



Structures of the alkanesulfonate monooxygenase MsuD provide insight into C–S bond cleavage, substrate scope, and an unexpected role for the tetramer

Received for publication, December 18, 2020, and in revised form, May 19, 2021. Published, Papers in Press, May 23, 2021.

<https://doi.org/10.1016/j.jbc.2021.100823>

Jeremy J. M. Liew¹, Israa M. El Saudi¹, Son V. Nguyen², Denyce K. Wicht², and Daniel P. Dowling^{1,*}

From the ¹Department of Chemistry, University of Massachusetts Boston, Boston, Massachusetts, USA; ²Department of Chemistry and Biochemistry, Suffolk University, Boston, Massachusetts, USA

Edited by F. Peter Guengerich

Bacterial two-component flavin-dependent monooxygenases cleave the stable C–S bond of environmental and anthropogenic organosulfur compounds. The monooxygenase MsuD converts methanesulfonate (MS^-) to sulfite, completing the sulfur assimilation process during sulfate starvation, but the mechanism of this conversion remains unclear. To explore the mechanism of C–S bond cleavage, we report a series of crystal structures of MsuD from *Pseudomonas fluorescens* in different liganded states. This report provides the first crystal structures of an alkanesulfonate monooxygenase with a bound flavin and alkanesulfonate, elucidating the roles of the active site lid, the protein C terminus, and an active site loop in flavin and/or alkanesulfonate binding. These structures position MS^- closest to the flavin N5 position, consistent with an N5-(hydro) peroxyflavin mechanism rather than a classical C4a-(hydro) peroxyflavin mechanism. A fully enclosed active site is observed in the ternary complex, mediated by interchain interaction of the C terminus at the tetramer interface. These structures identify an unexpected function of the protein C terminus in this protein family in stabilizing tetramer formation and the alkanesulfonate-binding site. Spurred by interest from the crystal structures, we conducted biochemical assays and molecular docking that redefine MsuD as a small- to medium-chain alkanesulfonate monooxygenase. Functional mutations verify the sulfonate-binding site and reveal the critical importance of the protein C terminus for monooxygenase function. These findings reveal a deeper understanding of MsuD's functionality at the molecular level and consequently how it operates within its role as part of the sulfur assimilation pathway.

Flavin-dependent monooxygenases catalyze the monooxygenation of a wide range of substrates using dioxygen and either a reduced flavin mononucleotide ($FMNH^-$) or reduced flavin adenine dinucleotide ($FADH^-$), making them incredibly useful biocatalysts (1–4). One family of substrates includes organosulfur compounds (5), which are of interest as the chemical properties of sulfur are crucial to many biological

processes. For example, sulfur's redox properties make the element key for electron carrier molecules such as glutathione (6) and iron-sulfur clusters (7). Biosynthetic pathways for incorporating sulfur use exogenous sulfate, which is reduced to sulfite and subsequently sulfide in the cell (Fig. 1A) (8). However, when deprived of typical sulfur sources, some bacteria produce sulfite from organosulfur compounds under the sulfur starvation response (9–11), using compounds including taurine (9), dibenzothiophene (12), long-chain alkanesulfonates (13), and C_1 -sulfur compounds (14) such as dimethylsulfide (DMS) (15), dimethyl sulfone ($DMSO_2$) (16), and methanesulfonate (MS^-) (17). $DMSO_2$ and MS^- originate from the oxidation of DMS (18) in either the top layer of the ocean (19) or in aerosols (20), and deposition by rainfall (5, 21) makes $DMSO_2$ and MS^- available for sulfur acquisition by soil-dwelling bacteria (Fig. 1A). In addition, the methyl groups can serve as a carbon source in methylotrophic bacteria through the serine or ribulose monophosphate pathways (22, 23). Enzymatic strategies for breaking the chemically stable C–S bond of organosulfur compounds include Rieske nonheme oxygenases (24, 25), α -ketoglutarate-dependent dioxygenases (9), and flavin-dependent monooxygenases (17, 26).

In pseudomonads, *msu* and *sfn* genes contain the flavin-dependent monooxygenases SfnG, MsuC, and MsuD that convert $DMSO_2$ to sulfite (Fig. 1B). SfnG converts $DMSO_2$ to methanesulfinate (MSI^-) (27), MsuC oxidizes MSI^- to methanesulfonate (MS^-) (28), and MsuD catalyzes the conversion of MS^- to sulfite (Fig. 1C) (17). Together SfnG and MsuD are responsible for sequential cleavage of the two C–S bonds of $DMSO_2$, and each methyl group is presumed to be oxidized to formaldehyde. The enzymes SfnG, MsuC, and MsuD are members of a small subset of flavin-dependent monooxygenases that are characterized by their use of reduced flavin as a cosubstrate rather than a cofactor. Termed two-component flavin-dependent monooxygenases, members of this family lack an NAD(P)H-binding site and therefore require a separate reduced NAD(P)H:oxidized flavin mononucleotide (FMN) oxidoreductase to provide the $FMNH^-$ cosubstrate. MsuE was identified within the *msu* operon of *Pseudomonas fluorescens* and *Pseudomonas aeruginosa* as the likely NADH:FMN oxidoreductase for MsuC (28) and MsuD

* For correspondence: Daniel P. Dowling, daniel.dowling@umb.edu.
Present address for Son V. Nguyen: Department of Chemistry, University of New Hampshire, Durham, New Hampshire 03824, USA.

Structural and mechanistic insights from ligand-bound MsuD

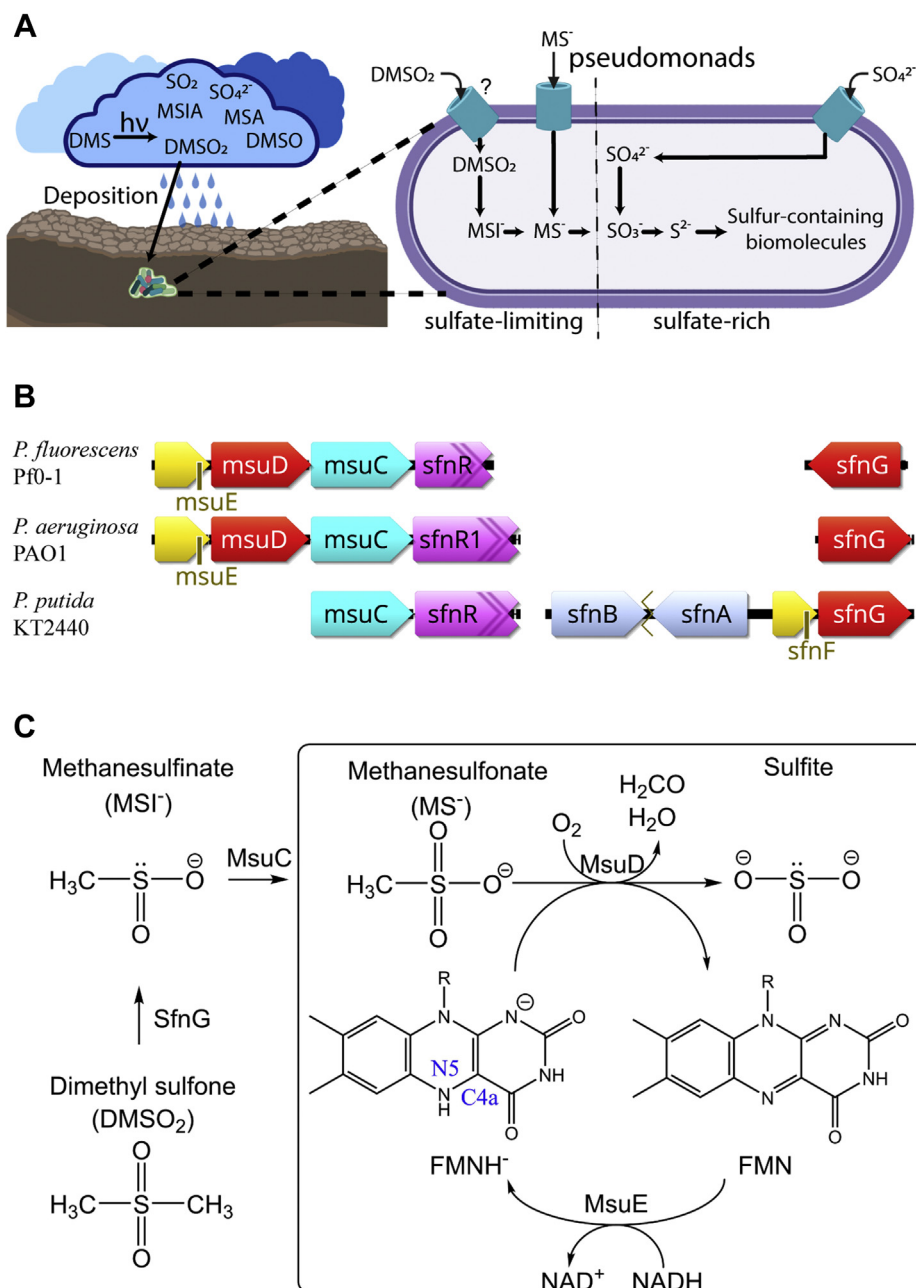


Figure 1. Bacterial sulfur assimilation of dimethyl sulfone and methanesulfonate. *A*, the cycling of atmospheric sulfur is depicted. Oxidation products of dimethylsulfide (DMS) include dimethyl sulfoxide (DMSO), dimethyl sulfone (DMSO₂), methanesulfinic acid (MSIA), methanesulfonic acid (MSA), sulfur dioxide (SO₂), and sulfate (SO₄²⁻). Pseudomonads can intracellularly convert DMSO₂ into methanesulfinic acid (MSi⁻), methanesulfonate (MS⁻), and, finally, sulfite. Transporters for DMSO₂ have yet to be characterized. Under sulfate-limiting conditions, sulfate is reduced to sulfite and subsequently sulfide (S²⁻) for biosynthetic reactions. Representation created with BioRender.com. *B*, representative gene clusters for utilization of DMSO₂ from *P. fluorescens* Pf0-1, *P. aeruginosa* PAO1, and *P. putida* KT2440 include three monooxygenases, *sfnG*, *msuC*, and *msuD*, and the oxidoreductases *msuE* or *sfnF*. Monooxygenases that break C–S bonds are in red, monooxygenases that form S–O bonds are in cyan, NADH:FMN oxidoreductases are in yellow, transcriptional regulators are in pink, and proteins of unknown function are in light blue. Initial studies of the *sfn* operon were reported for the genetically intractable *P. putida* DS1 (16, 29); therefore, *P. putida* KT2440 was used for alignment. The *sfn* operon in *P. putida* contains the uncharacterized *sfnAB* pair (79), which is separated by predicted ABC transporters and a hypothetical protein (represented as a jagged line); however, *sfnG* is a stand-alone gene within *P. aeruginosa* PAO1 (80) and *P. fluorescens* Pf0-1. The *msu* operon groups with *sfnR*, or *sfnR1* and *sfnR2* in *P. aeruginosa* (*sfnR2* omitted from alignment) (80). Gene information is in Table S1. *C*, the sulfur assimilation pathway of *P. fluorescens* converts DMSO₂ into sulfite. The NADH:FMN oxidoreductase MsuE is required to produce FMNH⁻ for SfnG, MsuC, and MsuD; however, MsuE is shown only for the MsuD reaction for clarity.

(17), and SfnF was identified in the *sfn* operon of *Pseudomonas putida* (29) (Fig. 1B). Structurally characterized two-component flavin-dependent monooxygenases that catalyze C–S bond cleavage include CmoJ, DmoA, BdsA, and SsuD,

which adopt a (β/α)₈ triose isomerase phosphate (TIM) barrel fold and are denoted as group C two-component flavin-dependent monooxygenases (1). This is separate from MsuC, a group D flavin-dependent monooxygenase that instead

exhibits the acyl-CoA dehydrogenase fold and catalyzes formation of a S–O bond. The group C monooxygenase DmoA functions on dimethylsulfide (30), CmoJ is involved in the cysteine salvage pathway (31), and BdsA is part of the 4S utilization pathway to convert dibenzothiophene to sulfite (32). However, MsuD is most closely related to *Escherichia coli* alkanesulfonate monooxygenase SsuD with 67% sequence identity. SsuD catalyzes C–S bond cleavage of C₂–C₁₀ alkyl-substituted sulfonates, in addition to alkyl chains with bulkier chemical groups and larger buffer molecules (26), whereas MsuD is reportedly most active with a C₁ sulfonate (17). To date, no known alkanesulfonate monooxygenase structure has been solved with flavin or an alkanesulfonate bound, and only the structure of *E. coli* SsuD without ligands has been reported. As such, the disorder of the SsuD active site has hindered our understanding of the vital alkanesulfonate monooxygenase reaction (33).

The overall monooxygenase reaction was originally thought to advance through a C4a-(hydro)peroxyflavin intermediate (3), but recent structural, computational, and biochemical studies of two monooxygenases, EncM from enterocin biosynthesis (34) and RutA from uracil degradation (35), support the use of either an N5-(hydro)peroxyflavin or flavin-N5-oxide. Although EncM is structurally similar to flavin oxidases/dehydrogenases (36) and RutA exhibits a TIM barrel fold (37), their structures identified that dioxygen and the substrate bind on opposite faces of the flavin, and a conserved dioxygen reactivity motif (F/L-T/S-NxV-A-F/L/Y, where x is a branched chain amino acid) was identified within RutA and a subset of group C monooxygenases, including CmoJ, DmoA, and DszA (catalyzes the same reaction as BdsA). Matthews *et al.* (35) investigated RutA *via* computational studies that support the thermodynamic possibility of inversion at the N5 position of the flavin, which would move the (hydro)peroxy group toward the substrate for catalysis to proceed. RutA and MsuD share only 25% sequence identity, and the dioxygen reactivity motif of RutA is not fully conserved in MsuD, SsuD, or SfnG; therefore, it is unclear if a similar N5-(hydro)peroxyflavin could be employed in enzymes that facilitate C–S bond cleavage. Observing the ordered active site would help elucidate this question. Likewise, a completely ordered model could help clarify regions of disorder that exist in RutA and other group C monooxygenases such as the active site lid and seemingly functionless C terminus.

