BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

Structure of T4moF, the toluene 4-monooxygenase ferredoxin oxidoreductase

Justin Finley Acheson, Hannah Moseson, and Brian G. Fox

Biochemistry, **Just Accepted Manuscript** • DOI: 10.1021/acs.biochem.5b00692 • Publication Date (Web): 26 Aug 2015

ACS Publications

Biochemistry is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Structure of T4moF, the toluene 4-monooxygenase ferredoxin oxidoreductase

*Justin F. Acheson, Hannah Moseson & Brian G. Fox**

Department of Biochemistry, College of Agriculture and Life Sciences, University of Wisconsin–Madison, Madison, Wisconsin 53706, USA

RUNNING TITLE Oxidoreductase structures in diiron hydroxylase reaction

Corresponding Author
* To whom correspondence should be addressed: Department of Biochemistry, 433 Babcock Dr., University of Wisconsin, Madison, Wisconsin 53706. Phone: (608) 262-9708. Fax: (608) 262-3453. E-mail: bgfox@biochem.wisc.edu

ABBREVIATIONS

BMM, bacterial multicomponent monooxygenase; T4MO, four-protein toluene 4-monooxygenase complex from *Pseudomonas mendocina* KR1; T4moH, hydroxylase component of T4MO; T4moD, effector protein of T4MO; T4moHD, stoichiometric complex of T4moH and T4moD; T4moC, Rieske-type ferredoxin of the T4MO complex; T4moHC, stoichiometric complex of T4moH and T4moC; T4moF, NADH oxidoreductase of the T4MO complex; RMSD, root mean square deviation.

ABSTRACT The 1.6 Å crystal structure of toluene 4-monooxygenase reductase T4moF is reported. The structure includes ferredoxin, flavin, and NADH binding domains. The position of the ferredoxin domain relative to the other two domains represents a new configuration for the iron-sulfur flavoprotein family. Close contacts between the C8 methyl group of FAD and [2Fe-2S] ligand Cys36-O represent a plausible pathway for electron transfer between the redox cofactors. Energy-minimized docking of NADH and calculation of hinge-like motions between domains suggest how simple coordinated shifts of residues at the C-terminus of the enzyme could expose the N5 position of FAD for productive interaction with the nicotinamide ring. The domain configuration revealed by the T4moF structure provides an excellent steric and electrostatic match to the obligate electron acceptor, Rieske-type [2Fe-2S] ferredoxin T4moC. Protein-protein docking and energy minimization of the T4moFC complex indicate that T4moF [2Fe-2S] ligand Cys41 and T4moC [2Fe-2S] ligand His67, along with other electrostatic interactions between the protein partners, form the functional electron transfer interface.

KEYWORDS diiron enzyme; oxidoreductase, iron-sulfur flavoprotein, electron transfer; protein-protein complex.

Bacterial multicomponent monooxygenases (BMMs) belong to a class of soluble diiron enzymes that carry out hydroxylation of unactivated C-H bonds and aromatic rings (*1*). Often, these enzymes use O_2 and exogenous electrons to oxidize a substrate as the initial activation step of a microbial metabolic pathway (*2*). The BMMs are of great interest due to their high reactivity, and the unique mechanistic and structural features of their diiron centers. Further motivations for their study include potential biotechnological uses in the remediation of xenobiotic compounds and in the synthesis of fine chemicals.

BMMs require a minimum of three proteins: a flavin-containing ferredoxin:NAD^+ oxidoreductase; a diiron hydroxylase; and a cofactorless effector protein that is unique to this family of enzymes (*2, 3*). Four component BMMs require an additional Rieske-type ferredoxin that serves as an intermediate electron carrier between the oxidoreductase and the hydroxylase (*4*). Typically, the oxidoreductase accepts a hydride equivalent from NADH and then delivers reducing equivalents via the intermediacy of its [2Fe-2S] cluster to either the diiron center or the Rieske-type ferredoxin. The effector protein, O_2, and substrate bind to the reduced hydroxylase to form a complex that carries out catalysis.

Toluene 4-monooxygenase (T4MO) is a four component BMM found in *Pseudomonas mendocina* KR1 (*5*). T4MO catalyzes regiospecific oxidation of toluene and produces *p*-cresol with ~97% yield (*6*). This initial oxidation step allows *P. mendocina* to grow on toluene as its sole carbon source. There are four T4MO proteins. T4moF, a 37 kDa FAD- and [2Fe-2S]-containing oxidoreductase transfers electrons to T4moC, a 12 kDa Rieske-type ferredoxin. T4moC transfers electrons to T4moH, a 210 kDa diiron hydroxylase. T4moD, an 11 kDa cofactorless effector protein, binds to T4moH to advance the catalytic cycle.

X-ray structures of T4moC (*7*), T4moD (*8*), T4moH (*9*), as well as T4moH bound to either T4moD (*9*) or T4moC (*10*) have been previously reported. Here we report the crystal structure of T4moF solved at a resolution of 1.6 Å. This is the first structure of an intact BMM oxidoreductase containing both Fd and FAD/NADH binding domains, and also completes the set of X-ray structures for all members of the T4MO complex. Although T4moF has structural similarities to the benzoate dioxygenase (BenC) (*11*) and phthalate dioxygenase (PDR) (*12*) reductases, the location of the Fd domain and closest approach between [2Fe-2S] and FAD differs from both of these earlier structures. Using this new structure as a starting point, we used molecular docking approaches to investigate the binding of NADH to T4moF and also the formation of the electron transfer complex between T4moF and T4moC. These new results provide a useful picture of these critical, early complexes in the BMM catalytic cycle.

MATERIALS AND METHODS

Plasmid construction. The T4moF gene was amplified from pT4moABEF (*13*) using primers that added the 5'-Sgf1 and 3'-PmeI restriction sites used for FlexiVector cloning. The amplified gene product and pVP68K, an expression plasmid that fuses maltose binding protein (MBP) to the N-terminus of a target protein (*14, 15*), were digested with Flexi-cloning blend restriction enzymes (Promega, Madison WI) and ligated with T4 DNA ligase (Promega, Madison, WI), creating expression plasmid pVP68KT4moF (Figure S3). pVP68KT4moF was further modified to introduce two mutations, Lys270Ser and Lys271Ser, using QuikChange site-directed mutagenesis (Stratagene, Santa Clara, CA). The mutated plasmid is called pF270_271S.

