The four-component toluene/o-xylene monooxygenase (ToMO) from Pseudomonas stutzeri OX1 is capable of oxidizing arenes, alkenes, and haloalkanes at a carboxylate-bridged diiron center similar to that of soluble methane monooxygenase (sMMO). The remarkable variety of substrates accommodated by ToMO invites applications ranging from bioremediation to the regio- and enantiospecific oxidation of hydrocarbons on an industrial scale. We report here the crystal structures of the ToMO hydroxylase (ToMOH), azido ToMOH, and ToMOH containing the product analogue 4-bromo-phenol to 2.3 Å or greater resolution. The catalytic diiron(III) core resembles that of the sMMO hydroxylase, but aspects of the αββγγ tertiary structure are notably different. Of particular interest is a 6–10 Å-wide channel of ~35 Å in length extending from the active site to the protein surface. The presence of three bromophenol molecules in this space confirms this route as a pathway for substrate entrance and product egress. An analysis of the ToMOH active site cavity offers insights into the different substrate specificities of multicomponent monooxygenases and explains the behavior of mutant forms of homologous enzymes described in the literature.

Bacterial multicomponent monooxygenases (BMMs) comprise a family of carboxylate-bridged non-heme diiron enzymes capable of oxidizing a broad range of hydrocarbons including C2-C8 alkanes, alkenes, and aromatics (1, 2). Four characterized subclasses of multicomponent monooxygenases have been defined (2, 3). These are soluble methane monooxygenases (sMMOs), four-component alkene/arene monooxygenases or toluene monooxygenases (TMOs), three-component phenol monooxygenases (PHs), and αβ alkene monooxygenases (AMOs), of which all are believed to have evolved from a common ancestor. Bacteria containing multicomponent monooxygenases are capable of using specific hydrocarbon substrates as their primary source of carbon and energy (1, 2, 4). The remarkable range of substrate specificity exhibited by these enzymes endows these bacteria with the ability to bioremediate environmentally harmful substances such as trichloroethylene and petroleum spills (5, 6) and to regulate the global carbon cycle (4). BMMs can also perform regio- and stereospecific hydroxylations, making them useful for producing pure feedstocks for industrial synthesis (7). These enzyme systems, although highly homologous, have evolved different substrate specificities. Only soluble methane monooxygenase can activate the inert C–H bond of methane, which is one of the most difficult reactions to perform in nature (1), whereas the catalytic abilities of TMOs are limited to aromatics, alkenes, and some haloalkanes (2, 5).

Substrate hydroxylation in BMMs occurs at a dioxygen-activated, carboxylate-bridged diiron center in the α-subunit of a ~220–250 kDa hydroxylase component that is an (αβγγ) heterodimer or, in the case of one known AMO, an αβ monomer (1–3, 8, 9). Sequence identity comparisons and spectroscopic studies suggest that the diiron centers of the hydroxylase proteins from the different systems are structurally similar if not identical (1, 10–13). Catalysis at this important dimetallic center proceeds more efficiently in the presence of a small 10–16 kDa effector protein that alters the spectroscopic and redox properties of the diiron center and changes the regio-specificity of the reaction (1, 14). Electrons are supplied by NADH via a reductase comprising either one or two protein components (1, 2). The sMMO PH, and AMO families utilize a 38–40 kDa reductase consisting of an N-terminal ferredoxin domain and a C-terminal FAD domain. The four-component alkene/aromatic monooxygenase family requires a Rieske protein in addition to a reductase to facilitate electron transfer to the diiron center.

The methane monooxygenase hydroxylase (MMOH) is thus far the only member of the BMM family to have been structurally characterized (8, 15). It is presently unknown and of great interest to determine how the hydroxylase components from isopropyl-d-thiogalactoside; MES, 4-morphilinoethanesulfonic acid; MOPS, 3-N-morphilinopropanesulfonic acid; PDB, Protein Data Bank.
the different enzyme systems tune the diiron active site and organize the substrate binding pocket to select for specific hydrocarbons. Understanding the factors responsible for controlling the regio- and stereospecificity of substrate hydroxylation is important for they will prove valuable in developing strategies for the bioremediation of xenobiotics and the synthesis of targeted compounds. Moreover, since fully active TMOs and PHs can be expressed recombinantly and purified from *Escherichia coli*, whereas sMMO cannot, these enzyme systems are more useful for environmental and synthetic applications, and they offer advantages for mechanistic investigations of the hydroxylation/oxidation chemistry (16–19).

Knowledge of the structure of the hydroxylase component from one or both of these systems has thus been highly desired for some time. In the present article we describe our work on the toluene-xylene monooxygenase hydroxylase (ToMOH) from *Pseudomonas stutzeri* OX1. Toluene monooxygenases, which include monooxygenases from both the TMO and PH subclasses, have evolved to hydroxylate toluene specifically at the *ortho*, *meta*, and *para* positions, with high regiospecificity, and to epoxidize alkenes with high stereospecificity. We have crystallized ToMOH and *P. stutzeri* OX1, have relaxed toluene specificity (21), yet produce highly regiospecific products from alternate substrates like *o*-xylene, dimethylphenol, and *m*-cresol.2 The structure of a toluene monooxygenase hydroxylase offers the opportunity to understand how the enzyme tunes the reactivity of a carboxylate-bridged diiron center and adjusts the active site pocket to control product regiospecificity among the different TMOs and PHs. Here we report the structures of ToMOH in its oxidized, diiron(III) form as well as with bound 4-bromophenol and azide.