Here we report the first crystal structures and biochemical activity studies of *P. fluorescens* MsuD, which adopts a group C (β/α)₈ TIM barrel fold. Structural experiments entailing soaks and cocrystallization of MsuD with FMN and MS[−] enabled us to obtain the first ever crystallographic snapshots of an alkanesulfonate monooxygenase in different complexed states with its substrates. Because the FMNH[−] cosubstrate is aerobically sensitive (38, 39), we have employed oxidized FMN, a product of the monooxygenase, to obtain crystal structures that detail the flavin-binding site. We report the structures of unliganded MsuD, MsuD in complex with FMN, and MsuD in complex with both FMN and MS[−]. The 2.4-Å resolution structure of MsuD with flavin and MS[−] has enabled us to revisit the

chemical mechanism to assess if C–S cleavage likely advances through the C4a or N5-(hydro)peroxyflavin intermediates, as well as put into context many of the biochemical experiments that have been difficult to rationalize without an experimentally determined binding mode for the flavin and alkanesulfonate moieties. Our studies further have led us to redetermine the substrate scope for MsuD. The resulting structures reveal an exquisite molecular connection between the flavin- and MS[−]-binding sites, the active site lid, and the protein C terminus, providing an unrealized function for the C-terminal extension that may be used in other group C flavin monooxygenases.

Results

Overall structure of MsuD from *P. fluorescens*

The 2.4-Å resolution crystal structure of MsuD in complex with FMN and MS[−] (ternary-MsuD) was obtained by cocrystallization with greater than 5-fold molar excess of both ligands. MsuD adopts a classic (β/α)₈ TIM barrel fold (Fig. 2A) (40), with four insertion segments (ISs) to the (β/α)₈ core of MsuD: IS-1 includes two β -strands (labeled $\beta_{1a/b}$); IS-2 includes an α -helix (labeled α_{4a}); IS-3 includes a β -hairpin (labeled $\beta_{4a/b}$); and IS-4 includes three α -helices (labeled $\alpha_{7a/b/c}$), a 3_{10} helix (labeled η_{7d}), and three short β -strands (labeled $\beta_{7a/b/c}$). In addition, MsuD contains an extended protein C terminus. All but the last four residues of the protein C terminus were able to be placed into electron density along with the ligands FMN and MS[−].

Ternary-MsuD crystallized in the space group *P*6₁ with four MsuD chains per asymmetric unit, arranged as a dimer-of-dimers (Fig. 2B). The dimer and tetramer interfaces bury approximately 1960 and 2940 Å² per molecule, respectively, determined by the PISA server (41). The buried surface area is consistent with size exclusion data that support MsuD is a tetramer (Fig. S1A). The dimer interface consists of portions of the (β/α)₈ core, IS-2, and IS-3 (Fig. 2C), where the $\beta_{4a/b}$ -hairpin packs against α_{4a} of the dimer molecule. The main tetramer interface is made up of components from the (β/α)₈ core, where α_1 and its preceding loop as well as α_8 combine with a symmetry-related molecule to generate a four-helix bundle (Fig. 2D). The ordering of the C terminus constitutes approximately 1650 Å² of the tetramer interface, making the tetramer interface the largest interface between monomeric subunits; however, as will be discussed below, the C terminus appears to be a mobile element that is related to substrate binding. It is intriguing that truncation of the last 16 residues from the C terminus affects migration of the protein by gel filtration (Fig. S1B), causing MsuD to behave as a dimer.

As was predicted, the closest reported structure to MsuD is *E. coli* SsuD (42), which aligns with an RMSD of 1.3 Å for 330 C α atoms (43). Related structures with less than 32% sequence identity are included in Table S2, and sequence alignment of MsuD with SsuD and a more distantly related homolog representative, RutA, is shown in Figure 3. Like MsuD, both SsuD and BdsA crystallize as homotetramers (Fig. S3), whereas the other structural homologs are reported homodimers. The

Structural and mechanistic insights from ligand-bound MsuD

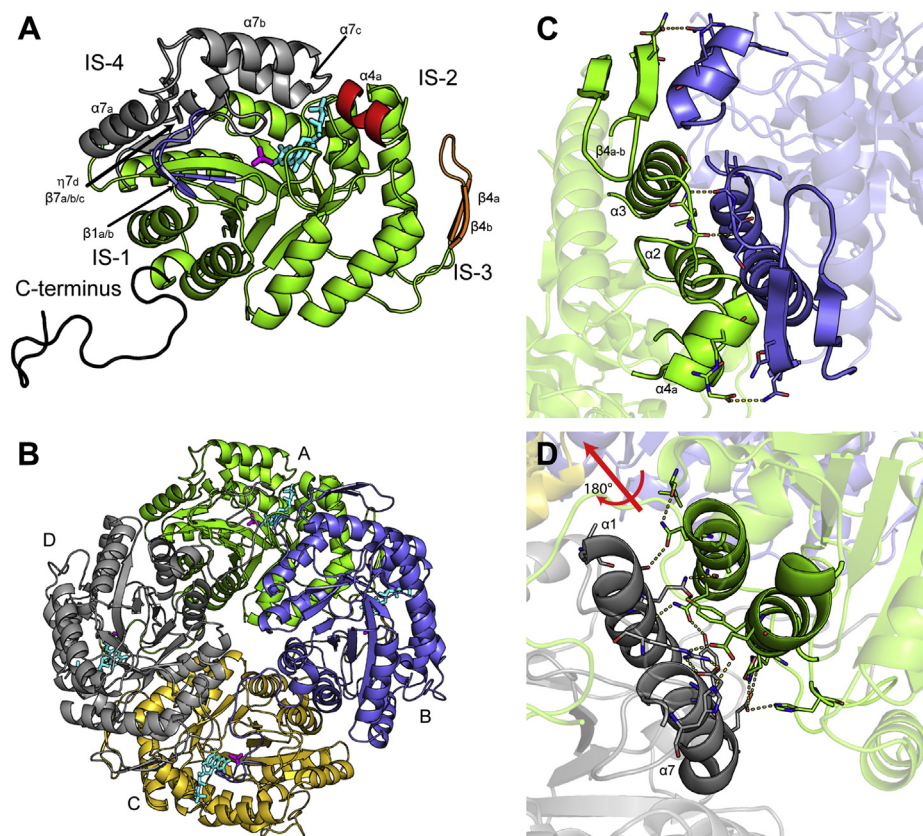


Figure 2. Structure of MsuD from *P. fluorescens*. A, MsuD adopts a classic $(\beta/\alpha)_8$ barrel (green) with four insertion regions, labeled IS-1 (blue), IS-2 (red), IS-3 (orange), and IS-4 (gray). The extended protein C terminus is colored black. FMN (cyan) and MS^- (magenta) are displayed as sticks. B, MsuD adopts a dimer-of-dimers between chains A and B and chains C and D. Protein chains are labeled A–D and colored green, blue, yellow, and gray, respectively. C, the dimer and D, tetramer interfaces are displayed with labels for only half of the symmetrical interfaces for clarity. Hydrogen bonding networks are drawn as yellow dashed lines.

observed protein–protein interfaces are highly conserved with SsuD (33); however, there is less than 15% conservation of interfacing residues with other structurally characterized group C monooxygenases. Therefore, this enzyme family can accommodate greatly divergent sequences that still preserve the overall quaternary structure.

The active site of MsuD is solvent exposed without FMN bound

In the absence of ligands, MsuD crystallizes in space group $P2_1$ with two MsuD tetramers (chains A/B/C/D and E/F/G/H) per asymmetric unit. The 2.8-Å resolution crystal structure of unliganded MsuD provides a snapshot of the enzyme prior to $FMNH^-$, MS^- , and O_2 binding. The two tetramers align with an overall RMSD of 0.34 Å for 1188 α atoms, indicating their structures are largely identical. Of interest, residues in all chains within IS-4 (between α_{7a} and β_{7a}) and from L354 to the C terminus lack any electron density, preventing their modeling and resulting in a very open structure (Fig. 4A). In each chain of ternary-MsuD, α_{7b} and α_{7c} from IS-4, as well as the protein C terminus of an adjacent chain, are ordered and observed to enclose the C-face of the $(\beta/\alpha)_8$ barrel, covering FMN and MS^- and blocking solvent access to the active site (Fig. 4B and electron density is shown in Fig. S4, A–C). The region spanning residues 250 to 282 in SsuD is similarly disordered (33) and reported as a “mobile loop region” (44).

We will refer to the α_{7b} – α_{7c} region (D250–L282) as a “lid.” In ternary-MsuD, we observe that the two lid helices pack antiparallel to each other and are connected by a three-residue linker (F263–S265). Both hydrophobic and polar interactions contribute to stabilizing the antiparallel positioning of α_{7b} and α_{7c} , which are splayed apart at an angle of 32.7° .

To further explore the results of ligand binding on MsuD, FMN alone or FMN with MS^- was soaked into crystals of unliganded MsuD using excess concentrations of each ligand. Ligand binding and protein ordering results are summarized in Table S3. Datasets for crystals soaked with FMN were prepared in two ways: an overnight soak with >1 mM FMN (binary-soak MsuD) and an overnight soak titrated with a low concentration of FMN until crystals became pale yellow (binary-titrated MsuD). Both crystals contained 2 mM FMN in their cryoprotectants, yielding datasets of approximately 2.8-Å resolutions. In both structures, six FMNs were placed but distributed differently between the two tetramers. Binary-soak MsuD contained FMN in chains A/B/C/D of the first tetramer and chains E/G of the second tetramer, whereas binary-titrated MsuD contained FMN in chains A/C/D and E/G/H. Chains A, C, and E contain the strongest electron density for FMN in both structures (Fig. S4, D and E). FMN binding led to ordering of the active site lid only when FMN bound with nearly full occupancy, and the B-factor values for portions of

Structural and mechanistic insights from ligand-bound MsuD

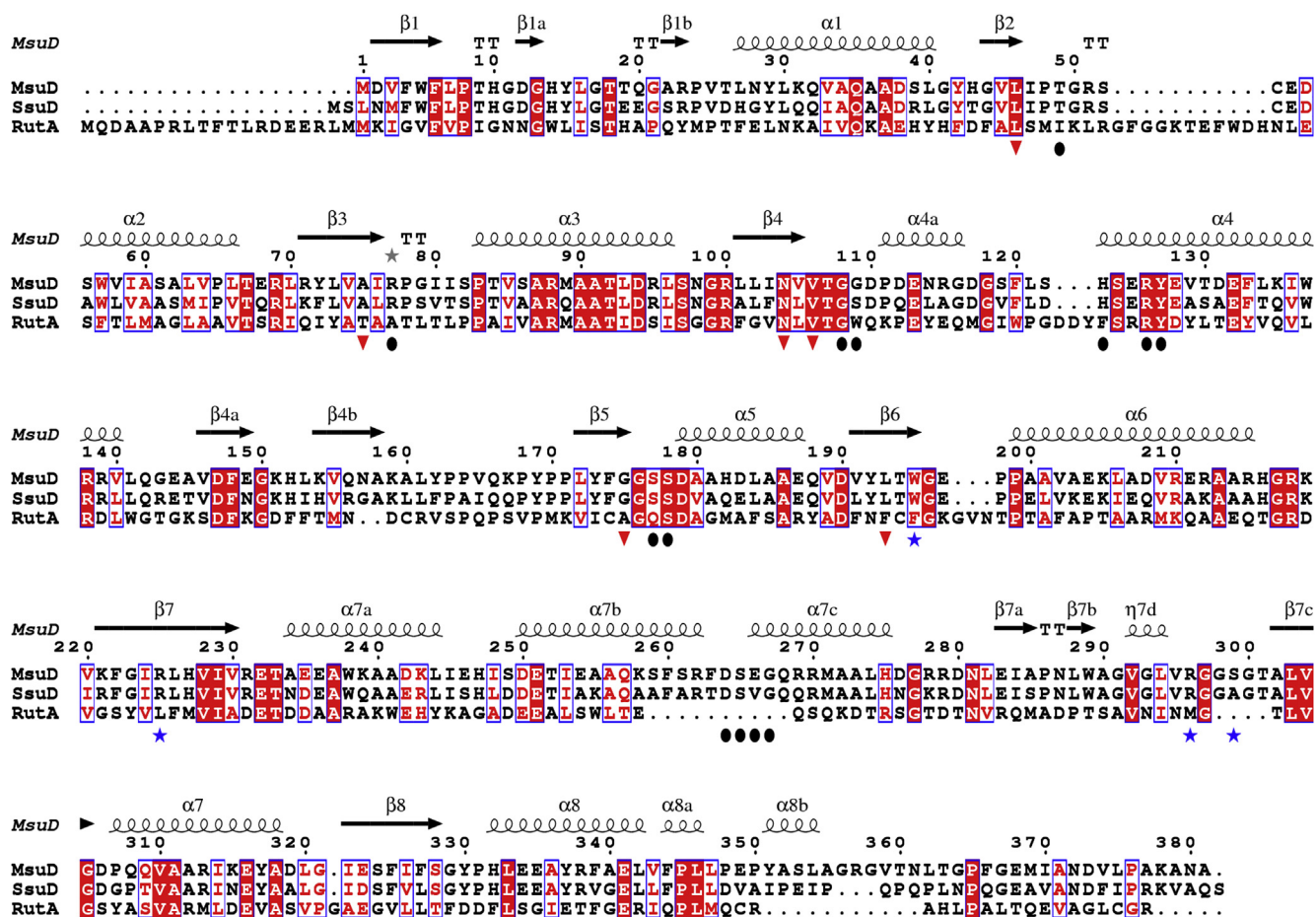


Figure 3. Sequence alignment of *P. fluorescens* MsuD with *E. coli* SsuD and *RutA*. Circles represent residues involved in FMN binding, red triangles are part of the proposed dioxygen binding motif, and blue stars are residues directly involved in MS⁻ binding within MsuD. An extended sequence alignment with structural homologs is included in Fig. S2. Blue boxes indicate conserved and similar residues, with strictly conserved residues highlighted in red and similar residues in red text. Secondary structural elements are shown for MsuD.

these lids are nearly double the average protein B factor, indicative of lid mobility within the crystal. Only one or three C termini were observed in both FMN-soaked structures: chain H in binary-soak MsuD extends into the active site of FMN-bound chain E, and chains B, D, and H in binary-titrated MsuD extend into the FMN-bound active sites of chains C, A, and E, respectively. For each active site that has an inserted C-

terminal tail, electron density is within the sulfonate-binding site that best refines as a succinate from the crystallization condition, likely serving as a mimic of MS⁻.