Protein Expression. pF270S_271S was transformed into *E. coli* BL21 RILP (Stratagene) and the cells were grown in a Bioflow 110 bioreactor (New Brunswick Science, NJ) as

previously described for T4moH (*10*). Briefly, 2 × 2 mL of non-inducing minimal medium was inoculated in the morning and grown at 37˚C (*15*). In the evening, 2 × 500 mL of the same medium was inoculated with the 2 mL starter cultures and grown overnight at 25˚C. In the morning, both 500 mL cultures were transferred to a Bioflow 110 14 L vessel containing 9 L of 5M medium (*15*) at 37˚C with O_2 levels maintained by agitation. When the cells reached an OD_{600} of ~2.5, the temperature was reduced to 25˚C and expression was induced by the addition of 200 μM IPTG, 20 g Casamino acids, and 36 g of lactose. The total expression time was 5 h, after which the medium was collected and centrifuged in 1 L jars by centrifugation at 4200 rpm for 25 min. The 10 L fermentation typically yielded ~45 g of cell paste.

Protein Purification. The cell paste was thawed on ice, and then suspended in 2 mL of 25 mM MOPS, pH 7.5, containing 50 mM NaCl and 2% (v/v) glycerol (buffer A) per gram of cell paste. The cell suspension was sonicated on ice with alternating 15 s on and 30 s rest periods for a total of 10 min of sonication on time. The sonicated cell suspension was centrifuged at 48000 × *g* and 4˚C in a JA-20.50 rotor and Avanti J30-I centrifuge (Beckman Coulter) for 60 min. The supernatant was then diluted two-fold with buffer A and applied to a 45 mm × 250 mm DEAE Sepharose column (GE Healthcare, Piscataway NJ) at 5 mL/min. The column was washed with 2 column volumes of buffer A and T4moF was then eluted in a 1200 mL linear gradient of 50 to 450 mM NaCl in buffer A at 5 mL/min. The fractions containing a brownish orange color, indicative of the FAD and [2Fe-2S] cluster of T4moF, eluted at ~300 mM NaCl. These fractions were assayed for activity and pooled based on the T4MO-catalyzed oxidation of nitrobenzene to *p*-nitrophenol (*16*).

Pooled fractions were applied to an XK 26/40 (GE Healthcare, Piscataway NJ) amylose column (New England Biolabs, Ipswitch, MA) installed on an Akta Purifier (GE Healthcare,

Piscataway NJ) equipped with Unicorn 5.0 software. The column was equilibrated with 25 mM MOPS, pH 7.5, containing 250 mM NaCl and 2% (v/v) glycerol (buffer B) at 2 mL/min. Unbound protein was removed by washing with buffer B until the absorbance at 280 nm returned to baseline for at least 2 column volumes. The bound protein was eluted in buffer B with the addition of 10 mM maltose. The orange-brown peak was collected, yielding nearly pure MBP-T4moF fusion protein.

The MBP-T4moF fusion protein was diluted to 100 mL and 3C protease was added at 4˚C for ~24 h to allow liberation of T4moF. Proteolysis was confirmed by SDS-PAGE. Imidazole was added to the proteolyzed sample to 25 mM and then the sample was passed over a Ni-NTA column equilibrated with buffer B containing 25 mM imidazole. The column flow-through, containing T4moF, was collected and concentrated using either a Centricon YM-10 or Amicon 10K spin column. The concentrated protein was then diluted with 10 mM MOPS, pH 7.5, containing 50 mM NaCl and reconcentrated. The buffer-exchanged protein was concentrated to ~15 mg/mL, drop frozen in liquid N_2, and stored at -80˚C.

Crystallization and Structure Determination. T4moF crystals were obtained by hanging drop vapor diffusion after mixing 1.5 μL of a ~15 mg/mL protein solution with 1.5 μL of 100 mM Bis-Tris, pH 5.5, containing 200 mM ammonium acetate, 16% PEG 3350, 5 mM $NiCl_2$, and 50 mM guanidine-HCl. Tapered yellow-brown rods grew within 2-3 days to a size suitable for diffraction experiments. Crystals were cryo-protected by passing through Fomblin 2500 and frozen in liquid N_2. Diffraction data were collected at Life Sciences–Collaborative Access Team (LS–CAT) at the Advance Photon Source (APS), Argonne National Laboratory. The data were indexed, integrated, and scaled using HKL2000 (*17*). Molecular replacement was achieved by using Swissmodel (*18*) to create homology models to the Fd and FAD/NADH

binding domains of PDB 1KRH [BenC, benzoate dioxygenase reductase (*11*)], and by then treating both domains as independent units in Phaser-MR (*19*). The electron density was modeled using iterative rounds of Phenix.refine (*19*) and Coot (*20*). Molprobity (*21*) was used to assess statistics and overall quality of the structure. Figures were prepared with Pymol (*22*).

NADH Docking. T4moF was aligned with FNR complexed with $NADP^+$ (PDB 1QFZ) (*23*) using PyMOL to provide an initial model for how NADH would bind to T4moF. Some alterations in the position of the predicted position of NADH bound to T4moF were introduced after inspection of protein-ligand contacts. After the initial placement was complete, the position and geometry of the NADH placed into the T4moF structure was minimized using Phenix Geometry Minimization (*19*). The protein-protein docking web server *ClusPro* (http://nrc.bu.edu/cluster/) was used to obtain 3D-models of the complex of T4moF with T4moC (*24, 25*). This server provides algorithms that filter docked conformations based on shape complementarity, desolvation and electrostatic energies. The PDB coordinates obtained for T4moF in this work were used as the static "receptor " molecule and the file 1VM9 contained the coordinates for T4moC (*7*) used as the moving "ligand" molecule.

RESULTS AND DISCUSSION

Vector Construction. The expression and purification of T4moF has been challenging compared to other T4MO proteins. In the natural multi-cistronic message, the expression of T4moF is attenuated by the presence of rho-independent terminator sequences (*26*). During overexpression in *Escherichia coli*, it formed inclusion bodies, and during purification it often lost the FAD cofactor. An early vector, pUCT4moF, gave an appreciable amount of soluble protein (~1 mg/g of cell paste). However, the low yield of an untagged enzyme made purification

tedious (*13*). The adoption of pVP68K (*14*) alleviated these problems by providing tight control of basal expression in the presence of non-inducing medium, strong induction in the presence of IPTG and lactose, and improved handling as an MBP fusion protein. Furthermore, the use of amylose affinity chromatography in the second purification step avoided the need to expose the enzyme to high concentrations of imidazole. These process improvements increased the final yield of T4moF to ~2 mg/g of cell paste after release of intact T4moF by proteolysis of the fusion protein. However, even with this streamlined purification and increased yield, extensive screening for crystallization yielded no positive results.

Surface Entropy Reduction. A homology model prepared in Swissmodel (*18*) from the closest homolog BenC (*11*) (PDB 1KRH) was inspected for possible residue targets for surface entropy reduction (*27*). From this, Lys270 and Lys271 were located on a predicted surface loop, and were replaced with serine. After expression and purification as described in Materials and Methods, crystallization screening with the doubly mutated enzyme yielded brown rod-like crystals.