**EXPERIMENTAL PROCEDURES**

**Expression and Purification of ToMOH—**ToMOH was expressed from the plasmid pET22BEA in BL21(DE3) *E. coli* cells (17). Cells were grown at 37 °C in 10 2.1-liter baffled flasks containing 1 liter of LB media. At an OD_{600} of 0.8 the temperature was lowered to 25 °C, and ToMOH expression was induced by adding IPTG to a final concentration of 25 μM. At t = 0, 1, and 2 h after induction, 1 ml of a 100 mM Fe(NH₄)₂(SO₄)₂.6H₂O solution was added. After 4 h, the cells were pelleted by centrifugation at 6,000 rpm for 5 min, frozen in liquid nitrogen, and stored at −80 °C for further use.

Selenomethionine-derivatized ToMOH was created by growing cells in LeMaster’s media substituted with 90 mg/liter selenomethionine (22). Escherichia coli/pET22BEA ToMOH was expressed and stored as described above.

Purification of ToMOH was carried out essentially as described by Cafaro et al. (17). Thawed cells were sonicated in an ice slurry using six 1-min pulses with 1-min intervals between pulses to avoid excessive heating of the protein. The cracking solution contained buffer A (25 mM MOPS, pH 7.0, 50 mM NaCl, 10% glycerol, 2 mM cysteine, 200 μM Fe(NH₄)₂(SO₄)₂, 6H₂O, 8 mM thiglycolate) plus Pefabloc, phenylmethylsulfonyl fluoride, 20 units of DNase, and 5 mM MgSO₄. Debris was removed from the supernatant by centrifugation at 35,000 rpm for 45 min. The supernatant was decanted, filtered through a 0.22-μm membrane, and then loaded onto a DEAE fast flow ion exchange column (26 mm × 40 cm) equilibrated in buffer A. The protein was eluted using a 50–400 mM NaCl gradient. Fractions containing ToMOH were pooled, concentrated, and loaded onto an S300 size exclusion column (26 mm × 100 cm) preconditioned in 25 mM MOPS, pH 7.0, 100 mM NaCl, 10% glycerol. ToMOH fractions were pooled and loaded onto a phenyl-Sepharose column (26 mm × 40 cm) equilibrated in 1.3 mM NaCl, 25 mM MOPS, pH 7.0, and 10% glycerol. Before loading, NaCl was added to the ToMOH fractions from the S300 column to bring the salt concentration to ~1.3 M. The protein was eluted from the column by using a salt gradient from 1.3 to 0.1 mM NaCl. The purest fractions were identified by SDS-PAGE, pooled, and used for crystallization and activity assays. The protein was stored at −80 °C. The average yield was 15 mg of pure protein per liter.

**Purification of ToMOD—**ToMOD was expressed from the plasmid pET22D(∗)todD in BL21(DE3) gold *E. coli* cells (24). Cells were grown in eight 1-liter baffled flasks in LB media at 37 °C until an OD_{600} of ~0.7 and then induced with 100 μM IPTG at 30 °C. After 3–4 h of expression, the cells were harvested by centrifugation (5 min at 6,000 rpm) and frozen at −80 °C. Purification of ToMOD was carried out as described by Scognamiglio et al. (24).

**Activity and Iron Assays—**Steady state activity assays were performed according to previously published procedures (17), and the iron content of the protein was assessed by the ferrozine assay (25). The activity and iron content were 1.156 ± 127 nmol/min/mg (6.1 ± 0.3 s⁻¹) and 4.1 ± 0.1 Fe/dimer, respectively.

**Crystallization and Data Collection—**ToMOH was exchanged into buffer containing 10 mM MES, pH 7.1, and 10% glycerol and then complexed with ToMOD to a 50 μM ToMOH and 100 μM ToMOD protein solution for crystallization. This solution was then mixed with an equal volume (2–4 μl) of a precipitant solution consisting of 100 mM HEPES, pH 7.5, 2.1–2.5 μl (NH₄)₂SO₄, and 2–4% polyethylene glycol 400. Crystallization was achieved at 20 °C by using the hanging drop vapor diffusion method. Crystals of ToMOH alone grew within 0.5–2 days. Well diffracting crystals of ToMOH could be grown under these conditions in the absence of ToMOD. Both the recombinant ToMOH and the Se-Met derivative crystallized under the same conditions. The Se-Met protein, however, diffracted better and was used for most of the studies. The cryogenic solution contained the precipitant solution plus 20–25% glycerol. The 4-bromophenol derivative was obtained by soaking crystals in 50 mM 4-bromophenol dissolved in crystallization buffer for 15 min. X-ray diffraction data were collected at ALS on BL 8.2.2. The HKL suite of programs was used to integrate and scale the data (26).

**Structure Determination and Model Refinement—**The heavy atom positions and experimental phases were determined in CNS (27) by SAD phasing on the selenium peak data using iterative cycles of manual peak picking from anomalous difference Fourier maps (Table I). Solvent flattening using CNS improved the electron density map quality. The initial model was traced by Arp/wARP 6.0 (28). All refinement and manual rebuilding were performed with CNS and Xtalview (29). Rigid body refinement was carried out with CNS to generate initial models for the azide- and bromophenol-bound forms of ToMOH. The models were refined as described above.