MsuD soaked with FMN and MS⁻ (ternary-soak MsuD) was solved to 2.7-Å resolution, and four FMN with MS⁻ pairs were placed within all protomers of the first tetramer (chains A/B/C/D) (electron density displayed in Fig. S4F). Within the

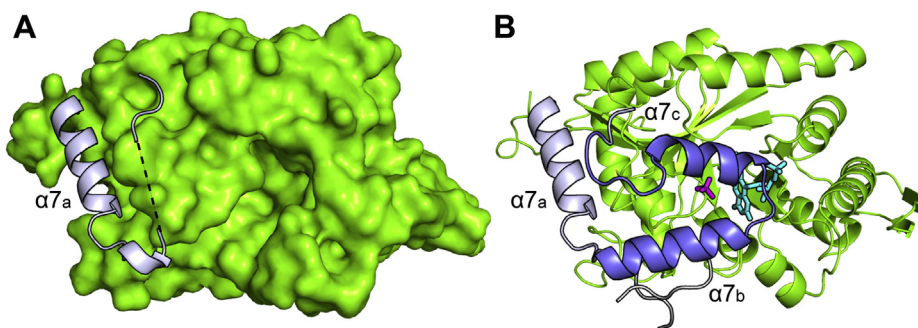


Figure 4. The lid of MsuD is important for active site enclosure. A, the active site lid is disordered in the unliganded MsuD structure, which is displayed as a green surface with portions flanking the lid region as ribbon (blue-white). The disordered region is represented as a dashed black line. B, the ternary-MsuD structure contains an ordered lid (blue) and an ordered C terminus from a separate protein molecule (gray). FMN (cyan) and MS⁻ (magenta) are displayed as sticks.

Structural and mechanistic insights from ligand-bound MsuD

second tetramer, three FMN molecules were placed within chains E/G/H, and MS^- could only be placed in chain E. Ligand binding similarly led to ordering of the active site lid with FMN bound at nearly full occupancies; the lid was observed in chains A/B/C/D of the first tetramer and chain E of the second tetramer. Chains D, G, and H have varying degrees of disorder for the lid region (Table S3). The C terminus is only observed for chain H, where it extends into the active site of the FMN-bound chain E. Of importance, weaker electron density for FMN and MS^- were observed in protomers containing a disordered lid region.

Taken together, ordering of the lid and the C terminus is linked to binding of FMN at near full occupancy. FMN binding appears to favor binding fully to one molecule of each homodimer related by C_2 symmetry (chains A/C and E/G in these structures) prior to binding in the remaining tetramer active sites. Binding of a succinate within the sulfonate-binding site is observed only with an ordered interaction of both the lid of the same protomer and the C terminus from an adjacent monomer, whereas MS^- requires ordering of the lid but can bind without an ordered C terminus. Variable numbers of ordered C termini were observed between the soaked structures, indicative of the higher mobility of the C terminus. We also cannot rule out that ordering of the lid and C terminus within an entire tetramer may be perturbed by the crystal lattice, which would explain why cocrystallization of ternary-MsuD yielded a tetramer with four ordered lids and four ordered C termini.

FMN binding to MsuD

Chain A of ternary-MsuD was used to analyze FMN binding as it is the most ordered chain of the tetramer and displays the lowest average B factor value. FMN binds at the C-terminal face of the TIM barrel and is positioned to engage in several conserved hydrogen-bonding and van der Waals interactions (Fig. 5A and Fig. S5). The phosphoryl moiety of FMN engages in seven or eight hydrogen-bonding interactions with the side chains of H123, Y127, and S178; the backbone NH groups of S177, S178, E266, and G267; and an ordered water molecule observed within chains A, B, and C (labeled "a"). The ribityl 3'-OH is approximately 3.5 Å from the backbone NH groups of G108 and G109, and the 4'-OH interacts with the hydroxyl of S265 and water "a." The FMN 2'-OH is positioned at the *re*-face of the isoalloxazine ring where it interacts with the backbone C=O of G175 and a crystallographic water (labeled "b"). Of interest, S265, E266, and G267 are located within the active site lid, and D264 of the same region is observed to interact with a conserved arginine (R126) from near the phosphoryl-binding site. The observed binding mode of FMN is expected to be similar for FMNH⁻ and is consistent with mutagenesis studies in the homolog BdsA, where mutations to residues corresponding to H123, R126, Y127, and S177 reduced the turnover rate by 45% to 95% (45).

The isoalloxazine ring of FMN binds with the O4 and N3 of its pyrimidine moiety positioned within hydrogen bonding distance to the backbone NH and C=O of T49 and its O2 near the

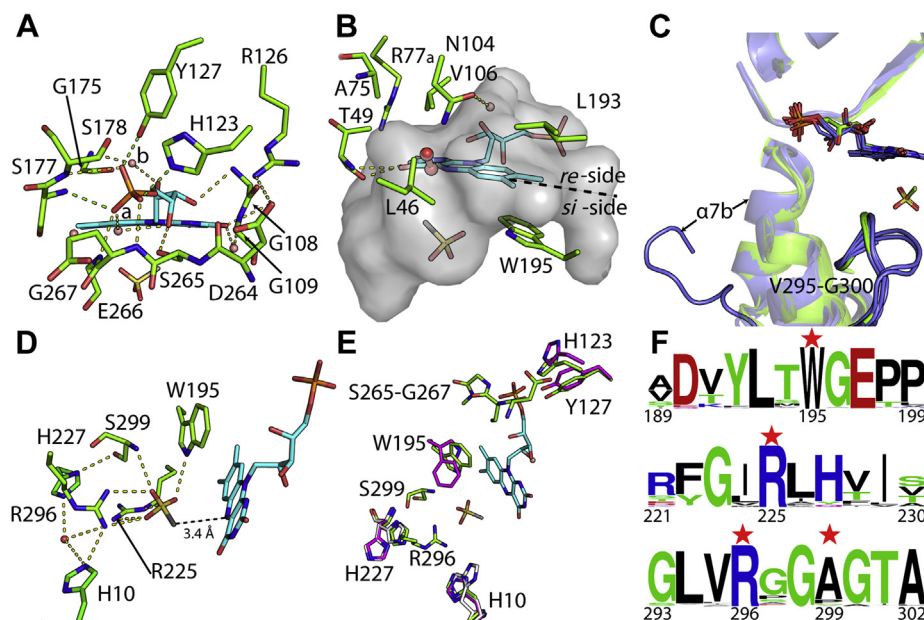


Figure 5. Residues involved in flavin and MS^- binding and a potential dioxygen-binding site. A, extensive polar contacts are made to the phosphate head and hydroxyls of the ribityl arm. B, Isoalloxazine interacting residues are shown. Residues of the dioxygen reactivity motif generate a partial pocket at the *re*-face of FMN where dioxygen is proposed to bind. Dioxygen is modeled in *red sphere and stick* representation from alignment of the FMN isoalloxazine ring of MsuD with the FMN of O₂-bound RutA (Protein Data Bank ID 6SGG). The active site cavity is displayed as a *gray surface*. C, orientations of the ribityl arm and the loop of V295-G300 show alternate conformations within different chains of binary-soak MsuD (*blue*) in comparison with ternary-MsuD (*green*). An alternative orientation for α_7b of the lid region in chain H is observed to direct the active site lid away from the FMN-binding site. D, MS^- binding positions the methyl group of MS^- approximately 3.4 Å from the N5 of FMN. E, an overlay of ternary-MsuD (chain A), binary-MsuD (FMN-bound chain A and unliganded chain H), and unliganded MsuD (chain A) shows residues that have alternate conformations when FMN is not present. Colors are as follows: ternary-MsuD carbons, *green*; binary-MsuD carbons, *white*; unliganded-MsuD carbons, *magenta*; FMN carbons, *cyan*; MS^- carbon, *gray*; oxygen, *red*; nitrogen, *blue*; sulfur, *yellow*; phosphorus, *orange*; and waters, *pink*. F, sequence logos are shown for the sulfonate-binding motif of flavin-dependent alkanesulfonate monooxygenases with *red stars* for interacting residues.

guanidinium of R77 (Fig. 5B); however, R77 is more disordered and observed in multiple orientations within chains A, B, and D. A π -stacking interaction is observed on the *si*-face of FMN with W195 (which adopts a different conformation with high B factor values in structures without FMN). The corresponding aromatic residue in unliganded homologous structures is directed away from the flavin-binding site (Fig. S6A). A pocket is formed by L46, T49, A75, N104, V106, and L193 on the *re*-face of the isoalloxazine ring where dioxygen likely binds (Fig. 5B), and comparatively no protein atoms are located within 5.4 Å of the C4a or N5 flavin positions that could stabilize dioxygen binding at the *si*-face. Indeed, dioxygen-binding studies within RutA (35) have identified that the *re*-face of the isoalloxazine ring is where dioxygen is expected to bind within several group C flavin-dependent monooxygenases and be positioned to react with the N5 as opposed to the C4a of flavin. The residues within MsuD that form the pocket share 50% sequence identity with RutA and are largely consistent in atomic coordinates; however, N104 and L46 are positioned closer to N5 of FMN than in RutA and result in a smaller binding pocket for dioxygen. An overlay of the FMN isoalloxazine rings from dioxygen-bound RutA and MsuD reveals the dioxygen site is too small. This binding mode is likely a result of having the product, oxidized FMN, bound. No other pocket is observed that could facilitate dioxygen binding. Therefore, the likely dioxygen-binding site in MsuD is on the *re*-face of FMN at the dioxygen reactivity motif, and formation of an N5-peroxyflavin would be consistent with the positioning of the dioxygen pocket closest to the N5 position on the *re*-face of the flavin and the need for inversion to position the reactive peroxy group on the *si*-face, within reaction distance of MS^- (detailed below).

Although group C flavin-dependent monooxygenases contain highly conserved flavin-binding pockets, the ribityl moiety of FMN is observed in different orientations. MsuD binds FMN with two ribityl torsion angles less than 80° and the 3'-OH 2.8 Å from N1 of the isoalloxazine ring, most closely resembling the structure of LadA (Fig. S6B and Table S4) (46). Of interest, in MsuD the crystallographic water near the FMN 2'-OH group is positioned to interact with N104 from the expected dioxygen-binding site (Fig. 5B). Therefore, the configuration of the FMN ribityl moiety may communicate to the proposed dioxygen-binding site by ordering N104 through a water-mediated hydrogen bond. In support of this hypothesis, the ribityl moiety of FMN observed within binary-soaked MsuD has less well-defined electron density than in ternary-MsuD (Fig. 5C and Fig. S4, C and D), particularly in monomers that have FMN partially bound but without an ordered lid. It is intriguing that alternate configurations for the loop V295-G300 near MS^- are observed and a second conformation is observed for residues S249–R272 ($\alpha 7_b$) in the lid of chain H. In this alternate lid conformation, the lid is directed away from the active site and the residues no longer adopt an ordered α -helical structure.

Methanesulfonate binding to MsuD

Within the structure of ternary-MsuD, a tetrahedral-shaped electron density peak is observed at the *si*-face of FMN in each

molecule of the tetramer. Placement of MS^- within the electron density refines nicely with B factors that are comparable with the FMN and nearby protein residues. The binding mode of MS^- positions its oxygen atoms within an extensive hydrogen-bonding network with the guanidiniums of R296 and R225, the indole NH of W195, and the backbone NH of S299 (Fig. 5D), resulting in placement of the methyl group 3.4 and 3.9 Å from the N5 and C4a atoms of FMN (respective average distances for the tetramer), whereas the sulfur is approximately 4.4 Å away from the N5. Therefore, the observed binding mode supports the methyl of MS^- as the site of oxygenation, and the closer distance to the N5 of FMN is consistent with use of an N5-peroxyflavin intermediate (35). To further validate the sulfonate-binding site, the R225A, R296A, and W195A variants of MsuD were generated (Fig. S7) and assayed in a coupled reaction with MsuE for conversion of MS^- to sulfite. In comparison with wildtype MsuD, which generated $28.5 \pm 1.1 \mu M$ sulfite within a 5-min reaction, no measurable sulfite was detected for R225A, R296A, or W195A.

Apart from R225, many of the residues that interact with MS^- are observed in alternative conformations or are completely disordered in unliganded MsuD (Fig. 5E), which supports binding of flavin prior to MS^- . The W195 interaction uses π -stacking to orient the indole NH group of FMN. The positioning of R225, R296, and S299 in ternary-MsuD is stabilized by two histidine residues. H10 adopts a nonstandard rotamer to interact with R296, and H227 packs edge-to-face against the guanidinium of R296 and is within hydrogen bonding distance of S299 and an ordered water molecule (Fig. 5D). A mutation in SsuD of the residues analogous to H10 and H227 showed little or no change in activity relative to wildtype, supporting a role for both H10 and H227 as secondary shell interactions to the MS^- -binding site (47). Both R296 and S299 are located within the V295-G300 loop connecting η_{7d} and β_{7c} after the lid region, and these residues are completely disordered in unliganded MsuD or are observed in alternate conformations with FMN absent or at partial occupancy (Fig. 5, C–E). As this loop contains half of the residues directly interacting with MS^- , it is being termed the sulfonate-binding loop, with the consensus sequence RGGGA identified from an alignment of 6428 sequences demonstrating between 30% and 40% sequence identity with MsuD (Fig. 5F). Indeed, the MsuD R296A variant reported here demonstrates no sulfite production, and mutations in SsuD of the arginine analogous to R296 to alanine or cysteine similarly resulted in a loss of enzyme function, with a lysine substitution yielding a 30-fold decrease in catalytic efficiency (44). Our structure of ternary-MsuD therefore has identified the sulfonate-binding motif of group C flavin-dependent alkanesulfonate monooxygenases to include W195 and R225 of the (β/α)₈ core and R296 and A299 (S299 in MsuD) from the sulfonate-binding loop.