Structure Determination. T4moF crystals belonged to the C2 space group and diffraction data extending to 1.6 Å were collected. Molecular replacement using BenC (PDB 1KRH) as the template was unsuccessful. In order to achieve the molecular replacement, the loop connecting the Fd and FAD/NADH binding domains was removed and homology models for the two separated domains were then treated as independent proteins for the molecular replacement. Using the split domains, a phasing solution was identified and the structure was solved. Ultimately, the results will show that the Fd and FAD/NADH binding domains in T4moF have a considerably different orientation than in BenC, which likely caused failure of the standard molecular replacement to identify a phasing solution.

The data collection and refinement statistics for T4moF are presented in Table 1. The diffraction data had I/σ of 13.08 (3.94 in the highest shell) and was 96.56% complete. The structure was refined to a resolution of 1.6 Å, and the R_{work} and R_{free} were 0.1447(0.1562) and 0.1882(0.2231), respectively, with values in highest resolution shell shown in parentheses. There is a single monomer of T4moF in the asymmetric unit, and the atoms in the FAD and [2Fe-2S] cluster had full occupancy. The addition of Ni^{2+} to the T4moF crystallization buffer improved the quality of the crystals, and we found that three solvent-exposed His residues, His12, His98 and His278, coordinate a single Ni^{2+} that was also coordinated by a molecule of Bis-Tris from the crystallization buffer. In addition, solvent-exposed His11 also coordinates a single Ni^{2+}, but no Bis-Tris molecule as a consequence of crystal contacts.

Table 1. Data collection and refinement statistics for T4moF

	T4moF 4WGM
Data collection	
Space group	C2
Cell dimensions	
a, *b*, *c* (Å)	84.33, 69.07, 68.22
α, β, γ (°)	90, 116.53, 90
Resolution (Å)	27.12 -1.615 (1.67 -1.62)*
R_{sym} or R_{merge}	0.066 (0.37)
I / σI	13.08 (3.94)
Completeness (%)	96.56 (91.2)
Redundancy	3.9 (3.7)

Refinement	
Resolution (Å)	27.12 -1.62 (1.67 -1.62)
No. reflections	43392 (4063)
R_{work} / R_{free}	0.145 (0.156)/ 0.188 (0.223)
No. atoms	3083
Protein	2560
Ligand/ion	103
Water	420
B-factors	
Protein	23.10
Ligand/ion	27.70
Water	34.90
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	1.15

*Values in parentheses are for highest-resolution shell, which was near to the detector edge. Data was collected from a single crystal.

Structure of T4moF. T4moF is a three-domain 326 residue monomeric enzyme (Figure 1) belonging to the large family of iron-sulfur flavoproteins. In T4moF, the N-terminal plant-type Fd domain consists of 90 residues (Figure 1, *gold* cartoon). The C-terminal domain, called an FNR-like domain, belongs to the ferredoxin-$NADP^+$ reductase superfamily (*28*), and contains both FAD (Figure 1, *blue* cartoon) and NADH binding domains (Figure 1, *maroon* cartoon) that

are separated by a 12 residue linker (*green* cartoon, residues 190-201). The Fd and FNR-like domains are connected by an 12 residue linker (Figure 1, *cyan* cartoon, residues 89-100).

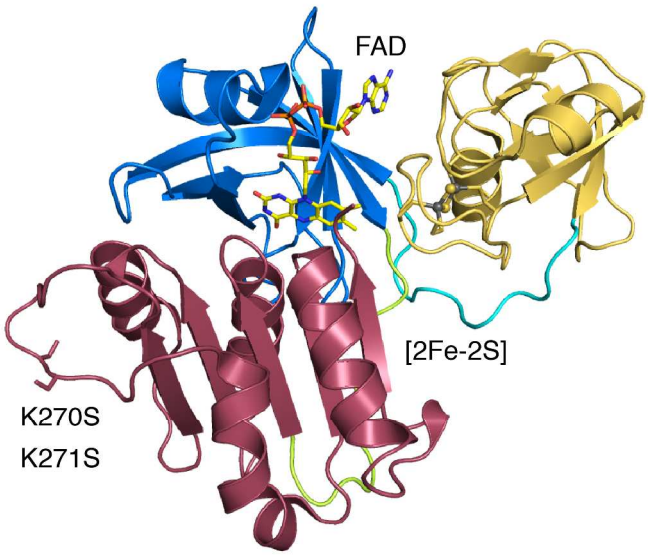

Figure 1. Structure of T4moF. The Fd domain is shown in *gold*, FAD binding domain in *blue*, NADH binding domain in *maroon*, and loops connecting the domains in *cyan* and *green*. The positions of bound FAD (*yellow* sticks), the [2Fe-2S] cluster (*grey* and *yellow* spheres), and mutations K270S and K271S, introduced to decrease surface entropy, are also shown.

Figure S1 shows a phylogenetic tree of the relevant subclade of the iron-sulfur flavoprotein family. Prior to this work, X-ray structures were known for BenC (*11*) and PDR (*12*), two homologous oxidoreductases that partner with Rieske and Fe^{2+}-dependent dioxygenases, and NMR structures were also known for the separated Fd (*29*) and FNR-like (*30*) domains of MmoC. Figure S2 shows a structure-annotated sequence alignment of enzymes most closely related to T4moF whose functions have been established (*31*). T4moF and TomoF are closely related oxidoreductases from aromatic ring hydroxylating BMMs, BenC and MmoC are closely related but from Rieske and Fe^{2+}-dependent dioxygenase and methane oxidizing BMM complexes, respectively, while PDR and PhP are more distantly related oxidoreductases from

two different enzyme complexes, i.e., a Rieske and Fe^{2+}-dependent dioxygenase and an aromatic ring hydroxylating BMM. T4moF, BenC and PhP have a similar arrangement of the Fd and FAD/NADH domains extending from the N-terminus of the enzyme, while PDR, the phylogenetically distant outlier, has a cyclic permutation of these domains to FAD/NADH followed by Fd.

FAD Domain and Binding. The FAD binding domain in T4moF has a similar structure to BenC (*11*), which is overall smaller than the comparable domains from other members of the FNR-like superfamily (*32*). A six-stranded β-sheet with an α-helix located between β5 and β6 provides the cleft needed to bind FAD (*blue* cartoon, Figure 2). Electron density of the FAD cofactor and the [2Fe-2S] cluster are also shown in Figure 2, along with the positions of other key residues in the active site. Residues from the FAD domain donate all observed hydrogen-bonding contacts to FAD isoalloxazine ring, with the identities of the bonding partners and the interatomic distances given in Table S2. The N1, N5 and N10 atoms in the bound FAD are planar (Figure 2), indicating that the flavin is in the oxidized state.