**RESULTS AND DISCUSSION**

**Global Fold and Topology of ToMOH—**ToMOH, crystallized from a solution containing the coupling protein (ToMOD), has a topology similar to that of MMOH (Fig. 1, a and b) with a crystallographically required C3 symmetry axis separating the αβγ protomers. The calculated r.m.s. deviations between the ToMOH and MMOH α and β-subunits are 1.63 and 1.48 Å, respectively. The α-subunit fold differs most significantly at regions responsible for mediating interactions with the β- and γ-subunits (Fig. 1c). The α-helices that comprise most of the α-subunits of ToMOH and MMOH, especially those that house the diiron center and form the active site cavity, have almost identical structural topology. Similarly, the helical regions of the ToMOH and MMOH β-subunits are almost identical except for loops at the αβ interface (Fig. 1, c and d). As in MMOH, interactions between the αβγ protomers result in the formation of a canyon at the dimer interface. The canyon in ToMOH is slightly larger than that in MMOH, owing to a broader angle of interaction between the subunits at the αβ interface within the individual protomers. A consequence of this broad angle is a larger aperture across the C2 axis in the center of the canyon (Fig. 1a). The ToMOH dimer is stabilized mainly by interactions between the A helices of each α-subunit and a pair of helix-helix contacts between the β-subunits at the base of the protein, whereas MMOH utilizes four pairs of overlapping helices to stabilize the β-β subunit interactions. In MMOH but not ToMOH, the N termini of the β-subunits extend across the α-subunits to form helix-helix interactions with each other at the northern end of the protein as depicted in Fig. 1b.

The fold and location of the γ-subunit are strikingly different from that of MMOH (Fig. 1, c and f). In ToMOH, the γ-subunit binding sites reside at the northeastern and northwestern corners of the protein and exclusively contact the α-subunits. The

---

1 V. Cafaro and A. Di Donato, unpublished results.
the Miller index (hkl)ase (T4MOH) to be a di(oxidized MMOH and the toluene 4-monooxygenase hydroxy-
dicted by sequence identity and spectroscopic comparisons to
using the EMBL-EBI SSM server reveals the
secondary structure mostly of
resemble the
this subunit, although it may either stabilize the protomer or
than 15% sequence similarity occurring for the methane, tolu-
component of the multicomponent monooxygenases, with less
(Fig. 3
identical to that of MMOH. The pseudooctahedral coordination
indicates that little or no interaction occurs between these
residues. Since Cys-151 of MMOH and Gln-141 of ToMOH also
align structurally with the radical-forming Tyr-122 residue of
the R2 subunit of ribonucleotide reductase (RNR-R2), it has
been speculated that the amino acid at this position in multi-
component monooxygenases may have an essential catalytic
function (31). Mutagenesis studies on T4MO, in which Gln-141
was replaced with either a Cys, Val, Ala, or Phe, yielded active
isozymes with unaltered regiospecificity toward toluene (Sup-
plementary Table I). This result suggested that the residue at
position 141 is not essential for catalysis in the TMO subclass
of enzymes (32–34). Replacing Gln-141 with Cys in T4MO did
not afford an enzyme capable of oxidizing methane (32).

The coordination sphere of Fe2 is formed by His-234, Glu-
197, Glu-231, and oxygen atoms from the bridging carboxylate,
thioglycolate, and hydroxide ligands already identified. Sur-
prisingly, His-234 coordinates to Fe2 using its
-alkoxodiiron(III) cen-
th measured diffraction intensity and
(see below). The terminal water li-
gand hydrogen bonds to the dangling oxygen atoms of both
Glu-104 and Glu-231, a feature that is observed in MMOH.
Unlike MMOH, Glu-104 participates in a more intricate hydra-
gen bonding network involving Gln-141 and His-96, of which
both are conserved among the members of the TMO subclass
(Fig. 3c and Table I). In MMOH, the amino acid analogous to
Gln-141 is Cys-151, of which the sulfur atom lies 4.5 Å from the
carboxylate oxygen atom of Glu-114 (Fig. 3b). This distance
indicates that little or no interaction occurs between these
residues. Since Cys-151 of MMOH and Gln-141 of ToMOH also

| Crystal growth conditions | Se-Met thioglycolate | Se-Met 4-bromophenol | Se-Met azide |
|---------------------------|----------------------|-----------------------|-------------|
| Data collection           |                       |                       |             |
| Beamline                  | ALS 8.2.2             | ALS 8.2.2              | ALS 8.2.2   |
| Wavelength (Å)            | 0.979                 | 1.127                 | 0.920       |
| Space group               | P3,21                 | P3,21                 | P3,21       |
| Unit cell dimensions (Å)  | 182.81 × 182.81 × 68.04 | 183.24 × 183.24 × 67.67 | 182.70 × 182.70 × 67.78 |
| Total reflections         | 267.124               | 369.342               | 354.564     |
| Unique reflections        | 69,535                | 66,417                | 58,500      |
| Completeness (%)         | 99.3 (93.6)           | 98.4 (86.4)           | 93.4 (91.7) |
| Rfree (%)                 | 22.5 (27.3)           | 27.3                  | 26.7        |
| Number of atoms           | 7,348                 | 7,365                 | 7,382       |
| Water                     | 250                   | 287                   | 418         |
| r.m.s. deviation bond angle (°) | 0.0061 | 0.0068 | 0.0065 |
| Average B-value (Å)      | 35.0                  | 53.8                  | 33.7        |