An identified function of the protein C terminus in flavin monooxygenases

Available structures of group C FMN-dependent monooxygenases all contain a relatively open active site, despite positioning of the lid region when flavin binds. These proteins

Structural and mechanistic insights from ligand-bound MsuD

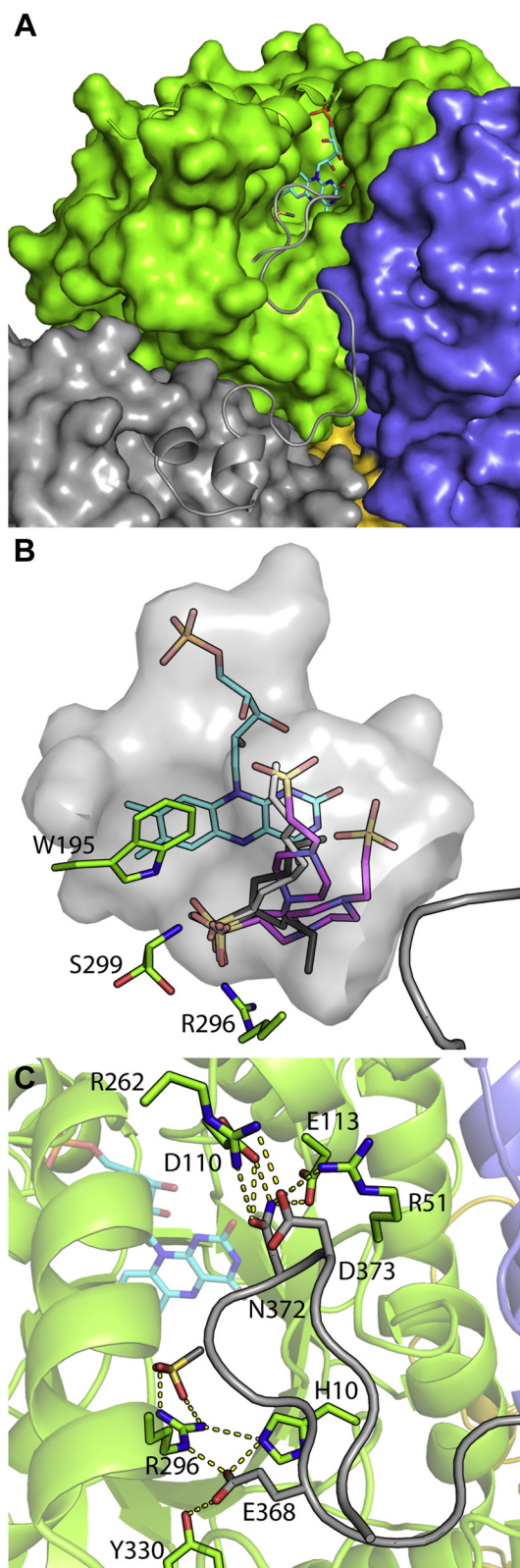


Figure 6. C terminus interactions within ternary-MsuD. A, the active site of MsuD chains A and B are displayed as green and blue surfaces with ribbons corresponding to the active site lid and the C terminus of chain D (gray). An identical interaction is observed for all chains in ternary-MsuD. B, example poses of docked C5 (gray) and C8 (light gray) sulfonates, and the buffer Pipes (magenta), position the sulfonate moiety similarly to MS^- , but with variable conformations of the substrate. The active site cavity is displayed as a gray surface with a gap where the C terminus encloses the

contain a long, disordered protein C terminus of unidentified function (30, 33, 37, 45, 48). To our surprise, the dataset of ternary-MsuD contained contiguous electron density for residues A355-A377 of the C terminus. The C terminus includes two short helices, a 3_{10} helix from F344-L347 and an α -helix from Y351-A355, followed by a long meandering loop that extends from each protein molecule, inserting itself within the active site of the adjacent molecule at the tetramer interface (Figs. 2B and 6A). Solvent-exposed portions of this tail display higher B factors and weak electron density; however, residues that interact with the adjacent protein molecule are well ordered with electron density for side chains.

When both the lid and C terminus are ordered and bound in ternary-MsuD, the active site appears completely enclosed from bulk solvent. The apparent volume is larger than MS^- , consistent with previously observed activity against larger sulfonate substrates. Therefore, molecular docking of substrates ranging in size from pentanesulfonate to Pipes was explored. Docking returned possible poses with the sulfonate moiety in a similar orientation as observed for MS^- , but with variable positioning of alkyl groups (Fig. 6B).

MsuD from P. fluorescens is an alkanesulfonate monooxygenase

To explore if the docked poses for larger alkanesulfonates could be competent for catalysis, recombinantly expressed MsuD from *P. fluorescens* was assayed with C_1 , C_5 , and C_8 sulfonates and the buffer molecules Hepes, Pipes, and Mops in a coupled reaction with MsuE. MsuD converted MS^- to sulfite with an activity of $26 \pm 2 \mu\text{mol min}^{-1} \mu\text{mol enzyme}^{-1}$, which is dramatically more active than the previous report of $0.035 \mu\text{mol min}^{-1} \text{mg}^{-1}$ from cell lysates in *P. aeruginosa* (17). Of importance, MsuD tested here was active with C_1 , C_5 , and C_8 alkanesulfonates (Table 1). Less activity was observed for the buffer molecules, indicating that, although each molecule could be successfully docked within ternary-MsuD, the lower enzyme activity suggests that Hepes, Pipes, and Mops bind in a less catalytically competent state.

C terminus interaction with the identified sulfonate-binding loop

Besides length, the overall sequences of the C termini within group C flavin-dependent monooxygenases lack any apparent conservation. However, their sequences contain multiple proline, glycine, and alanine residues, which are indicative that these segments can adopt sharp bends and turns (Fig. 3 and Fig. S2). In MsuD, five prolines enable the C terminus to adopt a kinked and winding path between protein molecules. Within the C-terminal sequence are three conserved amino acids with SsuD whose side chains appear important for binding (Fig. 6C). Residues N372 and D373 are presented to interact with conserved side chains of R51, D110, E113, and R262. The

active site. R296 is omitted for clarity. Hepes and Mops docking results had similar placements of the sulfonate moiety but are omitted for clarity as both represent smaller buffer molecules compared to Pipes. C, polar interactions of the C terminus are shown.

Table 1
Enzyme activity of *P. fluorescens* MsuD

Enzyme activity	Substrate					
	Methane sulfonate	Octane sulfonate	Pentane sulfonate	Hepes	Pipes	Mops
Rate $\mu\text{mol min}^{-1} \mu\text{mol enzyme}^{-1}$	26 \pm 2	30 \pm 3	19.4 \pm 1.4	3.6 \pm 1.2	4 \pm 2	5 \pm 2
% Relative activity	86 \pm 7	100 \pm 10	62 \pm 5	12 \pm 4	14 \pm 7	19 \pm 7

The desulfonation reaction was measured in a coupled reaction with 0.4 μM MsuD, 1 μM MsuE, 1 μM FMN, 500 μM NADH, and 500 μM of each respective sulfonate as detailed in Materials and Methods. The results are the average \pm standard error of three determinations, and relative activity is normalized to the activity obtained for octanesulfonate.

third residue, E368, extends directly into the active site and is positioned to interact with Y330 and R296. The importance of R296 in the sulfonate-binding loop suggests that ordering of the C terminus within an adjacent protomer's active site is linked to binding of both flavin and MS^- .

To further test the importance of the protein C terminus, a truncation mutation of MsuD was generated in which the last 16 residues were deleted (MsuD $^{\Delta\text{C-16}}$) by means of incorporating a stop codon (Fig. S7). It is surprising that the truncated MsuD variant showed lower detected sulfite formation at $0.45 \pm 0.01 \mu\text{mol min}^{-1} \mu\text{mol enzyme}^{-1}$ MsuD $^{\Delta\text{C-16}}$ under conditions demonstrating activity for wildtype MsuD ($5.7 \pm 0.02 \mu\text{mol min}^{-1} \mu\text{mol enzyme}^{-1}$). Therefore, the C terminus of MsuD has three previously unidentified roles in monooxygenase function: contributing to the tetramer interface, ordering active site residues for MS^- binding and catalysis, and closing off the active site from solvent in conjunction with the lid region.

Discussion

Flavin-dependent monooxygenases catalyze a breadth of oxygenation reactions, and group C flavin-dependent monooxygenase reactions include light emission, Baeyer-Villiger oxidation, epoxidation, sulfoxidation, hydroxylation, and desulfurization (1). Despite their chemical utility, there are many outstanding questions regarding how these enzymes function at the molecular level, in part due to the limited structures for group C monooxygenases with substrates bound. The structures of MsuD reported here represent the first crystal structure of MsuD, and we have obtained crystallographic snapshots of three protein states: before flavin binding, after flavin binding, and after flavin and MS^- binding.

Previous studies with MsuD from *P. aeruginosa* (87% sequence identity) showed desulfonation rates of larger sulfonates between 1.3% and 31% the rate for MS^- (Table S5) (17), therefore MsuD was identified as a methanesulfonate monooxygenase. Our results, however, support MsuD from *P. fluorescens* as an alkanesulfonate monooxygenase with similar activity against both small (C_1) and medium (C_5 and C_8) alkanesulfonates. SsuD, in comparison, shows little activity against a C_1 substrate but a range of activity with medium and larger substrates (26). The active site residues within MsuD from *P. aeruginosa* and *P. fluorescens*, as well as for *E. coli* SsuD, are conserved; therefore, it is likely that residues outside of the primary active site define substrate preference.

MsuD and other group C monooxygenases adopt a $(\beta/\alpha)_8$ TIM barrel fold containing insertions involved in oligomerization and catalysis. In MsuD, IS-2 and IS-3 are involved in

forming the MsuD homodimer, and two homodimers arrange to form a dimer-of-dimers architecture. Truncation of the C terminus of MsuD yielded a dimeric enzyme with decreased function, indicating that MsuD activity is dependent on tetramerization. Homotetramers are only observed for MsuD, SsuD (33), and BdsA (32), all of which catalyze C–S bond cleavage. DmoA similarly catalyzes C–S bond cleavage of dimethylsulfide; however, unliganded DmoA is reported to be a homodimer (30). The similar or conserved residues within group C flavin-dependent monooxygenases primarily are involved in flavin and dioxygen binding. No sequence conservation is observed in the MS^- -binding site outside of alignment with SsuD, attesting to the diverse scope of substrates recognized by group C flavin-dependent monooxygenases.

The structures of MsuD with and without ligands support ordered binding for FMNH^- and MS^- , and the preferential binding of FMN first within chains A/C and E/G is suggestive of possible cooperativity. Without ligands, the active site lid, the sulfonate-binding loop, and the protein C terminus are disordered. With FMN added, we observe binding of FMN within a subset of MsuD active sites that increases with higher concentrations of FMN. Binding of FMN leads to a partial ordering of the active site lid within some MsuD protomers, but it is not until MS^- (or a mimic such as succinate) is also bound that ordering of the sulfonate-binding loop, lid, and the protein C-terminal tail occurs, fully enclosing the MsuD active site, consistent with biochemical observations for the SsuD system (47, 49, 50). Studies in SsuD highlight the lid's importance in the production of sulfite and protecting bound FMNH^- from oxidation (51). An analogous lid region is observed in all homologs of MsuD containing between three and five helices (Fig. S8A), and structures of EDTA monooxygenase EmoA (48) similarly describe an open and closed lid conformation that is linked to FMN binding. Therefore, flavin binding likely stabilizes the closed conformation of the lid in a shared general mechanism for linking flavin binding to lid closure prior to catalysis within this enzyme family (44). It remains to be seen if other group C flavin-dependent monooxygenases will utilize their extended protein C termini. In the crystal structures of MsuD homologs, the C terminus is either entirely disordered or partially built but in a different orientation (Fig. S8B). Truncation of the MsuD C terminus resulted in a 13-fold decrease in MS^- conversion to sulfite. Indeed, structures reveal that the protein C terminus donates E368 as a second coordination sphere contact to the MS^- -binding site that is conserved in SsuD. The C terminus is highly flexible, consistent with a role in closing off the active site during catalysis. A tight interaction would interfere with substrate and

Structural and mechanistic insights from ligand-bound MsuD

product diffusion. It is possible that the observed interactions of the C terminus are a mechanism for sensing active site lid closure as a result of flavin and MS^- binding. Alternatively, only when the C terminus is ordered was succinate, a substrate mimic, observed in the sulfonate-binding site, implying the C terminus may play a role in sulfonate affinity. Future studies are needed to explore the role of the MsuD C terminus and the possible cooperativity in flavin binding.