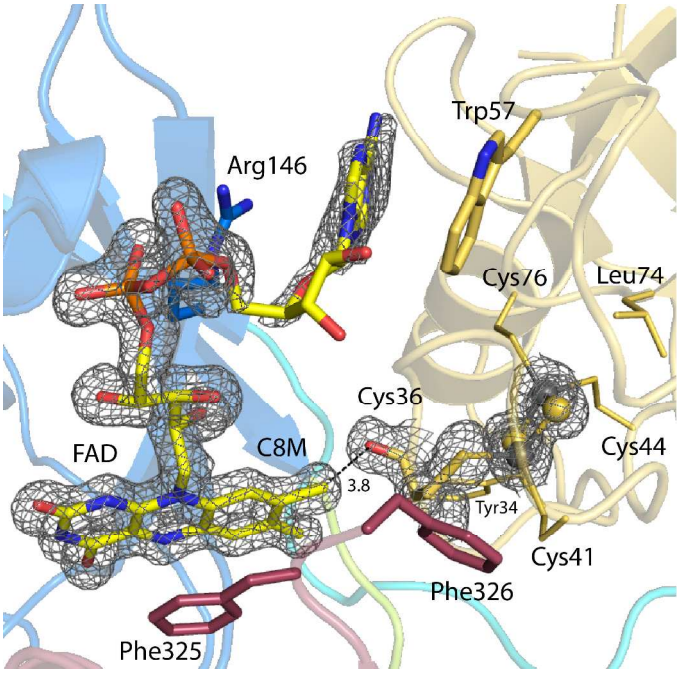

Figure 2. Relationship between redox cofactors in T4moF. Colors of individual domains are as in Figure 1. Electron density of redox cofactors at 1.5σ. Cysteine residues 36, 41, 44, and 76 coordinate the irons of the [2Fe-2S] cluster. Cys36-O provides the closest approach to FAD C8M, 3.8 Å. Trp-57 (*gold* sticks) from the Fd domain and Arg146 from the FAD domain position the adenine between the Fd and FAD domains. Phe325 from the NADH domain (*maroon* sticks) makes an important contact with FAD between the FAD/NADH domains. Phe326 is the C-terminus of the enzyme.

The atoms of the FAD isoalloxazine ring have lower average B-factors (~10) than the rest of the enzyme (~23-28, Table 1), indicating they are well ordered. Notable π-stacking interactions that hold the FAD isoalloxazine ring are provided by Tyr135 from the FAD domain and Phe325 from the NADH domain. Phe325 is homologous to Phe335 from BenC, Tyr308 from FNR, and Phe342 from MmoC (Figure S1). The role of Phe325 in NADH binding is described below.

The atoms of the adenine moiety of the FAD in T4moF have higher average B-factors (~50) than the remainder of the enzyme, likely indicating this portion of the cofactor may adopt multiple configurations. Indeed, there are relatively few interactions between the adenine and T4moF, with a cation-π interaction with Arg146 and a π-stacking with Trp57 of the Fd domain (Figure 2) providing the main interactions. This contribution of the Fd domain is unique compared to what is observed in other FNR-type proteins (MmoC PDB 1GAQ, FNR-Fd-1 PDB 1EWY, BenC PDB 1KRH, and cytochrome b_5 reductase PDB 1NDH) where the stabilizing interactions are provided by either the FAD domain only or by a combination of the FAD and NADH domains (*11, 30, 33-35*).

Ferredoxin Domain. In T4moF, the Fd domain contains a single [2Fe-2S] cluster coordinated by four cysteines in the Cys-X_4-Cys-X_2-Cys//Cys plant-type ferredoxin motif (*36*). Cysteine residues 36, 41, 44 and 76 coordinate the iron atoms of the [2Fe-2S] cluster, while residues 35, 37 and 39-42 provide hydrogen bonds to the sulfur atoms in the [2Fe-2S] cluster (see Table S1). The [2Fe-2S] cluster is also flanked by Tyr34 and Leu74, which protect it from solvent. The 12-residue coil (*cyan* cartoon) that connects the Fd (*gold* cartoon) and FAD (*blue* cartoon) domains makes extensive contacts with the FAD domain, and only minor contacts with the NADH binding domain (*maroon* cartoon).

Comparison of the structure of the T4moF Fd domain with other plant-type ferredoxins revealed that Cys36 has a binding conformation that is most similar to that observed in $1e^-$ reduced $[2Fe\text{-}2S]^{1+}$ clusters (*37*), i.e., where the carbonyl projects away from the [2Fe-2S] cluster (Figure 2). In the T4moF crystals, which were prepared in the oxidized state, photo-reduction from synchrotron radiation (*37*) may have reduced the [2Fe-2S] cluster without reducing the FAD, which is planar. In the observed configuration, Cys36-O provides the closest approach to

C8M of FAD (3.8 Å, Figure 2), suggesting the preferred contact for inter-domain electron transfer. HARLEM, a software package used to predict likely electron transfer pathways, supports that these two atoms are likely participants in the electron transfer reaction (*38*). However, since binding of NADH may alter the distances between the FAD and [2Fe-2S] cofactors, this may correspondingly change the output of the HARLEM calculation.

NADH Binding Interactions. The NADH binding domain contains a 5-stranded β-sheet with two α-helices on either side (Figure 1), which is typical of the FNR-like superfamily (*39*). Attempts to solve a T4moF structure with either NAD^+ or NADH bound were unsuccessful, but presumably for different reasons. Soaking with NAD^+ seemingly had no effect on the integrity of the crystals, but no electron density corresponding to bound NAD^+ was observed. In contrast, soaking with NADH caused the crystals to rapidly disintegrate, supporting evidence from other studies that NADH binding promotes a structural rearrangement in this family of enzymes, perhaps associated with a redox reaction (*40*). Moreover, co-crystallization of T4moF in the presence of either NAD^+ or NADH was not successful. Examination of the T4moF structure suggested that penultimate C-terminal residue Phe235 from the NADH domain would need to be displaced in order for the nicotinamide ring from NADH to bind in a productive manner for hydride transfer (Figure 3, *gray* cartoon and lines showing the position of Val193, Tyr276, Phe325 and Phe326 in the ligand-free enzyme).

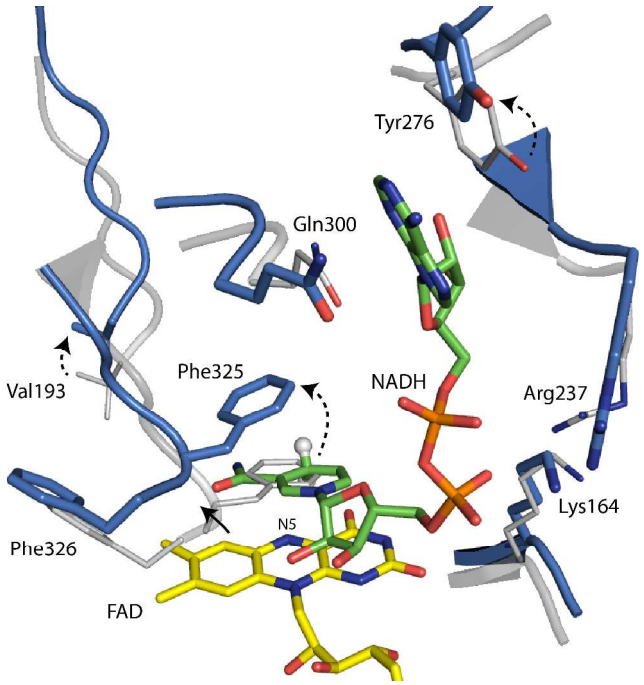

Figure 3. Docked position of NADH in T4moF. The ligand-free (*gray* cartoon and lines) and NADH-bound structures (*blue* cartoon and sticks) were aligned by their FAD (*yellow* sticks) groups. Positioning of the nicotinamide ring (*green* sticks) near to FAD N5 requires shifts in positions of Phe325, Phe326 and other residues indicated by *black* arrows.