- Values in parentheses are for the highest resolution shell.
- \( R_{cryst} = \frac{\sum |I(hkli) - \langle I(hkli) \rangle|}{\sum I(hkli)} \), where \( I(hkli) \) is the measured diffraction intensity and \( \langle I(hkli) \rangle \) is the mean of the intensity for the Miller index \( (hkli) \).
- \( R_{free} = \frac{\sum |I(hkli)| - |F_{hkli}|}{\sum |F_{hkli}|} \).
- \( R_{free} = R_{cryst} \) for a test set of reflections (5% in each case).

### Crystal Structure of the Toluene Monooxygenase Hydroxylase

**Table I**

| Phasing | Method | Number of sites/asymmetric unit | Figure of merit | Refinement |
|---------|--------|---------------------------------|-----------------|------------|
|         | SAD    | 32                              | 0.65            |            |
| Phasing | Molecular replacement | Molecular replacement |                |            |

Ce-Met thioglycolate

Se-Met 4-bromophenol

Se-Met azide

See Table I for details.
interactions between Glu-134, Glu-197, and Glu-111 in ToMOH (Fig. 3a), whereas the amino group from the side chain of Gln-140 performs this role in MMOH (Fig. 3b) (8). The hydronium ion assignment preserves charge neutrality and is supported by the out-of-plane displacement of the O-atom from the H-bond acceptor O-atoms in Glu-111, Glu-134, and Glu-197. Since Glu-111 of ToMOH and Gln-140 of MMOH do not align sequentially with one another, the proteins seemed to have evolved different strategies to preserve these interactions. Both Glu-111 and Gln-140 are part of localized hydrogen bonding networks linking these residues to amino acids on helix F, which forms part of the canyon region proposed for binding the coupling protein (37). Disruption of these hydrogen bonds due to binding of the coupling protein may allow this regulatory component to alter the chemical reactivity of the diiron center. These newly identified local hydrogen bonding networks in both the MMOH and ToMOH active sites highlight features not previously thought to be significant to the function of the enzyme.

The bridging ligands in the ToMOH active site are best fit as a hydroxide ion at the position cis to the coordinating histidines and a thioglycolate ion at the trans position that is accessible to solvent from the substrate binding cavity. We tentatively assign the latter ligand as thioglycolate, a component in the purification buffer of the protein, based on the size and shape of its electron density (Fig. 2c). Some residual difference density remains in the vicinity of the sulfur atom even after thioglycolate is modeled, suggesting possible partial oxidation to the sulfinate or perhaps positional disorder. Other buffer components such as dithiothreitol, sulfate ion, glycerol, ethylene glycol, and water, do not fit the density as well as thioglycolate and leave even more residual difference density. Nonetheless,
it is clear that the species in this bridging position binds symmetrically to both iron atoms and must be a monoanion if the neutrality of the dimetallic center is to be preserved. In the structure of oxidized ToMOH determined from crystals grown in the presence of 1.5 mM NaN₃, but not ToMOD, the density at the bridging position is significantly altered (Fig. 4). The new density is nicely fit by a monodentate, asymmetrically bound azide anion, a more strongly coordinating ligand than the thioglycolate anion. Because the buffers used in the purification of ToMOH and ToMOD were very similar, the absence of the thioglycolate feature cannot be attributed to the ToMOD buffer components. The identity of the electron density as azide is supported by the observation of a charge-transfer band at 450 nm in the UV-visible spectrum of ToMOH following the addition of NaN₃ (Supplementary Fig. 1). A similar spectroscopic feature arises when azide is added to MMOH, stearoyl-acyl carrier protein (ACP)/H₉₀₀₄/H₉₀₁₄ desaturase, and ruberythrin (38–40). The geometry of the asymmetrically bound azide ligand in ToMOH is similar to that of azide bound in the crystal structure of a reduced RNR-R2 Y208F mutant (Fig. 4b) (41). Although further spectroscopic evidence is desired to support the assignment of azide binding to iron in ToMOH, the above findings are consistent with previous structural and spectroscopic work.

In the catalytic cycle of MMOH, the first spectroscopically observable intermediate has been assigned as a ω-(peroxo)diron(III) species (31). Although there is presently no evidence for such an assignment in ToMOH, we may surmise that it will exist. Protein and small molecule model compound crystal structures of the azido- and/or peroxo-bound forms of the hemerythrin diron(III) and hemocyanin dicopper(II) metal centers demonstrate that azide-bound structures are good mimics for peroxide binding at metalloenzyme active sites and are capable of predicting the metal site at which O₂ is activated (42). The azido ToMOH structure suggests that O₂ may initially attack the Fe₂ site to become activated. The observed
1,1-semibridging or bent end-on coordination of the bound azide ion suggests that a peroxide intermediate may adopt one of these two geometries before reacting with substrate or further rearranging to a di(μ-oxo)diiron(IV) Q-like intermediate. Density functional theory calculations detailing the reaction of O₂ with the diiron center in reduced MMOH predict initial attack at Fe₂ for dioxygen activation, consistent with the above argument (43).

