and Lipscomb, J. D. (2001) Intermediate Q from soluble methane monooxygenase hydroxylates the mechanistic substrate probe norcarane: evidence for a stepwise reaction, *J. Am. Chem. Soc. 123*, 11831−11837.
67. Brazeau, B. J., and Lipscomb, J. D. (2000) Kinetics and activation thermodynamics of methane monooxygenase compound Q formation and reaction with substrates, *Biochemistry 39*, 13503−13515.
68. Mitchell, K. H., Studts, J. M., and Fox, B. G. (2002) Combined participation of hydroxylase active site residues and effector protein binding in a para to ortho modulation of toluene 4-monooxygenase regiospecificity, *Biochemistry 41*, 3176−3188.
69. Kelley, L. A., and Sutcliffe, M. J. (1997) OLDERADO: on-line database of ensemble representatives and domains, *Protein Sci. 6*, 2628−2630.

BI047459G