To gain further insight into the possible binding mode used by NADH, FNR Tyr308Ser bound with $NADP^+$ [PDB 1QFZ, (*23*); with the indicated mutation at the position corresponding to T4moF Phe325], was aligned to T4moF and NADH was then modeled into the T4moF structure using $NADP^+$ as the template (*23*). Further adjustment of the position of the docked NADH was carried out after inspection of interactions with the surrounding residues (*22*). After initial placement in Pymol, energy minimization was carried using Phenix (*19*). The energy-minimized docking placed the adenine moiety in the cleft between Tyr276 and Gln300, while residues Lys164 and Arg237 coordinated the phosphate groups. Additionally, the phenyl ring of Phe325 was displaced by ~3.5 Å so that C4N of nicotinamide could reside within less than 3 Å

of FAD N5 in a position to facilitate transfer of the pro-R hydride to FAD. The hydride transfer reaction should also be facilitated by polarization of FAD N1, which is provided by a hydrogen bonding network to this position involving Tyr148-OH, FAD-O4, an active site water molecule, and FAD-O2. Docking of NADH also caused a shift in the positions of C-terminus Phe326 (shown) and Arg324, which has two conformations including a hydrogen bonding interaction with Glu93 (not shown). Notably, both Phe325 and Phe326 reside in a position between the FAD and Fd domains within 3 Å of both Cys36 and Cys41 bound to the [2Fe-2S] cluster.

The energy-minimized placement of NADH also caused the FAD and NADH binding domains to open slightly, as has already been predicted in other FNR-type proteins (*12*). To search for residues that could possibly contribute to this motion, a prediction was carried out with HingeProt (*41*). This analysis revealed Val193 (Figure 3) might be a hinge site in the linker connecting the FAD and NADH domains and thus participate in the needed repositioning; incidentally, a second hinge was predicted by this software at Gly94 between the Fd and FAD domains (not shown).

Normal mode analysis (*42*) supported the possibility of interdomain motions corresponding to those predicted by the docked structure and also by the hinge calculation. Thus mode 10 calculated by the elNemo server (frequency 1.74; collectivity 0.3519) moves the FAD and NADH domains apart, with the motion originating near Val193. This motion also pulls Phe325 away from its resting position above FAD N5, potentially facilitating binding of the nicotinamide ring in the proper position for the flavin reduction reaction. Another motion in this mode, localized near Gly94, moves the Fd domain away from the FAD domain, giving additional space for movement of Arg234 and Phe326 along with Phe325. Interestingly, mode 9 (frequency 1.32; collectivity 0.5860) moves the FAD and NADH domains back together, also

with hinge motion originating near Val193. This motion results in movement of Phe325 back toward its resting position above FAD N5, and so suggests how domain motions may also participate in the displacement of the nicotinamide ring of NAD^+ after reduction of the flavin.

Comparison with Other Iron-sulfur Flavoproteins. NMR structures of the separate ferredoxin (*29*) (Fd) and FAD/NADH binding (*30*) domains from the related oxidoreductase (MmoC) of three-component methane monooxygenase (MMO) have been reported. Early on, cross-linking of MmoC revealed an interaction with the beta subunit of MmoH (*43*). Later, solution NMR studies of the interactions between the MmoC-Fd domain and MmoH revealed that a small α-helix (α2) on MmoC-Fd was also likely to be important in the protein-protein interaction (*29*). This helix is ~25 Å away from the [2Fe-2S] cluster in MmoC. Although the Fd domain of T4moF lacks this helix, it does use a small loop in the corresponding region of the structure aligned with MmoC-Fd to interact with the adenine ring of FAD. Furthermore, in the T4moHC complex, T4moC makes contacts with a flexible loop of the T4moH beta chain ~25 Å from the [2Fe-2S] cluster (*10*). These distal contacts provide an important contribution to formation of a productive electron-transfer complex in T4moHC, and so distal interactions detected between MmoC and MmoH may potentially contribute in a similar way. Variations in distal recognition thus provide one way that the electron transfer proteins used in various BMMs and other redox-dependent enzymes can be more specifically targeted to different interaction partners (*10*).

A comparison of the structures of T4moF and the homologous, full-length oxidoreductases BenC (PDB 1KRH) and PDR (PDB 2PIA) was carried out. Alignment of the individual Fd domains (Table S3) showed an excellent agreement, with T4moF and BenC aligning with root mean square deviation (RMSD) of 0.624 $Å^2$. Likewise, comparison of the

individual FNR-like domains gave RMSD values less than 2 $Å^2$. The alignment of the FNR-like domains in the structures of full-length enzymes was used as a basis to investigate inter-domain relationships in the three enzymes. Since MmoC is closely related to BenC, the following comments on the structural homology of full-length BenC and T4moF may apply to MmoC as well.

Alignment of the FNR-like domains of T4moF with PDR (Figure 4A) and BenC (Figure 4B), revealed differences in the positions of their Fd domains. Inspection of the contacts to the Fd domains in these three structures suggested their positioning was unlikely to be an artifact of crystal packing. The T4moF structure revealed a new interaction between the Fd and FAD domains that brings Cys36-O, the first Cys residue in the plant-type [2Fe-2S] ferredoxin motif (indicated by roman numeral I), to within 3.7 Å of C8M of FAD and Cys36-CB to with 3.8 Å. Although the [2Fe-2S] cluster in PDR was positioned ~3 Å away from the [2Fe-2S] clusters in both T4moF and BenC, PDR had the O and CB atoms of the first Cys residue in the plant-type ferredoxin motif within 4.7 and 4.1 Å of C8M of FAD, respectively (Cys272 in PDR, indicated by roman numeral I). Further comparison revealed that the 5.4 Å closest approach to FAD C8M in BenC was provided by equidistant positioning of the O and CB atoms from the second Cys residue in the motif (Cys46 in BenC, indicated by roman numeral II).