Open Channel Access to the Diiron Active Site—Within the ToMOH α-subunit lies a large channel of ~30–35 Å in length and 6–10 Å in width that connects the diiron center to the surface of the protein (Fig. 5). Starting at the diiron center, the channel traverses the four-helix bundle between helices E and B and extends toward the surface through a space created by the interface of helices D, E, B, G, and H (Figs. 5b and 6c). At the surface of the protein, amino acid side chains create a fork in the channel, of which the arms lead to two adjacent openings (Fig. 5a and Supplementary Fig. 2). These openings are located at the northern end of the molecule at the ridge of the canyon next to the C-terminal and N-terminal loops of helices E and H, respectively. The channel diameter of 6–10 Å is large enough to accommodate aromatic substrates or products as they move toward or away from the active site, respectively. Depending upon the crystal investigated, 3–4 ordered water molecules line the channel surface, which comprises both hydrophobic and hydrophilic residues (Figs. 6a and 5b). Crystals soaked with the product analogue 4-bromophenol, chosen to facilitate identification of the hydroxyl group, contained 3 molecules in the channel (Fig. 6b and Supplementary Fig. 3). The 4-bromophenol molecules are positioned such that the hydroxyl moieties are directed toward the surface. Their orientations in the channel appear to be determined by a combination of both hydrogen bonding interactions between the bromine atom, protein side chains, and adjacent bromophenol phenyl rings. Although it is largely unknown whether the channel can direct the orientation of substrate molecules as they try to gain access to the active site pocket, or how such substrate steering might be accomplished, the positioning of the bromophenol molecules suggests that a mechanism of this

![Crystal Structure of the Toluene Monooxygenase Hydroxylase](http://www.jbc.org/Downloaded from)
FIG. 4. Active site structures of ToMOH and RNR-R2 Y208F with bound azide. a, active site of oxidized ToMOH with azide bound asymmetrically depicted as in Fig. 2a. The Fe2-N-N bond angle is 147° in ToMOH and the Fe-Fe distance is 3.0 Å. b, the active site of RNR-R2 Y208F, depicted as in Fig. 2a, with azide bound asymmetrically. The Fe2-N-N bond angle is 130°, and the iron atoms are 3.4 Å apart. c, stereoview of the azide ligand bound to ToMOH accompanied by a 2Fo - Fc sa-omit map contoured to 1.2 σ.

TABLE II
Comparison of active site pocket residues from four BMM families

| Family | Residues |
|--------|----------|
| ToMO   | His-96   Ile-100 Glu-103 Ala-107 Ala-110 Gln-141 Phe-176 Met-180 Leu-192 Phe-196 Thr-201 Phe-205 |
| T4MO   | His-96   Ile-100 Gly-103 Ala-107 Gly-110 Gln-141 Phe-176 Ile-180 Leu-192 Phe-196 Thr-201 Phe-205 |
| T3MO   | His-96   Ile-100 Gly-103 Ala-107 Ala-110 Gln-141 Phe-176 Phe-180 Leu-192 Phe-196 Thr-201 Phe-205 |
| PH4    | His-102 Thr-106 Val-109 Ala-113 Met-116 Asp-147 Phe-182 Val-202 Thr-207 Phe-211 |
| AlkM   | Phe-101 Val-105 Leu-108 Ala-112 Gly-115 Gln-146 Met-180 Ala-184 Val-196 Phe-200 Thr-205 Phe-209 |
| IsoM   | Phe-107 Val-111 Leu-114 Ala-117 Gly-120 Gln-151 Phe-181 Ala-185 Val-197 Phe-201 Thr-206 Phe-210 |
| PH3    | Val-103 Leu-107 Gly-113 Ala-117 Ala-120 Cys-151 Phe-188 Phe-192 Leu-204 Gly-208 Thr-213 Ile-217 |
| BMO    | Met-87   Leu-91 Ala-94 Ala-98 Gly-101 Gly-101 Leu-132 Gly-167 Leu-184 Ala-188 Thr-193 Ile-197 |

* ToMO, toluene-o-xylene monooxygenase, Pseudomonas stutzeri OX1.
* T4MO, toluene 4-monooxygenase, Pseudomonas mendocina KR1.
* T3MO, toluene 3-monooxygenase, Ralstonia pickettii PKO1.
* PH4, four-component phenol hydroxylase, Ralstonia eutropha JMP134.
* AlkM, alkene monooxygenase, Xanthobacter sp. Py2.
* IsoM, isoprene monooxygenase, Rhodococcus sp. AD45.
* PH3, three-component phenol hydroxylase, Pseudomonas sp. CF600.
* T2MO, toluene 2-monooxygenase, Burkholderia cepacia JS150.
* DMSH, dimethyl sulfide hydroxylase, Acinetobacter sp. 20B.
* MMO, soluble methane monooxygenase, composite of all known strains.
* BMO, butane monooxygenase, Pseudomonas butanovora.
* AMO, alkene monooxygenase, Gordonia rubripertinctus B-276.
kind exists. It is clear from these results that the channel is solvent accessible and large enough to accommodate ToMO substrates and products. We therefore propose this route as the major pathway for substrate entrance to and product egress from the diiron active site, perhaps through the separate openings. Whether or not O₂ reaches the active site principally by this or some other pathway is less obvious, since it is small enough to diffuse through the surface helices to reach the diiron center.