The structure of ternary-MsuD enables us to envision how the components of an alkanesulfonate monooxygenase are oriented for catalysis and to revisit the proposed chemical mechanism. The classical monooxygenase proposed mechanism invokes formation of a C4a-(hydro)peroxyflavin intermediate (3)

capable of reacting with MS^- (Fig. 7); however, the structure identifies the reactive dioxygen motif recently reported by Matthews *et al.* (35) on the *re*-face of the flavin. Studies of RutA support formation of an N5-peroxyflavin intermediate from dioxygen bound on the *re*-face, and inversion at N5 would bring the reactive peroxy moiety within reaction distance to the substrate bound on the *si*-face. This model demonstrates how MsuD and related monooxygenases can have variable substrate-binding sites on the *si*-face and still retain the residues needed for dioxygen binding on the opposite side of the flavin. The reactive dioxygen motif of MsuD is conserved in the positioning of residues in RutA, except for a shift of L46 and N104 closer to the N5 position of the oxidized flavin. Therefore, the

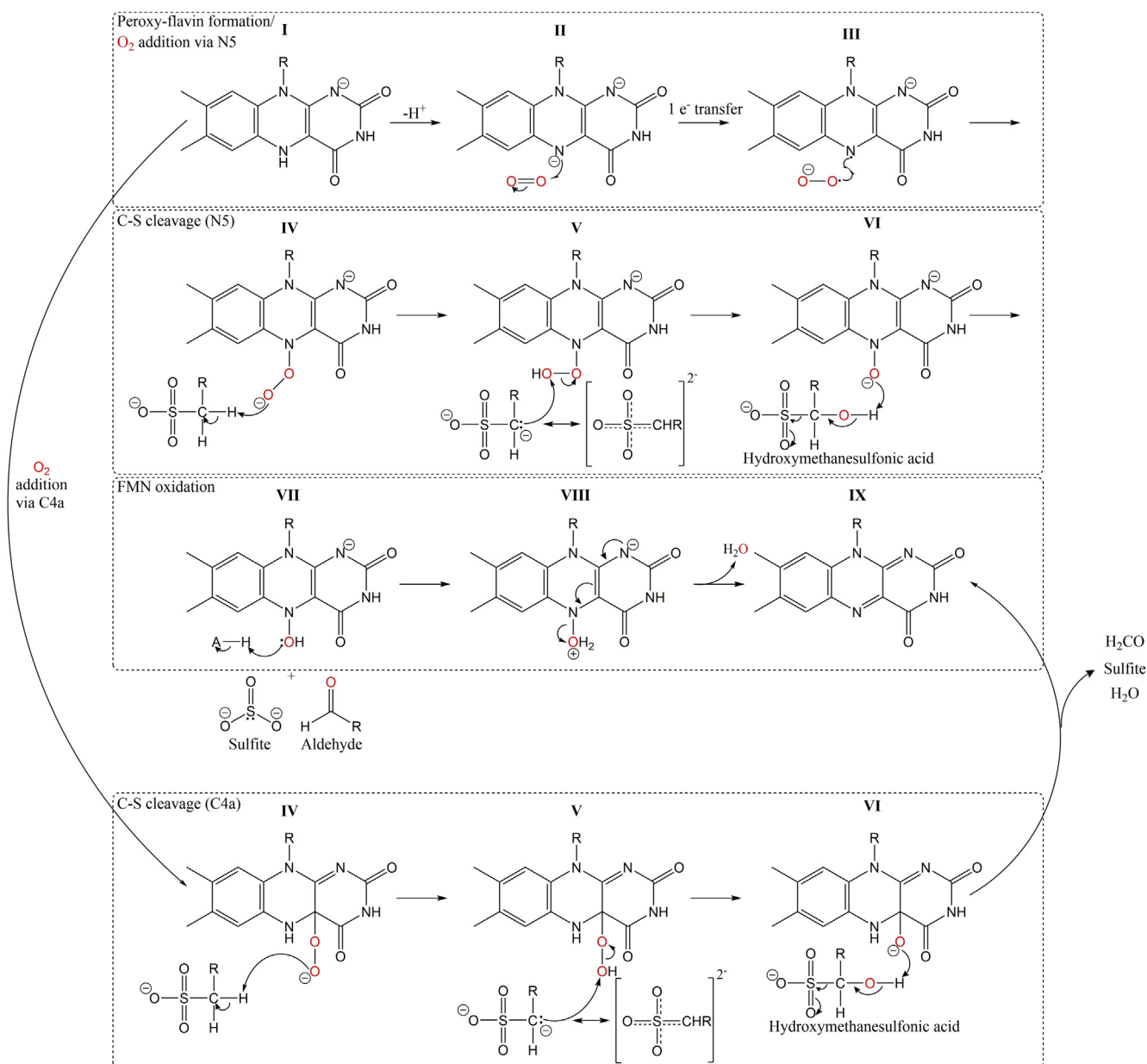


Figure 7. Proposed mechanism for two-component flavin-dependent alkanesulfonate monooxygenases based on the Ternary-MsuD structure. Dioxygen is colored red, and R can be a hydrogen or larger alkyl chain. The proton donor AH is likely a water molecule. Alternatively, hydroxymethanesulfonic acid may nonenzymatically convert to sulfite and aldehyde. Positioning of MS^- is consistent with oxygenation at the methyl, with both the N5- and C4a-pathways included for comparison. Reduced flavin is generated by the NADH:FMN oxidoreductase MsuE (not depicted).

dioxygen-binding site in ternary-MsuD appears to be partially occluded, which is likely a result of having the product, oxidized FMN, bound in place of the substrate, reduced FMN. Nonetheless, there are no additional pockets on the *re*-face of the flavin for dioxygen to bind, and a large active site cavity where variable alkanesulfonates would bind is present on the *si*-face without any residues to position dioxygen to react with FMNH⁻ at the C4a position; therefore, dioxygen is predicted to bind in a similar mode as observed in RutA, positioned to form an N5-peroxyflavin (Steps I to III, Fig. 7). Within the substrate alkanesulfonate there are only two possible reactive sites: the C₁ position and the sulfur (49). As the product is sulfite, it would be unexpected for the peroxyflavin to directly attack the sulfur (52). The methyl group of MS⁻ is closest to the N5 position of FMN at 3.4 Å apart, supporting the C₁ position as the reactive site. Therefore, the N5-peroxyflavin can deprotonate the C₁ position, generating a resonance stabilized carbanion that can in turn attack the generated N5-hydroperoxy flavin, yielding 1-hydroxyalkanesulfonate and a reduced flavin-N5-oxide (Steps IV and V, Fig. 7). At this point the MsuD mechanism has diverged from the general proposed mechanisms for RutA, DszA, and HcbA1 (35), in which the organic substrates and products are in the same oxidation state and the cosubstrate FMNH⁻ forms flavin-N5-oxide; a reducing equivalent of NADH is needed to reform FMN (52–54). However, in MsuD the starting alkanesulfonate is more reduced than the product, formaldehyde, and the cosubstrate FMNH⁻ stays in its reduced form after hydroxylation. In effect, two more electrons must be lost from the substrate MS⁻ in comparison with the RutA, DszA, and HcbA1 reactions, and these electrons enable the formation of reduced flavin-N5-oxide, which can reform FMN after protonation and loss of water, precluding the need for an additional NADH molecule (Steps VII to IX, Fig. 7, and Fig. S9). This mechanistic scheme requires two protonation events. The generated 1-hydroxyalkanesulfonate could serve as a proton source in Step VI or alternatively convert to sulfite and an aldehyde nonenzymatically as has been proposed for the degradation of *n*-alkane-1-sulphonates (55) and taurine (9). No residues are observed in the ternary-MsuD structure to be near the C₁-alkanesulfonate position, and N104 is the only polar residue near the N5 position of FMN. Considering asparagine is not a strong general acid, solvent is the most likely proton donor. After catalysis completes, the mobile C terminus and active site lid must move to open the active site to bulk solvent and allow product release.

Our mechanistic proposal is consistent with previous kinetic isotope experiments (50) and biochemical studies of SsuD (47). A primary kinetic isotope effect for C₁ deuterated octanesulfonate in SsuD supports deprotonation of C₁ as the rate-limiting step, and solvent isotope effect studies support a protonation event after the rate-limiting step. The pH profile of SsuD identified the need for a deprotonated group with a pK_a value of 6.6 and a protonated group with a pK_a value of 9.5. The ternary-MsuD structure does not reveal any residues that could serve as the source for either pK_a value, leaving only flavin as a candidate. The higher pK_a value is near the reported pK_a of 9.8 for the C4a-(hydro)peroxyflavin group of

p-hydroxyphenylacetate-3-hydroxylase (56), suggesting that an N5-(hydro)peroxyflavin may have a similar pK_a. The pK_a of free FMNH₂ is reported to be 6.7 (57); therefore, the need for FMNH⁻ matches the lower pK_a. The ribityl moiety of FMN within MsuD is in a compact state that positions the 3'-OH to form an intramolecular hydrogen bond with N1 of the isoalloxazine ring. The binding mode of FMN connects to the active site lid, the MS⁻-binding site, the protein C terminus, and the proposed site for dioxygen binding; therefore, maintenance of the ribityl moiety in a catalytically competent configuration may serve as the source of the acidic limb on the pH profile of SsuD. Currently, it is unclear why MsuD and SsuD have different preferences in substrate size, and future studies will be necessary.

The series of crystal structures of MsuD from *P. fluorescens* provides a timely addition to enzymes proposed to use an N5-(hydro)peroxyflavin intermediate, and the structure of ternary-MsuD has enabled us to revisit the alkanesulfonate monooxygenase reaction. Although we cannot fully rule out use of a C4a-(hydro)peroxyflavin intermediate, the separation of the MS⁻-binding site from the predicted oxygen-binding motif residues is consistent with a mechanism that engages both N5-peroxyflavin and N5-hydroperoxyflavin intermediates in the hydroxylation at the C₁ position of an alkanesulfonate, and future work to explore possible intermediates in the reaction will be illuminating. Alkanesulfonates represent a family of substrates that undergo oxidation to an aldehyde and sulfite; therefore, unlike the reactions of the group C two-component flavin-dependent monooxygenases RutA, DszA, and HcbA1, a second equivalent of NADH would not be necessary to complete the catalytic cycle. Rather, oxidation of the alkanesulfonate substrate is the likely source of electrons. The structures of MsuD characterize movements of the protein linked to FMN and MS⁻ binding, including ordering of a sulfonate-binding loop and active site lid, as well as identify a novel function for the extended C terminus of the monooxygenase in closing off the enzyme active site after MS⁻ binds and maintaining the functional tetramer.

Experimental procedures

Materials

Recombinant DNA, molecular cloning, and microbiological procedures were carried out using standard procedures (58). PCR primers were supplied from Integrated DNA Technologies and Life Technologies Corporation, and *P. fluorescens* was kindly made available by Katharina Ribbeck from the Department of Biological Engineering at MIT. GoTaq green master mix, blue/orange 6x loading dye, 1 kb DNA ladder, and LigaFast Rapid DNA Ligation System were purchased from Promega. Alkaline phosphatase was purchased from Roche Applied Science. Chemically competent *E. coli* cells were purchased from Invitrogen. DNA purification kits were purchased from Qiagen and New England Biolabs and used as instructed. Restriction enzymes, polynucleotide kinase, and T4 DNA ligase were purchased from New England Biolabs. Ni

Structural and mechanistic insights from ligand-bound MsuD

Sepharose resin was purchased from GE Healthcare Biosciences. Expression vector pET28a was purchased from Novagen. All crystallization reagents were purchased from Hampton Research, and other chemicals were purchased from commercial suppliers and used as received.

Construction of the MsuD_{PflU} expression plasmid

The gene *Pfl01_3916* (GenBank accession ABA75653.1) was amplified by colony polymerase chain reaction (PCR) using GoTaq green master mix and the DNA primers shown in Table S6 for cloning into expression vector pET28a. The amplified DNA was analyzed by agarose gel electrophoresis and visualized using ethidium bromide staining. Crude PCR product was purified using a Qiagen QIAquick spin purification kit and stored at -20°C . The amplified DNA and pET28a were digested with restriction enzymes NheI and HindIII, purified as described above, and ligated using the LigaFast Rapid DNA Ligation System. The resulting pET28msuD_{PflU} vector was transformed into XL10 *E. coli* cells. Colonies were screened by PCR, and amplified DNA was analyzed by agarose gel electrophoresis. Colonies that exhibited amplified DNA consistent with the desired gene were used to inoculate 5 ml sterile LB broth with kanamycin sulfate (50 $\mu\text{g}/\text{ml}$) in 10 ml culture tubes and incubated overnight at 37°C and 250 rpm. The culture tubes were centrifuged for 20 min at 3000g and 4°C , and the supernatant was decanted. The plasmid DNA of the pelleted cells was isolated and purified using a Qiagen QIAprep spin miniprep kit and stored at -20°C . The produced pET28msuD_{PflU} vector was verified by Sanger sequencing (MIT Biopolymers Laboratory).

Generation of MsuD_{PflU} variants

MsuD mutants W195A, R225A, R296A and the truncation of 16 residues from the C terminus (MsuD ^{$\Delta\text{C-16}$}) were constructed by site-directed mutagenesis of the previously described MsuD plasmid with designed primers listed in Table S6. PCR products were treated in reactions containing approximately 2.5 ng/ μl of DNA, polynucleotide kinase, T4 DNA ligase, and T4 DNA ligase reaction buffer, and the ligated plasmids were used to transform T7 express and BL21(DE3) cells. Overnight 10 ml cultures containing LB broth with 50 $\mu\text{g}/\text{ml}$ kanamycin sulfate were inoculated with single bacterial colonies and allowed to grow overnight at 37°C and 250 rpm. Cell cultures were harvested by centrifugation for 20 min at 3000g and 4°C , and the supernatant was decanted. Plasmid DNA was isolated and purified using the Monarch Plasmid Miniprep Kit, and the desired mutations were verified by Sanger sequencing (Quintara Biosciences).

Production and purification of MsuD and MsuD variants

A starter culture (10 ml) of *E. coli* BL21(DE3) cells transformed with pET28msuD_{PflU} was transferred into 1 l lysogeny broth (LB)-Miller containing 50 $\mu\text{g}/\text{ml}$ kanamycin sulfate. The culture was grown at 37°C with shaking at 250 rpm until the OD₆₀₀ reached between 0.5 and 0.6, at which point the culture was cooled on ice. Cells were induced with isopropyl β -D-1-

thiogalactopyranoside (IPTG) at a final concentration of 50 μM and allowed to grow overnight at 16°C with shaking at 250 rpm prior to harvesting at 3000g, 4°C for 15 min. The cell pellets were resuspended in 10 ml of wash buffer (20 mM Tris, 100 mM NaCl, pH 8.0) and centrifuged at 10,000g, 4°C for 10 min. The pellets were transferred to a 50-ml Falcon tube to yield approximately 4.0 g wet cell paste, frozen in liquid N₂, and stored at -80°C .