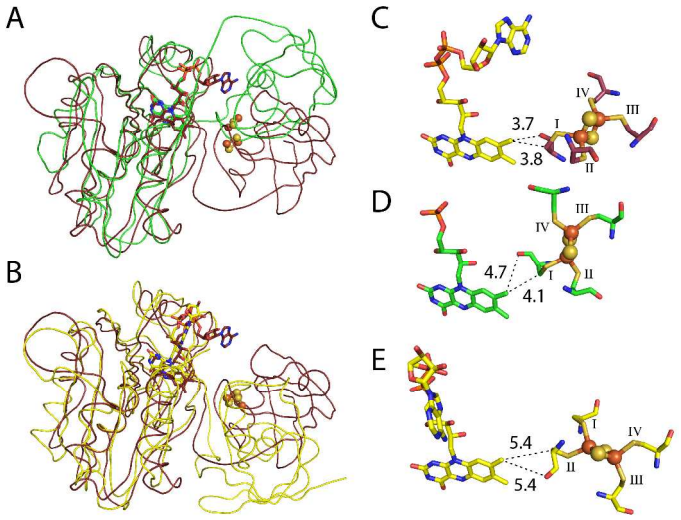

Figure 4. Alignment of T4moF with BenC and PDR. The FNR-like domain of T4moF (*red* ribbon) was aligned with BenC (panel A, *yellow* ribbon, PDB 1KRH) or PDR (panel B, *green* ribbon, PDB 2PIA). RMSDs for the alignments are found in Table S3. These alignments show that the Fd domain in T4moF adopts a different position with respect to the FNR-like domain than observed in the prior two structures. C, D and E, distances of the closest approach between flavin and [2Fe-2S] in T4moF, BenC and PDR, respectively. Roman numerals indicate the position of the residue in the plant-type [2Fe-2S] ferredoxin motif.

Interestingly, PDR has an FAD/NADH-Fd domain architecture, which is a cyclic permutation of the Fd-FAD/NADH domain arrangement observed in BenC and T4moF. As reported earlier (*12*), cyclic permutation of the three contributing domains does not dramatically alter the relative positions of the redox cofactors in the tertiary structure. However, the cyclic permutations may provide different possibilities for interactions of the Fd domain with subsequent proteins in their respective enzyme complexes, providing another mechanism for evolutionary adaption in the iron-sulfur flavoprotein family (*11*).

Predicted complex of T4moF and T4moC. With the availability of the T4moF structure, we asked whether the domain positions just described might provide a favorable shape for protein-protein interactions with T4moC, the Rieske-type ferredoxin that accepts electrons from T4moF. Even with no adjustment of the domain positions in T4moF, visual inspection showed there was remarkably good steric and electrostatic matching when T4moC was placed into the cleft between the Fd and NADH binding domains in T4moF. Correspondingly, ClusPro v2.0 protein-protein docking (*24, 25*) gave a number of plausible structures for the T4moFC complex. One of these plausible structures is shown in Figure 5. Relative to the image of T4moF shown in Figure 1, the orientation of the docked complex has been rotated around the *z*-axis by 180˚. Thus T4moC makes its key interactions by approach from the opposite side of T4moF than does NADH. Also notable, the docked structures have the [2Fe-2S] clusters in T4moF and T4moC separated by ~12 Å, within a predicted favorable distance for biological electron transfer (*44*).

Figure 5B shows a close-up of the predicted interface between T4moF and T4moC. The energy-minimized interface has several intriguing features. The Rieske [2Fe-2S] ligands His47 and His67, which are solvent-exposed in T4moC alone and make essential inter-protein contacts in the T4moHC complex (*10*), are also potentially involved in the T4moFC complex through interactions with T4moF Glu93 and Glu35, respectively. Moreover, the docking model predicts that T4moC Arg65 can have a hydrogen bonding interaction with Cys41, which is a ligand to the [2Fe-2S] cluster in T4moF, and with T4moF Gly62-O. The importance of these interactions is supported by earlier mutagenesis studies showing that T4moC Arg65Ala had ~7-fold drop in k_{cat}/K_M for reduction by T4moF, which was primarily driven by an increase in K_M from ~2 μM to

~20 μM (*45*). It is also possible that Arg324 from T4moF, which shows two conformations in the T4moF electron density, can make electrostatic interactions with T4moC Ala66-O.

The electrostatic contributions to the interface between T4moF and T4moC predicted by the docking are consistent with the disrupting influence that increasing ionic strength had on the rate of inter-protein electron transfer (*45*). Interestingly, the T4moC Arg65Ala mutation had no effect on the ability of chemically reduced T4moC to transfer electrons to T4moH (*45*). This is consistent with the structure of the T4moHC complex, which shows that T4moC Arg65 is not included in the electron transfer interface between T4moC and T4moH, but instead projects into solvent (*10*). However, formation of a trinary T4moHCF complex is not plausible, given the involvement of T4moC Rieske [2Fe-2S] ligands His47 and His67 in the formation of complexes with both T4moF and T4moH.

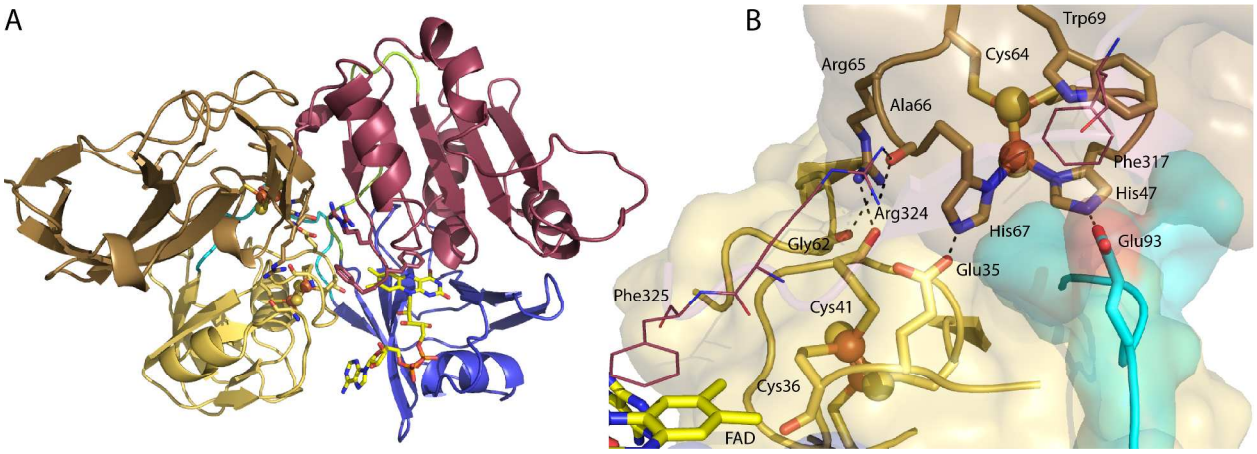

Figure 5. Predicted protein-protein interaction between T4moF and T4moC. T4moF has a prominent cleft between the Fd (*gold*) and NADH binding (*maroon*) domains that appears to be well-shaped to interact with T4moC (*brown*). A, energy-minimized T4moFC complex. The [2Fe-2S] clusters (spheres) are ~12 Å apart. B, close up of the interface showing predicted key

interactions, including T4moC Arg65 (*brown* sticks) interacting with T4moF Cys41 and Gly62; T4moC His67 interacting with T4moF Glu35; and T4moC His47 interacting with T4moF Glu93.