In the α-subunit of MMOH, protein side chains separating three major cavities prevent them from forming an analogous channel linking the diiron center with the surface of the protein (Fig. 6c) (44). Previous structural studies of MMOH revealed that substrate analogues including xenon, iodoethane, and dibromoethane (44) as well as products like bromoethanol⁢³ can occupy these cavities (Fig. 6f). Further analysis of these MMOH structures indicates that cavities 2 and 3 are linked in crystal form I MMOH and separated in crystal form II, demonstrating that residues within the core of the α-subunit can shift to gate substrate or product movement through the protein (Fig. 6, c and d). Comparison of the ToMOH channel with the MMOH cavities, with and without bound substrate analogues, reveals that the paths through the core of the α-subunit are very similar for the two proteins (Figs. 5b, 6c, 6e, 6f). The only difference between the pathways is that the final MMOH cavity extends between helices D and G, reaching the surface on the other side of helix H. We suggest that they trace a universal route for molecular access to the diiron sites of other carboxylate-bridged diiron proteins of this class.

In MMOH, access to the hydrophobic cavity at the diiron center is blocked by Leu-110 and Phe-188, a feature previously termed the "leucine gate" (Fig. 7b) (45). The small aperture between these residues suggested initially that only small substrates like methane would be capable of gaining access to the diiron center via this route. In the structures of two different MMOH crystal forms, Leu-110 adopts alternate rotomeric conformations, revealing how larger substrates, like aromatics and C₆-C₈ alkanes, might approach the active site. ToMOH has similar residues, Ile-100, Phe-176, and Phe-196, marking the entrance to the active site cavity, but the opening to the diiron center of this enzyme is much larger (Fig. 7a). As a result, the iron atoms are directly accessible to solvent and substrates from the channel. Among the different carboxylate-bridged diiron enzymes, O₂ and H₂O₂-derived intermediates have not yet been spectroscopically detected for TMOs, PHs, and ruberythrin, proteins in which the diiron center is accessible, or predicted to be accessible, to bulk solvent. Stearoyl-ACP Δ⁹ desaturase also has a large substrate channel similar to that of ToMOH that is very accessible to solvent. In the absence of the stearyl-acyl carrier protein (ACP), however, intermediates are not observed in Δ⁹ desaturase. We therefore suggest that the binding of the ACP allows for the accumulation of a spectroscopically observable μ-1,2-peroxodiiron(III) species (46, 47), presumably because binding of the ACP blocks buffer access to the Δ⁹ desaturase diiron center, which might quench this intermediate. Similarly, the high valent di(μ-oxo)diiron(IV) intermediate Q of MMOH and the Fe(III)Fe(IV) intermediate X of RNR-R2 are housed in relatively secluded hydrophobic environments. We further speculate that the failure to observe O₂- or H₂O₂-derived intermediates in TMO, PH, and ruberythrin may be due to reactions with buffer components that can more readily access and react with these species in their active sites.

Mutagenesis experiments on T4MOH and toluene 2-monoxygenase (T2MO) targeting Ile-100 and Val-106, respectively, provide strong evidence to support the channel as the primary pathway for substrate entrance (32–34). Changing Ile-100 to residues with larger side chains such as Arg and Trp inhibited toluene hydroxylation, presumably by blocking access to the active site cavity (Supplementary Table I). An I100Q mutant, in which Ile was replaced with a polar side chain of comparable size, was active, whereas the activity of an I100C mutation was compromised. The I100Q and I100C mutants indicate that the residue in this position need not be hydrophobic for TMOs to be efficient catalysts and suggest that a small side chain is not beneficial, perhaps due to limited gating ability and greater exposure to solvent. A Val to Ala mutation of the analogous gate residue in T2MO from Burkholderia cepacia G4, a member of the three-component PH family, resulted in enhanced ability to degrade naphthalene and three-ring fused aromatics including phenanthrene, fluorene, and anthracene (48). In the light of the present structures, these findings strongly suggest that members of the PH subclass have a similar substrate channel. Furthermore, as the work on T2MO demonstrates, the ability to engineer TMOs and PHs to act on bigger aromatic substrates is most likely a consequence of modulating the dimensions of the channel.

Residues Phe-176 and Phe-196 at the active site also appear to be essential for activity (32, 34). A F176A T4MO mutant could not oxidize toluene, TCE, or butadiene, whereas a F176L mutant was 100, 25.5, and 38.4% as active toward these substrates, respectively, as compared with the wild-type enzyme. The T4MO mutants F196Y and F196L were as active as the wild-type protein, whereas the F196I and F196G mutants were less so. These results imply that the bulky side chains of Phe-176 and Phe-196 (Fig. 7a) are important for controlling the activity of the enzyme, perhaps by limiting exposure of the iron

³ M. H. Sazinsky and S. J. Lippard, unpublished results.
centers to undesired solution components, or for guiding the bound substrate to the diiron center. Further work is required to clarify their roles.

**Substrate Specificity of Four-component Alkenel/Arene Monoxygenases**—For TMOs the key to regiospecific hydroxylation is the topology of the active site pocket as defined by the residues
that form its lining. The dimensions of the substrate channel may also be important. Although the coupling protein can affect the regiospecificity of the enzyme, it most likely functions in this manner by adding rigidity to the active site to enforce pre-programmed constraints in the enzyme, rather than by dictating specific regiochemistry (14).