The harvested cell paste was thawed, resuspended in 22 ml of lysis buffer (20 mM Tris, 500 mM NaCl, pH 8.0), and supplemented with 5 $\mu\text{g}/\text{ml}$ lysozyme, 10 $\mu\text{g}/\text{ml}$ RNase A, 10 $\mu\text{g}/\text{ml}$ DNase I, and a Pierce protease inhibitor tablet (Thermo Fisher Scientific). The cells were lysed at 24,000 psi *via* three passages through an Avestin EmulsiFlex-B15 high-pressure homogenizer. After centrifugation at 10,000g, 4°C for 20 min, the supernatant was loaded at 4°C onto a 2.1-cm³ hand-packed Ni Sepharose gravity column preequilibrated with lysis buffer. The column was washed with ten column volumes of wash buffer (20 mM Tris, 300 mM NaCl, 20 mM imidazole, pH 8.0). MsuD was eluted with six column volumes of elution buffer (20 mM Tris, 500 mM NaCl, 250 mM imidazole, 10% (w/v) glycerol, pH 8.0), and fractions were qualitatively analyzed for protein using the Bradford method and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Fractions containing protein were combined and buffer exchanged into storage buffer (20 mM Tris, 100 mM NaCl, pH 8.0) and concentrated using a 30,000-Da MWCO centrifugal filter unit (MilliporeSigma). MsuD was flash frozen in liquid nitrogen and stored at -80°C temporarily prior to purification over a HiLoad Superdex 75 pg 16/600 column (GE Scientific) equilibrated with size exclusion buffer (50 mM Tris, 100 mM NaCl, pH 8.0). Protein-containing fractions were analyzed *via* SDS-PAGE, confirming a purity $>90\%$ *via* ImageJ analysis (59). The protein fractions were combined and concentrated to 8 mg/ml with a 10-kDa centrifugal filter (Vivaspin) at 4°C , aliquoted, flash frozen in liquid nitrogen, and stored at -80°C . Protein concentration was determined by absorbance at 280 nm, using the calculated molar absorptivity 52,370 M⁻¹ cm⁻¹ (60).

The MsuD variants were expressed and purified in the same manner as wildtype MsuD, except using a Nuvia 1-ml column (Bio-Rad) for the first purification step. Protein concentrations were determined using the calculated molar absorptivity value of 46,870 M⁻¹ cm⁻¹ for W195A and 52,370 M⁻¹ cm⁻¹ for MsuD ^{$\Delta\text{C-16}$} , R225A, and the R296A mutants (60). The protein fractions were combined and concentrated to between 100 and 750 μM with a 10-kDa centrifugal filter (Vivaspin) at 4°C , aliquoted, flash frozen in liquid nitrogen, and stored at -80°C .

Activity assays for MsuD

MsuD variants were assayed against MS⁻ in the presence of recombinantly produced MsuE (27), and the liberation of sulfite was monitored by reaction with 5,5-dithio-bis-(2-nitrobenzoic acid) (DTNB) (17) using the molar extinction coefficient of the TNB anion of 14.1 mM⁻¹ cm⁻¹ (61). Reactions of 400 μM MS⁻, 1 μM MsuE, 1 μM FMN, 100 mM NaCl, and 1 μM MsuD in a buffer of 50 mM Tris-HCl at pH

7.5 were prepared in a multiwell plate with five replicates. Reactions were initiated by the addition of NADH to a final concentration of 500 μM , and final reactions contained a volume of 100 μl and were allowed to run for 5 min with shaking at 1000 rpm in a MixMate (Eppendorf) before quenching with final concentration 2 M urea. DTNB was added to each well to a final concentration of 1 mM, and absorbance at 412 nm was measured using the SpectraMax i3x Multi-Mode Microplate Reader (Molecular Devices). The background reaction of MS^- in the presence of MsuE, NADH, and FMN, but lacking MsuD, was subtracted to account for nonenzymatic reaction of FMNH^- with oxygen and MS^- . Reactions were repeated with 10 μM FMN and 3 μM MsuE to ensure that lack of reduced FMN was not a limiting factor to the assay.

To explore the substrate profile of MsuD, MsuD was assayed against MS^- , pentanesulfonate, octanesulfonate, Pipes, Mops, and Hepes as detailed above with minor modifications. Reactions contained 500 μM sulfonate, 500 μM NADH, 1 μM FMN, 100 mM NaCl, and 0.4 μM MsuD in a buffer of 50 mM Tris-HCl at pH 7.5 and were initiated by the addition of MsuE to a final concentration of 1 μM , and final reactions contained a volume of 100 μl and were allowed to run for 3 or 5 min with shaking at 500 rpm in a MixMate (Eppendorf) before quenching with 50 μl of 2 M urea. A longer reaction time was used for the buffer molecules, which showed lower rates of sulfite production by MsuD. The background reaction of each sulfonate in the presence of MsuE, NADH, and FMN, but lacking MsuD, was subtracted to account for nonenzymatic reaction of FMNH^- with oxygen and any sulfonate substrates. A control of each substrate in buffer was reacted with DTNB under the same conditions to account for any background sulfite from the commercial supplier.

Analytical gel filtration of MsuD and MsuD^{AC-16}

Purified MsuD and high-molecular-weight standards (Cytiva) were injected using a 100- μl loop at 4 °C onto an Enrich SEC 650 10 \times 300 column (Bio-Rad). The column was preequilibrated with 50 mM sodium phosphate and 150 mM NaCl at pH 7.2. Protein sample runs were performed at 0.75 ml/min and monitored at 280 nm.

Crystallization of MsuD

Initial crystallization conditions were identified in the polyethylene glycol (PEG)/ion HT Screen using a Phoenix pipetting robot (Art Robbins Instruments). Rectangular crystals formed within 24 to 48 h at 18 °C. Optimal crystals were obtained by sitting drop vapor diffusion in reservoir conditions of 12% to 18% (w/v) PEG 3350, and either 0.14 to 0.20 M succinate or 0.2 to 0.3 M sodium acetate set up at room temperature. Protein samples were prepared by spin filtration using a 0.22- μm Millipore CL Centrifugal Filter Unit. Drops at a ratio of 1 μl MsuD (8 mg/ml) to 1 μl reservoir were equilibrated against 600 μl of reservoir. Crystals ranging in size from 50 to 200 μm in the longest dimension were cryoprotected with increasing increments of glycerol to a final concentration

of 15% to 20% (v/v) with an increase of 3% to 5% (w/v) PEG 3350. Crystals were mounted in loops and cryocooled in liquid nitrogen.

Crystallization of the binary-soak MsuD and binary-titrate MsuD complexes with FMN

Crystals of MsuD in complex with FMN (MsuD•FMN; referred to as binary-soak MsuD) were prepared by soaking MsuD crystals grown in the succinate conditions with reservoir solution supplemented with FMN. A volume of 2 μl reservoir solution supplemented with 2 mM FMN was added to a 1- μl drop containing MsuD crystals and incubated for over 16 h at 18 °C. For the binary-titrated MsuD structure, FMN was introduced to crystals through an incremental addition of a 2 mM FMN solution using a crystallization loop until crystals turned a pale yellow. All crystals, soaked and titrated, were cryoprotected in reservoir supplemented with 2 mM FMN and increasing increments of glycerol to a final concentration of 15% to 20% (v/v) and an increase of 3% to 5% (w/v) PEG 3350. Crystals were mounted in loops and cryocooled in liquid nitrogen.

Crystallization of the ternary-MsuD and ternary-soak MsuD complexes with FMN and methanesulfonate

Structures of the ternary complex with FMN and MS^- (MsuD•FMN• MS^-) were obtained by both soaking (ternary-soak MsuD) and cocrystallization (ternary-MsuD) experiments, yielding ternary complex structures in monoclinic and hexagonal space groups, respectively. For cocrystallization, MsuD at 8 mg/ml was incubated with 2 mM FMN and 2 mM MS^- on ice for 15 min prior to spin filtration using a 0.22- μm Millipore CL Centrifugal Filter Unit. Crystallization was performed in a sitting drop vapor diffusion experiment with 1 μl MsuD•FMN• MS^- mixture and 1 μl of reservoir (0.24 M succinate and 12% (w/v) PEG 3350). Hexagonal crystals ranging in size from 50 to 200 μm in the longest dimension formed within 24 to 48 h at 18 °C. For soaking experiments, MsuD crystals were soaked in reservoir solution supplemented with FMN and MS^- . A volume of 2 μl reservoir solution supplemented with 2 mM FMN and 2 mM MS^- was added to a 1- μl drop containing MsuD crystals and incubated for over 16 h at 18 °C. Crystals were cryoprotected in the same manner as binary-MsuD, but with cryoprotectant supplemented with 2 mM FMN and 2 mM MS^- , and cryocooled in liquid nitrogen.

X-ray diffraction data, structure determination, and refinement

X-ray diffraction data for crystals of *P. fluorescens* MsuD were collected at the Advanced Photon Source, Argonne National Laboratory, beamlines 24ID-C and 24ID-E. Datasets were collected at 100 K on either an Eiger 16M detector or a Pilatus 6M detector, using 0.2° oscillations with an exposure time of 0.2 s. Crystals of unliganded MsuD diffracted to 2.8-Å resolution and indexed to space group $P2_1$ with eight molecules per asymmetric unit, yielding a Matthew's coefficient of

Structural and mechanistic insights from ligand-bound MsuD

2.43 Å³/Da and solvent content of 49.4%. Soaked crystals of binary-MsuD and ternary-MsuD complexes diffracted to approximately 2.8-Å resolution, with the same unit cell parameters and symmetry as unliganded MsuD crystals. Cocrystals of the ternary-MsuD complex diffracted to approximately 2.4-Å resolution and indexed to the higher symmetry space group of *P*6₁ with four molecules per asymmetric unit, a Matthew's coefficient of 2.37 Å³/Da, and a solvent content of 48.1%. Crystal parameters are reported in Table 2.

The datasets were processed using X-ray detector software (62), and data quality and model refinement statistics are reported in Table 2. Molecular replacement (MR) of the unliganded MsuD dataset was performed in PHASER (63) within Phenix (64) using the homolog SsuD from *E. coli* K12 (Protein Data Bank [PDB] ID 1NQG (33), seqID: 66.4%). The MR model was pruned using CHAINSAW (65), enabling PHASER to place eight protein chains within the asymmetric unit in the initial *P*2₁ dataset. Each protein molecule was refined with rigid body refinement prior to defining NCS matrices for further refinement. The initial model was refined with simulated annealing, energy minimization, and grouped B-factor refinement using strict noncrystallographic symmetry constraints within Crystallography & NMR System (66), and iterative model building was performed in COOT (67). Non-crystallographic symmetry was relaxed to torsional restraints, and final refinement cycles with iterative model building were carried out in Phenix (64) and COOT (67). Waters were added toward the end of refinement, and final model refinements included TLS groups (Translation-Libration-Screw-rotation model) generated by Phenix.

For MR of datasets containing ternary-MsuD, the built chain A from unliganded MsuD (PDB ID 7JV3) was used as the initial search model. MR placed four protein chains in the *P*6₁ asymmetric unit. FMN was added after rigid body refinement of the protein model. FMN was placed in 4/4 protein chains in the ternary-MsuD cocrystal. Subsequent rounds of refinement and iterative model building were carried out in COOT and Phenix (64, 67). Toward the end of refinement, waters and the MS⁻ ligand were placed and refined.

For MR of ternary-soak MsuD, eight protein chains were placed within the asymmetric unit in the *P*2₁ dataset, using chain A of ternary-MsuD (PDB ID 7JW9) as an initial search model. FMN was placed in 7/8 protein chains, and MS⁻ was placed in 5/8 protein chains; chain G and chain H each have a chloride ion refined in the MS⁻ site. Refinement and model building proceeded as for ternary-MsuD. Occupancy values were refined for FMN and MS⁻ in the later stages of refinement, and final model refinements included single TLS groups per protein chain.

Phases for both binary-soak and binary-titrated MsuD were solved by Fourier synthesis using the completed ternary-soak MsuD model (PDB ID 7K14), and model building and refinement were performed as described for ternary-soak MsuD. In binary-soak MsuD, FMN was placed in 6/8 protein chains, whereas only the phosphate group of FMN could be placed in chains F and H. In binary-titrated MsuD, FMN

was placed in 6/8 protein chains, and only the phosphate group of FMN could be placed in chains B and F. Succinate was refined in chains that contained an ordered C terminus interaction (chain E in binary-soak and chains A, C, and E in binary-titrated MsuD).

Within the series of liganded structures, placed FMN and MS⁻ molecules refined with B factors comparable with interacting protein residues. In molecules containing lower occupancy values, electron density is strongest for the flavin isoalloxazine ring and phosphate moieties, with weaker electron density for the ribityl moiety. In certain chains of the soaked structures the only density available for the FMN was about the phosphate moiety; therefore, a phosphate ion was placed instead. In ternary-soak MsuD a chloride ion refined best within two chains instead of MS⁻. At lower concentrations of FMN and in the absence of MS⁻, binary-titrated MsuD had appropriate electron density for succinate from the crystallization conditions within three chains where MS⁻ would have been bound. Table S3 summarizes the refined ligand occupancy information for each structure.