Summary of T4moF catalysis. By combining the predicted position for NADH binding obtained from this work with prior structural and spectroscopic results on PDR (*12, 40, 46*) and MmoC (*29, 30, 47, 48*), all three oxidoreductases will undoubtedly have a similar reductive half reaction with NADH. A minor structural arrangement, gated by predicted inter-domain motions, should lead to repositioning of Phe325 and allow NADH to bind the FNR-like domain (*40*) with the nicotinamide ring placed into a productive position in the active site. The results of energy-minimized docking of NADH also support hydride transfer to N5 of the isoalloxazine ring. Release of NAD^+, which has been shown to be a gating step for electron transfer in both PDR and MmoC, also seems likely to be gated by inter-domain motion. In this regard, the binding affinity for NAD^+ is ~2 orders of magnitude lower than NADH (*46*), and the crystal structure of PDR bound to NAD^+ showed electron density for all atoms of NAD^+ except the nicotinamide moiety. As a result, inter-domain motions that help Phe325 resume its resting position relative to FAD may weaken interactions with the nicotinamide ring, thus contributing to the release of NAD^+.

Electron-transfer between FAD and the [2Fe-2S] cluster may proceed through FAD C8M and Cys36-O as predicted by HARLEM, leading to formation of the reduced [2Fe-2S] cluster in T4moF and the semiquinone state of FAD. Subsequent electron transfer from the $[2Fe\text{-}2S]^{1+}$ cluster of T4moF to the Rieske-type [2Fe-2S] cluster in T4moC requires the formation of a new protein-protein interaction, and results presented here suggest that T4moF adopts an ideal configuration to achieve this. Interestingly, although electron transfer within T4moF likely involves C8M and Cys36-O, the most favorable interaction between T4moF and T4moC will

likely involve contacts of T4moF Glu35 and Cys41 with several residues on T4moC including Arg65, Ala66 and Rieske [2Fe-2S] ligand His67 (Figure 5B).

CONCLUSIONS

This work on T4moF provides a new structure of an intact BMM oxidoreductase from the iron-sulfur flavoprotein family. There is high structural homology within this family (*11*), albeit with little sequence identity and with cyclic permutation of the Fd and FNR-like domains. Oxidoreductases are generally thought to have optimized interactions with their cognate protein partners arising from specific electrostatic and steric interactions of residues that make up the respective electron transfer surfaces. This work also provides an example of how distal protein-protein interactions help to define the electron transfer complex. While electron transfer is not rate-limiting for T4MO catalysis (*45, 46*), this work suggests several conformational rearrangements that must occur in order for T4moF to be reduced by NADH. The structure of T4moF also identified a unique domain arrangement that supports formation of a highly favorable complex with its specific electron transfer partner T4moC. The predicted protein-protein interface is consistent with earlier mutagenesis results on electron transfer between T4moF and T4moC (*45*), lending support to the validity of the current structure-enabled analysis.

AUTHOR INFORMATION

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Financial Conflicts of Interest

None.

Funding Information

This work was funded by MCB-0843239 from the National Science Foundation to B.G.F. The use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science contract No. W-31-109-ENG-38. The use of the Life Science Collaborative Access Team (LS-CAT) was supported by the College of Agricultural and Life Sciences, Department of Biochemistry, and Office of the Vice Chancellor for Research and Graduate Education of the University of Wisconsin. J.F.A. was supported by a Wisconsin Distinguished Graduate Fellowship.

ACKNOWLEDGMENT

We would like to thank the Dane County School Consortium's Youth Apprenticeship Program for support provided to H.M.

Supporting Information

Hydrogen bonding distances for redox cofactors (Tables S1 and S2), RMSD values for alignment of iron-sulfur flavoprotein structures (Table S3), a sequence alignment (Figure S1), phylogenetic tree (Figure S2) and T4moF expression vector map (Figure S3).

REFERENCES

1. Lipscomb, J. D. (1994) Biochemistry of the soluble methane monooxygenase, *Annu Rev Microbiol 48*, 371-399.
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. 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.
4. Pikus, J. D., Studts, J. M., Achim, C., Kauffmann, K. E., Munck, 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.
5. Whited, G. M., and Gibson, D. T. (1991) Toluene-4-monooxygenase, a three-component enzyme system that catalyzes the oxidation of toluene to p-cresol in *pseudomonas mendocina* kr1, *J Bacteriol 173*, 3010-3016.
6. 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.
7. Moe, L. A., Bingman, C. A., Wesenberg, G. E., Phillips, G. N., Jr., and Fox, B. G. (2006) Structure of t4moc, the rieske-type ferredoxin component of toluene 4-monooxygenase, *Acta Crystallogr D Biol Crystallogr 62*, 476-482.
8. Lountos, G. T., Mitchell, K. H., Studts, J. M., Fox, B. G., and Orville, A. M. (2005) Crystal structures and functional studies of t4mod, the toluene 4-monooxygenase catalytic effector protein, *Biochemistry 44*, 7131-7142.
9. Bailey, L. J., McCoy, J. G., Phillips, G. N., Jr., and Fox, B. G. (2008) Structural consequences of effector protein complex formation in a diiron hydroxylase, *Proc Natl Acad Sci U S A 105*, 19194-19198.
10. Acheson, J. F., Bailey, L. J., Elsen, N. L., and Fox, B. G. (2014) Structural basis for biomolecular recognition in overlapping binding sites in a diiron enzyme system, *Nat Commun 5*, 5009.
11. Karlsson, A., Beharry, Z. M., Matthew Eby, D., Coulter, E. D., Neidle, E. L., Kurtz, D. M., Jr., Eklund, H., and Ramaswamy, S. (2002) X-ray crystal structure of benzoate 1,2-dioxygenase reductase from *acinetobacter* sp. Strain adp1, *J Mol Biol 318*, 261-272.
12. Correll, C. C., Batie, C. J., Ballou, D. P., and Ludwig, M. L. (1992) Phthalate dioxygenase reductase: A modular structure for electron transfer from pyridine nucleotides to [2fe-2s], *Science 258*, 1604-1610.
13. Bailey, L. J., Elsen, N. L., Pierce, B. S., and Fox, B. G. (2008) Soluble expression and purification of the oxidoreductase component of toluene 4-monooxygenase, *Protein Expr Purif 57*, 9-16.
14. Blommel, P. G., Martin, P. A., Seder, K. D., Wrobel, R. L., and Fox, B. G. (2009) Flexi vector cloning, *Methods Mol Biol 498*, 55-73.
15. Fox, B. G., and Blommel, P. G. (2009) Autoinduction of protein expression, *Curr Protoc Protein Sci Chapter 5*, Unit 5 23.
16. 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.
17. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode, *Macromolecular Crystallography, Pt A 276*, 307-326.
18. Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The swiss-model workspace: A web-based environment for protein structure homology modelling, *Bioinformatics 22*, 195-201.

19. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) Phenix: A comprehensive python-based system for macromolecular structure solution, *Acta Crystallogr D Biol Crystallogr 66*, 213-221.
20. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics, *Acta Crystallogr D Biol Crystallogr 60*, 2126-2132.
21. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) Molprobity: All-atom structure validation for macromolecular crystallography, *Acta Crystallogr D Biol Crystallogr 66*, 12-21.
22. DeLano, W. L. (2002) *The pymol molecular graphics system*, DeLano Scientific, San Carlos, CA, USA.
23. Deng, Z., Aliverti, A., Zanetti, G., Arakaki, A. K., Ottado, J., Orellano, E. G., Calcaterra, N. B., Ceccarelli, E. A., Carrillo, N., and Karplus, P. A. (1999) A productive nadp+ binding mode of ferredoxin-nadp+ reductase revealed by protein engineering and crystallographic studies, *Nat Struct Biol 6*, 847-853.
24. Comeau, S. R., Gatchell, D. W., Vajda, S., and Camacho, C. J. (2004) Cluspro: A fully automated algorithm for protein-protein docking, *Nucleic Acids Res 32*, W96-99.
25. Comeau, S. R., Gatchell, D. W., Vajda, S., and Camacho, C. J. (2004) Cluspro: An automated docking and discrimination method for the prediction of protein complexes, *Bioinformatics 20*, 45-50.
26. Yen, K. M., and Karl, M. R. (1992) Identification of a new gene, tmof, in the pseudomonas mendocina kr1 gene cluster encoding toluene-4-monooxygenase, *J Bacteriol 174*, 7253-7261.
27. Longenecker, K. L., Garrard, S. M., Sheffield, P. J., and Derewenda, Z. S. (2001) Protein crystallization by rational mutagenesis of surface residues: Lys to ala mutations promote crystallization of rhogdi, *Acta Crystallogr D Biol Crystallogr 57*, 679-688.
28. Karplus, P. A., and Bruns, C. M. (1994) Structure-function relations for ferredoxin reductase, *J Bioenerg Biomembr 26*, 89-99.
29. Muller, 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.
30. Chatwood, L. L., Muller, 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.
31. Gille, C., and Frommel, C. (2001) Strap: Editor for structural alignments of proteins, *Bioinformatics 17*, 377-378.
32. Bruns, C. M., and Karplus, P. A. (1995) Refined crystal structure of spinach ferredoxin reductase at 1.7 a resolution: Oxidized, reduced and 2'-phospho-5'-amp bound states, *J Mol Biol 247*, 125-145.
33. Kurisu, G., Kusunoki, M., Katoh, E., Yamazaki, T., Teshima, K., Onda, Y., Kimata-Ariga, Y., and Hase, T. (2001) Structure of the electron transfer complex between ferredoxin and ferredoxin-nadp(+) reductase, *Nat Struct Biol 8*, 117-121.

34. Morales, R., Kachalova, G., Vellieux, F., Charon, M. H., and Frey, M. (2000) Crystallographic studies of the interaction between the ferredoxin-nadp+ reductase and ferredoxin from the cyanobacterium *anabaena*: Looking for the elusive ferredoxin molecule, *Acta Crystallogr D Biol Crystallogr 56*, 1408-1412.
35. Nishida, H., Inaka, K., Yamanaka, M., Kaida, S., Kobayashi, K., and Miki, K. (1995) Crystal structure of nadh-cytochrome b5 reductase from pig liver at 2.4 a resolution, *Biochemistry 34*, 2763-2767.
36. Rypniewski, W. R., Breiter, D. R., Benning, M. M., Wesenberg, G., Oh, B. H., Markley, J. L., Rayment, I., and Holden, H. M. (1991) Crystallization and structure determination to 2.5-a resolution of the oxidized [2fe-2s] ferredoxin isolated from *anabaena* 7120, *Biochemistry 30*, 4126-4131.
37. Sevrioukova, I. F. (2005) Redox-dependent structural reorganization in putidaredoxin, a vertebrate-type [2fe-2s] ferredoxin from *pseudomonas putida*, *J Mol Biol 347*, 607-621.
38. Kurnikov, I. V. (2000) Harlem molecular modelling package, Department of Chemistry, University of Pittsburgh.
39. Ingelman, M., Bianchi, V., and Eklund, H. (1997) The three-dimensional structure of flavodoxin reductase from *escherichia coli* at 1.7 a resolution, *J Mol Biol 268*, 147-157.
40. Correll, C. C., Ludwig, M. L., Bruns, C. M., and Karplus, P. A. (1993) Structural prototypes for an extended family of flavoprotein reductases: Comparison of phthalate dioxygenase reductase with ferredoxin reductase and ferredoxin, *Protein Sci 2*, 2112-2133.
41. Emekli, U., Schneidman-Duhovny, D., Wolfson, H. J., Nussinov, R., and Haliloglu, T. (2008) Hingeprot: Automated prediction of hinges in protein structures, *Proteins 70*, 1219-1227.
42. Durand, P., Trinquier, G., and Sanejouand, Y.-H. (1994) A new approach for determining low-frequency normal modes in macromolecules, *Biopolymers 34*, 759-771.
43. 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.
44. Gray, H. B., and Winkler, J. R. (1996) Electron transfer in proteins, *Annu Rev Biochem 65*, 537-561.
45. Elsen, N. L., Moe, L. A., McMartin, L. A., and Fox, B. G. (2007) Redox and functional analysis of the rieske ferredoxin component of the toluene 4-monooxygenase, *Biochemistry 46*, 976-986.
46. Gassner, G., Wang, L., Batie, C., and Ballou, D. P. (1994) Reaction of phthalate dioxygenase reductase with nadh and nad: Kinetic and spectral characterization of intermediates, *Biochemistry 33*, 12184-12193.
47. Green, J., and Dalton, H. (1989) A stopped-flow kinetic study of soluble methane mono-oxygenase from *methylococcus capsulatus* (bath), *Biochem J 259*, 167-172.
48. Blazyk, J. L., and Lippard, S. J. (2002) Expression and characterization of ferredoxin and flavin adenine dinucleotide binding domains of the reductase component of soluble methane monooxygenase from *methylococcus capsulatus* (bath), *Biochemistry 41*, 15780-15794.

For Table of Contents Use Only

Structure of T4moF, the toluene 4-monooxygenase ferredoxin oxidoreductase

Justin F. Acheson, Hannah Moseson & Brian G. Fox