The T4MO and T3MO enzymes, of which both require a glycine at residue 103 and a bulky hydrophobic amino acid at position 180, are highly para-directing enzymes whereas PHs, which utilize a leucine and alanine at these positions, respectively, are strongly ortho-directing (Table II and Supplementary Table I). Mutagenesis work on T4MO has been carried out to probe how these active site residues affect the specificity of toluene monooxygenases (14, 32, 33). Substituting Gly-103 in T4MO with Leu increased the yield of o-cresol formation from 0.9 to 55.4%, confirming the importance of this amino acid in modulating the position of toluene ring hydroxylation (14). The size of the residue at position 103 can also affect the enantioselectivity of alkene epoxidation; the same G103L mutation increased the yield of (S)-butadiene epoxide from 67 to 90% (32). ToMOH, which differs significantly from the T4MO active site only at positions 103 and 180, has relaxed regiospecificity, as manifest by a product distribution of 36, 19, and 45% for o-, m-, and p-cresol, respectively. Residue 103 in ToMOH is a Glu and points away from the active site such that its charged carboxylate group is buried (Figs. 7a and 8). Only the α and β carbons face the substrate binding pocket. Since T4MO has no β carbon atom occupying this position, we propose that the relaxed regiospecificity of ToMO for toluene oxidation is partially due to the β carbon atom from Glu at position 103 steering the toluene C2 position toward the diiron center, presumably through steric interactions with the methyl group of the substrate. The result of these interactions will force the phenyl ring to cant to the left or right.

It is unknown whether the amino acid differences at position 180, Met versus Ile, affect ToMOH regiospecificity. Sequence comparison of the TMO and PH active sites suggests that smaller residues at positions 180 and 192 of T4MO, in addition to a larger residue at position 103, may be essential for the ortho-hydroxylating regiospecificity of PHs (Table II). Having an Ala at position 180 and an Ile or Val at position 192 will create extra space in the active site cavity and possibly provide a niche for the toluene methyl group. As toluene approaches the active site through the cavity, the bulky residue at position 103 may steer the methyl group on the phenyl ring toward the Fe2 side of the active site pocket where residues 180 and 192 reside. The result would be to expose the toluene C2 position to the diiron center with subsequent hydroxylation at the ortho position.

The above analysis suggests that it should be possible to design rationally BMM active sites to generate enzyme systems for specific industrial and environmental applications. Since structures are now available for both ToMOH and MMOH, two proteins having almost identical dimetallic centers, we are poised to address the question of why ToMOH cannot activate the inert C-H bond of methane. The present structure determination should also assist in the design of mutant forms of the protein in which intermediates in the dioxygen activation steps can be spectroscopically characterized, as in MMOH, and in revealing the factors responsible for hydroxylation specificity.

**Fig. 7.** Entryways to the MMOH and ToMOH and MMOH active sites. a, space-filling representation of the ToMOH active site entrance as substrate would approach it through the channel. Access to the diiron center is formed by residues 100, 176, and 196. The diiron center is depicted as orange spheres, whereas the rest of the atoms are colored by type. The α-subunit helices are colored blue. b, space-filling representation of the MMOH active site pocket depicting access to the diiron center as substrates would approach the leucine gate (residues Phe-188 and Leu-110) after passage through the MMOH α-subunit cavities. Atoms and ribbons are represented as in a.

**Fig. 8.** Stereoview of the ToMOH active site pocket from the other side of the diiron center. The iron atoms are depicted as orange spheres. Amino acid side chains forming the substrate pocket and iron ligands are colored by atom and are presented as sticks. The α-subunit helices are shown in green.
Acknowledgments—We thank Valeria Cafaro (Università di Napoli Federico II) for supplying the ToMO plasmids and a sample of ToMOF and Elisabeth Cadieux (MIT) for ToMOC.