All models were verified with composite omit electron density maps calculated in Phenix (64) during refinement, and Ramachandran angles were calculated using MolProbity (68). Solved structures of MsuD showed two major regions of disorder in some of the protein chains, from around L246-D280 and A355-A381 at the end of the C terminus. The missing electron density for these regions was apparent in all chains in unliganded MsuD, and partially for the binary-soak and ternary-soak MsuD. Both regions were able to be modeled in the ternary-MsuD cocrystal. Unliganded MsuD is missing residues in all monomers (A:248–281 and 295–299, B:248–281 and 296–299, C:251–281 and 297–299, D:254–281 and 297–299, E: 249–281 and 295–299, F: 251–281 and 297–299, G: 254–281 and 297–299, H: 251–281 and 296–299) with 355 to 381 missing in all chains and Ramachandran favored, allowed and outliers at 98%, 2%, and 0%, respectively. Binary-soak MsuD shows similar missing residues in about half of its monomers (A:357–381, B:255–279 and 357–381, C:356–381, D: 255–280 and 357–381, E:356–381, F: 256–281 and 356–381, G: 254–280 and 356–381, H:263–280, 297–298, 355–360 and 374–381) with Ramachandran favored, allowed, and outliers at 97.05%, 2.87%, and 0.07%, respectively. Binary-soak titrate MsuD shows similar missing residues in about half of its monomers (A:357–381, B:248–280 and 378–381, C:356–381, D: 248–280 and 378–381, E: 249–281 and 356–381, F: 256–281 and 356–381, G: 251–280 and 356–381, H:251–280 and 378–381) with Ramachandran favored, allowed, and outliers at 97.16%, 2.84%, and 0.07%, respectively. The ternary soak shows similar missing residues in about four of eight of its monomers (A, C, D, E: 356–381, B: 278–279 and 356–381, F: 249–280 and 356–381, G: 254–261, 278–279 and 356–381, H: 250–271 and 375–381) with Ramachandran favored, allowed, and outliers at 98.84%, 1.16%, and 0%, respectively. The ternary cocrystal only shows disorder at the end of the C terminus in its four monomers (378–381 within all chains) with Ramachandran favored, allowed, and outliers at 96.88%, 3.12%, and 0%, respectively. Three *cis*-peptides are

Table 2
Crystallography X-ray diffraction data collection and refinement statistics table

Parameters and Statistics	Monoclinic Unliganded MsuD	Monoclinic Binary-MsuD <i>titrated soak</i>	Monoclinic Binary-MsuD <i>soak</i>	Monoclinic Ternary-MsuD <i>soak</i>	Hexagonal Ternary-MsuD <i>cocrystal</i>
Data Processing					
APS Beamline	24-ID-E	24-ID-E	24-ID-C	24-ID-C	24-ID-C
Wavelength (Å)	0.9792	0.9792	0.9795	0.9795	0.9795
Resolution range (Å)	64.79–2.80 (2.87–2.80) ^a	65.03–2.80 (2.87–2.80)	83.16–2.76 (2.83–2.76)	82.21–2.75 (2.82–2.75)	46.24–2.39 (2.450–2.39)
Space group	<i>P</i> 1 2 ₁ 1	<i>P</i> 1 2 ₁ 1	<i>P</i> 1 2 ₁ 1	<i>P</i> 1 2 ₁ 1	<i>P</i> 6 ₁
a, b, c (Å)	94.65 210.05 94.76	94.01 212.01 94.46	94.2 214.38 95.2	94.00 211.93 94.35	92.47 92.47 320.54
α, β, γ (°)	90 119.57 90	90 118.85 90	90 119.13 90	90 119.05 90	90 90 120
Total reflections	353,328 (26,137)	483,126 (40,854)	515,608 (52,841)	482,748 (35,316)	779,854 (72,207)
Unique reflections	75,178 (5636)	78,079 (5561)	83,520 (5925)	82,800 (5987)	60,776 (4265)
Multiplicity	4.7 (3.0)	6.2 (5.3)	6.2 (6.2)	5.8 (5.9)	12.8 (12.5)
Completeness (%)	94.90 (96.20)	97.90 (94.90)	98.10 (94.10)	(99.10) (97.20)	99.60 (94.50)
Mean I/σ(I)	12.33 (1.60)	8.54 (2.51)	10.81 (1.830)	11.7 (1.77)	16.01 (1.36)
<i>R</i> _{sym} (%) ^b	0.084 (0.943)	0.127 (0.613)	0.1057 (0.969)	0.083 (0.985)	0.120 (1.341)
<i>R</i> _{meas} (%) ^c	0.095 (1.008)	0.138 (0.707)	0.116 (1.078)	0.091 (0.987)	0.125 (1.415)
<i>R</i> _{pim} (%) ^d	0.042 (0.511)	0.054 (0.286)	0.046 (0.418)	0.037 (0.397)	0.0348 (0.396)
CC1/2 (%) ^e	0.999 (0.606)	0.989 (0.867)	0.997 (0.682)	0.998 (0.700)	0.999 (0.764)
Refinement					
Unique reflections	75,153	78,024	83,498	82,785	60,658
<i>R</i> _{work}	0.1803	0.1993	0.1798	0.1715	0.1836
<i>R</i> _{free}	0.2212	0.2381	0.2240	0.2144	0.2208
Protein atoms	19,795	21,296	21,129	21,512	11,736
Ligands atoms	0	234	196	249	145
Solvent atoms	0	41	49	16	135
RMS(bonds) (Å)	0.003	0.005	0.002	0.006	0.002
RMS(angles) (°)	0.61	0.98	0.56	0.76	0.43
Average B-factor (Å ²)					
All atoms	85.3	74.8	77.4	91.6	57.4
Protein	85.3	74.8	77.3	91.7	57.4
Ligand and ions	–	80.6	88.5	91.7	55.3
Waters	–	53.0	78.1	67.5	52.2
Ramachandran					
Favored (%)	97.73	97.16	97.05	98.84	96.81
Allowed (%)	2.27	2.84	2.87	1.16	3.19
Outliers (%)	0.00	0.00	0.07	0.00	0.00
Rotamer outliers (%)	1.78	1.93	2.62	1.29	1.33

^a Highest resolution shell is shown in parentheses.

^b $R_{sym/merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$.

^c $R_{meas} = \frac{\sum_{hkl} [N/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$.

^d $R_{pim} = \frac{\sum_{hkl} [1/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$ where $I_i(hkl)$, $\langle I(hkl) \rangle$, and N represent the intensity measurement, the mean intensity, and the redundancy for reflection hkl , respectively.

^e $CC^* = [2CC_{1/2}/(1 + CC_{1/2})]^{1/2}$ where $CC_{1/2}$ is the correlation between two random halves of the datasets, each containing half of the measured intensities for each unique reflection, and CC^* is an approximation of the correlation coefficient for a noise-free dataset.

^f $R_{work} = \frac{\sum |F_{obs}(hkl) - F_{calc}(hkl)|}{\sum |F_{obs}(hkl)|}$, where $F_{obs}(hkl)$ and $F_{calc}(hkl)$ are the observed and calculated structure factor amplitudes of 95% of the reflections used for refinement. R_{free} was calculated from the 5% of total reflections that were omitted from the refinement.

Structural and mechanistic insights from ligand-bound MsuD

observed in the electron density: K167-P168, Y330-P331, and E349-P350. The region from approximately 250 to 280 missing in most models has higher B-factor values compared with the rest of the chain owing to the inherent flexibility of the region. Similarly, the C terminus from approximately 355 to 381 missing in most models is inherently more disordered compared with the rest of the model, but with ternary and binary complexes showing weak but clear electron density for the C terminus. Placements of ligands were verified using Polder omit maps calculated in Phenix (69). Final model validation was performed using Phenix and the RCSB validation server (www.rcsb.org). The FMN parameter file was generated using Phenix eLBOW with ideal FMN coordinates from the PDB, and ring plane parameters were generated using the Grade Web Server to enable the isoalloxazine moiety to bend slight, as is best defined by the electron density within MsuD. The MS⁻ parameter file was generated using the Grade Web Server (70) with a mol2 file of PDB ligand 03S prepared in Bioluminate 2020-3 (Schrödinger, LLC) (71–73).

Bioinformatic analysis

A DELTA-BLAST of the *P. fluorescens* MsuD sequence was done at a cutoff of 40% to 60% identity, resulting in over 5000 sequences. An alignment was generated with ClustalOmega (74), and both ends of the alignment were trimmed and gaps removed. The residue frequency in the regions of the sulfonate-binding loop were analyzed, and images were generated with WebLogo (75). Genome alignments were performed using Mauve (76) within Geneious Prime 2020.2 created by Biomatters (<https://www.geneious.com>).

Molecular docking

Rigid receptor molecular docking was performed in Bioluminate 2020-3 (Schrödinger, LLC). The ternary-MsuD tetramer and individual ligands were prepared for docking using the Protein Preparation tool (73), and docking was carried out in GLIDE (77). Ligands were prepared using Lig-Prep with states generated at pH 7.0, and the receptor grid was defined with a 15 × 15 × 15 Å³ inner box and a 35 × 35 × 35 Å³ outer box defined about the active site of MsuD chain A. The ligands explored include MS⁻, Pipes, Hepes, Mops, pentanesulfonate, and octanesulfonate.

Protein figures were generated using the software PyMOL (78), kinetics analyses and graphs were performed using OriginPro2020, and chemical mechanisms were constructed within ChemDraw 19.0 (PerkinElmer Informatics).

Data availability

The datasets for atomic coordinates, structure factors, crystal structure, and diffraction data generated during this study have been deposited in the RCSB Protein Data Bank and are available at www.rcsb.org, accession code 7JV3 for MsuD, 7K64 for binary-titrated MsuD, 7JYB for binary-soak MsuD, 7K14 for ternary-soak MsuD, and 7JW9 for ternary-MsuD cocrystal. The plasmid generated in this study has been

deposited to GenBank accession no. ABA75653.1. All remaining data are contained within the article.

Supporting information—This article contains [supporting information](#) (17, 26, 30, 31, 35, 37, 42, 43, 45, 46, 48, 81–87).

Acknowledgments—The authors kindly acknowledge Jess Soule and Reyaz Gonzalez for helpful discussions, the MIT Structural Biology Core facility, and Staff Scientist Robert Grant. This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the NIH (P30 GM124165). The Eiger 16M detector on 24-ID-E is funded by an NIH-ORIP HEI grant (S10OD021527). This research used resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

Author contributions—J. J. M. L., I. M. E., S. V. N., D. K. W., and D. P. D. data curation; J. J. M. L. and D. P. D. software; J. J. M. L. and D. P. D. formal analysis; J. J. M. L., I. M. E., S. V. N., D. K. W., and D. P. D. validation; J. J. M. L., I. M. E., S. V. N., and D. K. W. investigation; J. J. M. L. and D. P. D. writing—original draft; J. J. M. L., D. K. W. and D. P. D. conceptualization; D. K. W. and D. P. D. funding acquisition; D. K. W. and D. P. D. writing—review and editing; D. P. D. resources.

Funding and additional information—This work was supported by the National Science Foundation, USA (1807480 to D. K. W. and D. P. D.). Opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect the views of the National Science Foundation. I. M. E. was supported by a National Institutes of Health (NIH) IMSD grant award (2R25GM076321) and Undergraduate Research Funds. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: DMS, dimethylsulfide; DMSO₂, dimethyl sulfone; DTNB, 5,5-dithio-bis-(2-nitrobenzoic acid); FADH⁻, reduced flavin adenine dinucleotide; FMN, oxidized flavin mononucleotide; FMNH⁻, reduced flavin mononucleotide; IS, insertion segment; MR, molecular replacement; MS⁻, methanesulfonate; MSI⁻, methanesulfinate; TIM, triose isomerase phosphate.

References

- Huijbers, M. M., Montersino, S., Westphal, A. H., Tischler, D., and van Berkel, W. J. (2014) Flavin dependent monooxygenases. *Arch. Biochem. Biophys.* **544**, 2–17
- van Berkel, W. J. H., Kamerbeek, N. M., and Fraaije, M. W. (2006) Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *J. Biotechnol.* **124**, 670–689
- Walsh, C. T., and Wencewicz, T. A. (2013) Flavoenzymes: Versatile catalysts in biosynthetic pathways. *Nat. Prod. Rep.* **30**, 175–200
- de Gonzalo, G., Mihovilovic, M. D., and Fraaije, M. W. (2010) Recent developments in the application of Baeyer-Villiger monooxygenases as biocatalysts. *ChemBiochem* **11**, 2208–2231

5. Kelly, D. P., and Smith, N. A. (1990) Organic sulfur compounds in the environment. Biogeochemistry, microbiology, and ecological aspects. In: Marshall, K. C., ed. *Advances in Microbial Ecology*, Plenum Press, New York, NY: 345–385
6. Schafer, F. Q., and Buettner, G. R. (2001) Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. *Free Radic. Biol. Med.* **30**, 1191–1212
7. Braymer, J. J., and Lill, R. (2017) Iron-sulfur cluster biogenesis and trafficking in mitochondria. *J. Biol. Chem.* **292**, 12754–12763
8. Kertesz, M. A. (1999) Riding the sulfur cycle - metabolism of sulfonates and sulfate esters in Gram-negative bacteria. *FEMS Microbiol. Rev.* **24**, 135–175
9. Eichhorn, E., van der Ploeg, J. R., Kertesz, M. A., and Leisinger, T. (1997) Characterization of α -ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. *J. Biol. Chem.* **272**, 23031–23036
10. Quadroni, M., Staudenmann, W., Kertesz, M., and James, P. (1996) Analysis of global responses by protein and peptide fingerprinting of proteins isolated by two-dimensional gel electrophoresis. Application to the sulfate-starvation response of *Escherichia coli*. *Eur. J. Biochem.* **239**, 773–781
11. Kertesz, M. A., Leisinger, T., and Cook, A. M. (1993) Proteins induced by sulfate limitation in *Escherichia coli*, *Pseudomonas putida*, or *Staphylococcus aureus*. *J. Bacteriol.* **175**, 1187–1190
12. Gallagher, J. R., Olson, E. S., and Stanley, D. C. (1993) Microbial desulfurization of dibenzothiophene: A sulfur-specific pathway. *FEMS Microbiol. Lett.* **107**, 31–35
13. Thyse, G. J., and Wanders, T. H. (1972) Degradation of n-alkane-1-sulfonates by *Pseudomonas*. *Antonie Van Leeuwenhoek* **38**, 53–63
14. Schafer, H., Myronova, N., and Boden, R. (2010) Microbial degradation of dimethylsulphide and related C1-sulphur compounds: Organisms and pathways controlling fluxes of sulphur in the biosphere. *J. Exp. Bot.* **61**, 315–334
15. Omori, T., Saiki, Y., Kasuga, K., and Kodama, T. (1995) Desulfurization of alkyl and aromatic sulfides and sulfonates by dibenzothiophene-desulfurizing *Rhodococcus* sp. strain SY1. *Biosci. Biotechnol. Biochem.* **59**, 1195–1198
16. Endoh, T., Habe, H., Yoshida, T., Nojiri, H., and Omori, T. (2003) A CysB-regulated and σ^{54} -dependent regulator, *SfnR*, is essential for dimethyl sulfone metabolism of *Pseudomonas putida* strain DS1. *Microbiology (Reading)* **149**, 991–1000
17. Kertesz, M. A., Schmidt-Larbig, K., and Wuest, T. (1999) A novel reduced flavin mononucleotide-dependent methanesulfonate sulfonate encoded by the sulfur-regulated *msu* operon of *Pseudomonas aeruginosa*. *J. Bacteriol.* **181**, 1464–1473
18. Bentley, R., and Chasteen, T. G. (2004) Environmental VOSCs—formation and degradation of dimethyl sulfide, methanethiol and related materials. *Chemosphere* **55**, 291–317
19. Simó, R., and Pedrós-Alió, C. (1999) Short-term variability in the open ocean cycle of dimethylsulfide. *Global Biogeochem. Cycles* **13**, 1173–1181
20. Barnes, I., Hjorth, J., and Mihalopoulos, N. (2006) Dimethyl sulfide and dimethyl sulfoxide and their oxidation in the atmosphere. *Chem. Rev.* **106**, 940–975
21. Autry, A. R., and Fitzgerald, J. W. (1990) Sulfonate S: A major form of forest soil organic sulfur. *Biol. Fertil. Soils* **10**, 50–56
22. Kelly, D. P., Baker, S. C., Trickett, J., Davey, M., and Murrell, J. C. (1994) Methanesulphonate utilization by a novel methylotrophic bacterium involves an unusual monooxygenase. *Microbiology* **140**, 1419–1426
23. Borodina, E., Kelly, D. P., Rainey, F. A., Ward-Rainey, N. L., and Wood, A. P. (2000) Dimethylsulfone as a growth substrate for novel methylotrophic species of *Hyphomicrobium* and *Arthrobacter*. *Arch. Microbiol.* **173**, 425–437
24. Higgins, T. P., De Marco, P., and Murrell, J. C. (1997) Purification and molecular characterization of the electron transfer protein of methanesulfonic acid monooxygenase. *J. Bacteriol.* **179**, 1974–1979
25. Higgins, T. P., Davey, M., Trickett, J., Kelly, D. P., and Murrell, J. C. (1996) Metabolism of methanesulfonic acid involves a multicomponent monooxygenase enzyme. *Microbiology (Reading)* **142**, 251–260
26. Eichhorn, E., van der Ploeg, J. R., and Leisinger, T. (1999) Characterization of a two-component alkanesulfonate monooxygenase from *Escherichia coli*. *J. Biol. Chem.* **274**, 26639–26646
27. Wicht, D. K. (2016) The reduced flavin-dependent monooxygenase *SfnG* converts dimethylsulfone to methanesulfinate. *Arch. Biochem. Biophys.* **604**, 159–166
28. Soule, J., Gnann, A. D., Gonzalez, R., Parker, M. J., McKenna, K. C., Nguyen, S. V., Phan, N. T., Wicht, D. K., and Dowling, D. P. (2020) Structure and function of the two-component flavin-dependent methanesulfinate monooxygenase within bacterial sulfur assimilation. *Biochem. Biophys. Res. Commun.* **522**, 107–112
29. Endoh, T., Habe, H., Nojiri, H., Yamane, H., and Omori, T. (2005) The σ^{54} -dependent transcriptional activator *SfnR* regulates the expression of the *Pseudomonas putida sfnFG* operon responsible for dimethyl sulphone utilization. *Mol. Microbiol.* **55**, 897–911
30. Cao, H. Y., Wang, P., Peng, M., Shao, X., Chen, X. L., and Li, C. Y. (2018) Crystal structure of the dimethylsulfide monooxygenase DmoA from *Hyphomicrobium sulfonivorans*. *Acta Crystallogr. F Struct. Biol. Commun.* **74**, 781–786
31. Bhandari, D. (2017) *Studies on the Flavin-Mediated Cysteine Salvage Pathway, Structural Enzymology of Riboflavin Lyase and Mechanistic Investigation of the Radical S-Adenosyl-L-Methionine-Mediated Tryptophan Lyase*. Ph.D. thesis, Texas A&M University
32. Okai, M., Lee, W. C., Guan, L. J., Ohshiro, T., Izumi, Y., and Tanokura, M. (2017) Crystal structure of dibenzothiophene sulfone monooxygenase BdsA from *Bacillus subtilis* Wu-S2B. *Proteins* **85**, 1171–1177
33. Eichhorn, E., Davey, C. A., Sargent, D. F., Leisinger, T., and Richmond, T. J. (2002) Crystal structure of *Escherichia coli* alkanesulfonate monooxygenase SsuD. *J. Mol. Biol.* **324**, 457–468
34. Saleem-Batcha, R., Stull, F., Sanders, J. N., Moore, B. S., Palfey, B. A., Houk, K. N., and Teufel, R. (2018) Enzymatic control of dioxygen binding and functionalization of the flavin cofactor. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 4909–4914
35. Matthews, A., Saleem-Batcha, R., Sanders, J. N., Stull, F., Houk, K. N., and Teufel, R. (2020) Aminoperoxide adducts expand the catalytic repertoire of flavin monooxygenases. *Nat. Chem. Biol.* **16**, 556–563
36. Teufel, R., Miyana, A., Michaudel, Q., Stull, F., Louie, G., Noel, J. P., Baran, P. S., Palfey, B., and Moore, B. S. (2013) Flavin-mediated dual oxidation controls an enzymatic Favorskii-type rearrangement. *Nature* **503**, 552–556
37. Mukherjee, T., Zhang, Y., Abdelwahed, S., Ealick, S. E., and Begley, T. P. (2010) Catalysis of a flavoenzyme-mediated amide hydrolysis. *J. Am. Chem. Soc.* **132**, 5550–5551
38. Massey, V. (1994) Activation of molecular oxygen by flavins and flavo-proteins. *J. Biol. Chem.* **269**, 22459–22462
39. Gibson, Q. H., and Hastings, J. W. (1962) The oxidation of reduced flavin mononucleotide by molecular oxygen. *Biochem. J.* **83**, 368–377
40. Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., Wilson, I. A., Corran, P. H., Furth, A. J., Milman, J. D., Offord, R. E., Priddle, J. D., and Waley, S. G. (1975) Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 Å resolution using amino acid sequence data. *Nature* **255**, 609–614
41. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797
42. Zhang, R., Skarina, T., Savchenko, A., Edwards, A., Joachimiak, A., and Midwest Center for Structural Genomics (MCSG) (2003) *Structural Genomics, Crystal Structure of Alkanesulfonate Monooxygenase*, protein databank entry. <https://doi.org/10.2210/pdb1N9K/pdb>
43. Holm, L. (2020) DALI and the persistence of protein shape. *Protein Sci.* **29**, 128–140
44. Carpenter, R. A., Xiong, J., Robbins, J. M., and Ellis, H. R. (2011) Functional role of a conserved arginine residue located on a mobile loop of alkanesulfonate monooxygenase. *Biochemistry* **50**, 6469–6477
45. Su, T., Su, J., Liu, S., Zhang, C., He, J., Huang, Y., Xu, S., and Gu, L. (2018) Structural and biochemical characterization of BdsA from *Bacillus subtilis* Wu-S2B, a key enzyme in the “4S” desulfurization pathway. *Front. Microbiol.* **9**, 231
46. Li, L., Liu, X., Yang, W., Xu, F., Wang, W., Feng, L., Bartlam, M., Wang, L., and Rao, Z. (2008) Crystal structure of long-chain alkane monooxygenase (LadA) in complex with coenzyme FMN: Unveiling the long-chain alkane hydroxylase. *J. Mol. Biol.* **376**, 453–465

Structural and mechanistic insights from ligand-bound MsuD

47. Robbins, J. M., and Ellis, H. R. (2012) Identification of critical steps governing the two-component alkanesulfonate monooxygenase catalytic mechanism. *Biochemistry* **51**, 6378–6387
48. Jun, S. Y., Lewis, K. M., Youn, B., Xun, L., and Kang, C. (2016) Structural and biochemical characterization of EDTA monooxygenase and its physical interaction with a partner flavin reductase. *Mol. Microbiol.* **100**, 989–1003
49. Ellis, H. R. (2011) Mechanism for sulfur acquisition by the alkanesulfonate monooxygenase system. *Bioorg. Chem.* **39**, 178–184
50. Robbins, J. M., and Ellis, H. R. (2014) Steady-state kinetic isotope effects support a complex role of Arg226 in the proposed desulfonation mechanism of alkanesulfonate monooxygenase. *Biochemistry* **53**, 161–168
51. Xiong, J., and Ellis, H. R. (2012) Deletional studies to investigate the functional role of a dynamic loop region of alkanesulfonate monooxygenase. *Biochim. Biophys. Acta* **1824**, 898–906
52. Adak, S., and Begley, T. P. (2016) Dibenzo thiophene catabolism proceeds via a flavin-N5-oxide intermediate. *J. Am. Chem. Soc.* **138**, 6424–6426
53. Adak, S., and Begley, T. P. (2017) RutA-catalyzed oxidative cleavage of the uracil amide involves formation of a flavin-N5-oxide. *Biochemistry* **56**, 3708–3709
54. Adak, S., and Begley, T. P. (2019) Hexachlorobenzene catabolism involves a nucleophilic aromatic substitution and flavin-N5-oxide formation. *Biochemistry* **58**, 1181–1183
55. Thyse, G. J., and Wanders, T. H. (1974) Initial steps in the degradation of *n*-alkane-1-sulphonates by *Pseudomonas*. *Antonie Van Leeuwenhoek* **40**, 25–37
56. Chenprakhon, P., Trisrivirat, D., Thotsaporn, K., Sucharitakul, J., and Chaiyen, P. (2014) Control of C4a-hydroperoxyflavin protonation in the oxygenase component of *p*-hydroxyphenylacetate-3-hydroxylase. *Biochemistry* **53**, 4084–4086
57. Draper, R. D., and Ingraham, L. L. (1968) A potentiometric study of the flavin semiquinone equilibrium. *Arch. Biochem. Biophys.* **125**, 802–808
58. Maniatis, T., Fritsch, E. F., and Sambrooke, J. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York, NY
59. Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675
60. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. E., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) *Protein Identification and Analysis Tools on the ExPASy Server*. Humana Press, Totowa, NJ: 571–607
61. Robbins, J. M., and Ellis, H. R. (2019) Investigations of two-component flavin-dependent monooxygenase systems. *Methods Enzymol.* **620**, 399–422
62. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132
63. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
64. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkoczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., et al. (2019) Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877
65. Stein, N. (2008) CHAINSAW: A program for mutating pdb files used as templates in molecular replacement. *J. Appl. Crystallogr.* **41**, 641–643
66. Brunger, A. T. (2007) Version 1.2 of the crystallography and NMR system. *Nat. Protoc.* **2**, 2728–2733
67. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
68. Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., 3rd, Snoeyink, J., Adams, P. D., et al. (2018) MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315
69. Liebschner, D., Afonine, P. V., Moriarty, N. W., Poon, B. K., Sobolev, O. V., Terwilliger, T. C., and Adams, P. D. (2017) Polder maps: Improving OMIT maps by excluding bulk solvent. *Acta Crystallogr. D Struct. Biol.* **73**, 148–157
70. Smart, O. S., Womack, T. O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C., and Bricogne, G. (2011) *Grade v.1.2.9*
71. Greenwood, J. R., Calkins, D., Sullivan, A. P., and Shelley, J. C. (2010) Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *J. Comput. Aided Mol. Des.* **24**, 591–604
72. Shelley, J. C., Cholleti, A., Frye, L. L., Greenwood, J. R., Timlin, M. R., and Uchimaya, M. (2007) Epik: A software program for pK_a prediction and protonation state generation for drug-like molecules. *J. Comput. Aided Mol. Des.* **21**, 681–691
73. Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R., and Sherman, W. (2013) Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* **27**, 221–234
74. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539
75. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: A sequence logo generator. *Genome Res.* **14**, 1188–1190
76. Darling, A. C., Mau, B., Blattner, F. R., and Perna, N. T. (2004) Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* **14**, 1394–1403
77. Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., Sanschagrin, P. C., and Mainz, D. T. (2006) Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **49**, 6177–6196
78. Schrodinger, LLC (2015) *The PyMOL Molecular Graphics System, Version 1.8*
79. Habe, H., Kouzuma, A., Endoh, T., Omori, T., Yamane, H., and Nojiri, H. (2007) Transcriptional regulation of the sulfate-starvation-induced gene *sfmA* by a σ^{54} -dependent activator of *Pseudomonas putida*. *Microbiology* **153**, 3091–3098
80. Lundgren, B. R., Sarwar, Z., Feldman, K. S., Shoytush, J. M., and Nomura, C. T. (2019) SfnR2 regulates dimethyl sulfide-related utilization in *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* **201**, e00606-00618
81. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324
82. Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model* **51**, 2778–2786
83. Clifton, M. C., Edwards, T. E., and Seattle Structural Genomics Center for Infectious Disease (SSGCID) (2011) Structure of a nitrilotriacetate monooxygenase from *Burkholderia pseudomallei*, Protein data bank. <https://doi.org/10.2210/pdb3SDO/pdb>
84. Fisher, A. J., Thompson, T. B., Thoden, J. B., Baldwin, T. O., and Rayment, I. (1996) The 1.5-Å resolution crystal structure of bacterial luciferase in low salt conditions. *J. Biol. Chem.* **271**, 21956–21968
85. Campbell, Z. T., Weichsel, A., Montfort, W. R., and Baldwin, T. O. (2009) Crystal structure of the bacterial luciferase/flavin complex provides insight into the function of the β subunit. *Biochemistry* **48**, 6085–6094
86. Nguyen, Q. T., Trinco, G., Binda, C., Mattevi, A., and Fraaije, M. W. (2017) Discovery and characterization of an F₄₂₀-dependent glucose-6-phosphate dehydrogenase (Rh-FGD1) from *Rhodococcus jostii* RHA1. *Appl. Microbiol. Biotechnol.* **101**, 2831–2842
87. Domagalski, M. J., Chruszcz, M., Xu, X., Cui, H., Chin, S., Savchenko, A., Edwards, A., Joachimiak, A., Minor, W., and Midwest Center for Structural Genomics (MCSG) (2011) Crystal structure of the luciferase-like monooxygenase from *Bacillus cereus* ATCC 10987. Protein Data Bank. <https://doi.org/10.2210/pdb3RAO/pdb>