REFERENCES

1. Merku, M., Kopp, D. A., Sazinsky, M. S., Blazyk, J. L., Mihlig, J. M., and Lippard, S. J. (2001) Angew. Chem. Int. Ed. 40, 2782–2807
2. Leahy, J. G., Batchelor, P. J., and Morcom, S. M. (2003) FEMS Microb. Rev. 27, 449–479
3. Notomista, E., Lahm, A., Di Donato, A., and Tramontano, A. (2003) J. Mol. Evol. 56, 435–445
4. Han, X. S., and Hans, S. T. (1996) Microbiol. Rev. 60, 439–471
5. Chauhan, S., Barbieri, P., and Wood, T. K. (1998) Appl. Environ. Microbiol. 64, 3023–3024
6. Sullivan, J. P., Dickinson, D., and Chase, H. A. (1998) Crit. Rev. Microbiol. 24, 335–373
7. Archelas, A., and Furstoss, R. (1997) Annu. Rev. Microbiol. 51, 491–525
8. Rosenzweig, A. C., Frederick, C. A., Lippard, S. J., and Nordlund, P. (1993) Nature 366, 537–543
9. Muir, A., and Dalton, H. (1995) Biosci. Biotech. Biochem. 59, 853–859
10. Couf, D. E., Blazyk, J. L., Whittington, D. A., Wu, W., Rosenzweig, A. C., and Lippard, S. J. (2000) Eur. J. Biochem. 267, 2174–2185
11. Newman, L. M., and Wackett, L. P. (1995) Biochemistry 34, 14066–14076
12. Pisk, J. D., Studts, J. M., Achim, C., Kauffmann, K. E., Münck, E., Steffan, R. J., McClay, K., and Fox, B. G. (1996) Biochemistry 35, 9106–9119
13. Cadieux, E., Vrajmasu, V., Achim, C., Powlowski, J., and Münck, E. (2002) Biochemistry 41, 10689–10691
14. Mitchell, K. H., Studts, J. M., and Fox, B. G. (2002) Biochemistry 41, 3178–3186
15. Eland, N., Radzakiewicz, R., Froland, W. A., Wallar, B. J., Earhart, C. A., Lipscum, J. D., and Ohlendorf, D. H. (1997) Proc. Natl. Acad. Sci. 94, 565–566
16. Lloyd, J. S., Finch, R., Dalton, H., and Murrell, J. C. (1999) Microbiology 145, 461–471
17. Cafaro, V., Scognamiglio, R., Vignetti, M., and Carbone, E. F., Amoroso, A., Cashar, A., Pucci, P., and Di Donato, A. (2002) Eur. J. Biochem. 269, 5689–5699
18. Studts, J. M., Mitchell, K. H., Pisk, J. D., McClay, K., Münck, E., Steffan, R. J., and Fox, B. G. (2000) Protein Expr. Purif. 20, 58–65
19. Di Vito, L., Valetti, F., Cappello, P., Pessione, E., Cavaletto, M., Griva, E., Gribaudo, G., Gilardi, G., and Giunta, C. (2003) Eur. J. Biochem. 270, 2244–2253
20. McCaig, C., McCaig, H., and Steffan, R. J. (2000) Appl. Environ. Microbiol. 66, 1877–1882
21. Berton, G., Bolzoni, F., Galli, E., and Barbieri, P. (1996) Appl. Environ. Microbiol. 62, 3704–3711
22. Guo, N., and Poiesz, B. J. (1997) Electrophoresis 18, 2714–2723
23. LeMaster, D. M., and Richards, F. M. (1985) Biochemistry 24, 7263–7273
24. Scognamiglio, R., Notomista, E., Barbieri, P., Pucci, P., Dal Piaz, F., Tramontano, A., Di Donato, A. (2001) Proc. Natl. Acad. Sci. 98, 482–490
25. Gibbs, C. R. (1976) Annu. Rev. Biochem. 45, 2789–2807
26. Otwinski, Z., and Minor, W. (1997) in Methods in Enzymology (Sweet, R. M., ed) Vol. 276, pp. 302–326, Academic Press, New York
27. Brungger A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gras, P., Grasse-Kunsteufel, R. W., Jiang, J.-S., Kuszewski, J., Niles, P., Sane, U. S., Read, R. J., Ree, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D 54, 905–921
28. Lamzin, V. S., and Wilson, K. S. (1998) Acta Crystallogr. Sect. D 54, 129–147
29. McRee, D. E. (1999) J. Struct. Biol. 125, 156–165
30. Krienen, E., and Henrick, K. (2003) in Proceedings of the 5th International Conference on Structural Biology (Kungl, P. J., ed) p. 88, Vienna
31. Feig, A. M., and Lippard, S. J. (1994) Chem. Rev. 94, 785–805
32. Feig, M., and Lippard, S. J. (1995) J. Am. Chem. Soc. 117, 22346–22354
33. Sjo, E. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 2486–2490
34. Andersson, K. K., Sjo, E., and Nordlund, P. (1999) J. Am. Chem. Soc. 121, 2346–2352
35. Ai, J. Y., Broadwater, J. A., Lohr, T. M., Sanders-Lochr, J., and Fox, B. G. (1997) J. Biol. Inorg. Chem. 2, 37–45
36. Gherman, B. F., Baik, M.-H., Lippard, S. J., and Friesner, R. A. (2004) J. Am. Chem. Soc. 126, 2978–2990
37. Whittington, D. A., Rosenzweig, A. C., Frederick, C. A., and Lippard, S. J. (2001) Biochemistry 40, 3476–3482
38. Rosenzweig, A. C., Brandstetter, H., Whittington, D. A., Nordlund, P., Lippard, S. J., and Frederick, C. A. (1997) Proteins 29, 141–152
39. Broadwater, J. A., Achim, C., Münck, E., and Fox, B. G. (1999) Biochemistry 38, 12197–12204
40. Broadwater, J. A., Ai, J. Y., Lohr, T. M., Sanders-Lochr, J., and Fox, B. G. (1998) Biochemistry 37, 14664–14671
41. Canada, K. A., Iwashita, A., Shim, H., and Wood, T. K. (2002) J. Bacteriol. 184, 344–349
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

Matthew H. Sazinsky, Joel Bard, Alberto Di Donato and Stephen J. Lippard

J. Biol. Chem. 2004, 279:30600-30610.
doi: 10.1074/jbc.M400710200 originally published online April 19, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M400710200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2004/04/29/M400710200.DC1

This article cites 43 references, 6 of which can be accessed free at
http://www.jbc.org/content/279/29/30600.full.html#ref-list-1