Supporting Information for

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

Jeremy J.M. Liew¹, Israa M. El Saudi¹, Son V. Nguyen^{2,3}, Denyce K. Wicht², and Daniel P. Dowling^{1,*}

¹ Department of Chemistry, University of Massachusetts Boston, Boston MA 02125, USA

² Department of Chemistry and Biochemistry, Suffolk University, Boston MA 02108, USA

³ Current address: Department of Chemistry, University of New Hampshire, Durham NH 03824, USA

Correspondence to Daniel P. Dowling: <u>daniel.dowling@umb.edu</u> (D.P.D.)

Supporting Information Table of Contents

1. Supporting Figures:

Figure S1. Related to Figure 2. The results of analytical size exclusion chromatography of MsuD

Figure S2. Related to Figure 3. The results of a sequence alignment of MsuD structural homologs with a minimum sequence identity cut-off of 19%

Figure S3. Related to Figure 3. The results of comparison of class C flavin-dependent monooxygenases that crystallize as tetramers

Figure S4. Related to Figure 4. The results of representative polder omit electron density maps for structures of MsuD demonstrating the presence of the active site lid, protein C-terminus, and the ligands FMN, MS⁻, and succinate

Figure S5. Related to Figure 5. The results of analysis of residues and surrounding environment of (A) FMN and (B) MS⁻ in chain A of ternary-MsuD

Figure S6. Related to Figure 5. The results of (A) overlay of analogues of W195 from MsuD homologs and (B) overlay of FMN from different homolog structures demonstrating differences in ribityl tail positioning

Figure S7. Related to Figure 5. The purified protein samples for wild type MsuD and the W195A, R225A, R296A, and C-terminal truncation MsuD variants.

Figure S8. Related to Figure 6. The results of structural superimpositions of MsuD and homologs' (A) lid and (B) C-terminal regions

Figure S9. Related to Figure 7. The reactions of RutA, DszA, and HcbA1

2. Supporting Tables:

Table S1. Related to Figure 1. Genome and gene information for msu and sfn genes

Table S2. Related to Figures 3 and S2. Structural alignment results for MsuD from the DALI server, containing matches above or equal to 19% sequence identity

Table S3. Related to Figure 4. Summary of ligand occupancies and lid regions within MsuD crystal structures

Table S4. Related to Figure 5 and S6B. Ribityl tail torsion angles from different FMN-bound structures of the structural homologs of MsuD

Table S5. Related to Table 1. Relative activity of MsuD from *P. fluorescens* with different substrates compared to relative activity trends reported for MsuD from *P. aeruginosa* and SsuD in *E. coli*

Table S6. Related to Methods. Forward and reverse DNA primers used in PCR amplification of *msuD* of *P. fluorescens* Pf0-1 and in site-directed mutagenesis



Figure S1. MsuD behaves as a tetramer in solution. (A) Size exclusion chromatography of MsuD on the ENrich SEC 650 column compared to Cytiva high-molecular weight protein standards reveals MsuD travels as an ~200 kDa species, most in line with a tetramer (monomer chain MW: 44.2 kDa, and homotetramer MW: 176.9 kDa). Molecular weight standards include: Ovalbumin (MW: 44 kDa), Conalbumin (MW: 75 kDa), Aldolase (MW: 158 kDa), Ferritin (MW: 440 kDa) and Thyroglobulin (MW: 669 kDa). (B) An overlay of wild type MsuD and the C-terminal truncation mutant (MsuD ^{ΔC-16}) reveal the C-terminal tail is critical for tetramerization. MsuD^{ΔC-16} behaves as an ~113 kDa species, which corresponds to a dimer.

MsuD			la β1b	α1		۷
MsuD SsuD RutA 3RAO DmoA LadA RcaE BdsA CmoJ	MQDAAPRLTFTLRDEER MKKRIVL MKKRIVL MKGNTMKQLRF MAEQRQLHLAGFF MTRADFI	. MDVFWFLPTHGD MSINMFWFLPTHGD LMMKIGVFVPIGNN MEYGFWLPIFG NAFDMTCVSHQSAG NAFEMNCVGHIAHG GLFENAQTNDSGTA QFGAMIHGVGGTTD	Z GHYLGTEEGSRP GHYLGTEEGSRP GWLISTHAPQYM GWLRNVNDESMP TWRHPSSQAARY LWRHPENQRHRY TWRHPENQRHRY TWRHPDNQRHLF AWRHVGATNGFL GWRHPDVDPSAS	JU VTLNYLKQVAQAA VDHGYLQQIAQAA PTFELNKAIVQKA NDLEYWTNMAMEL IDLNYWTELAQLL DTLDYWRNIAQIC I.GEFYKKQIARTL TNIEFYMKKAQTA	40 DSLGYHGVLTPTGRSC DRLGYTGVLTPTGRSC EHYHFDFALSMIKLRC EQLGFSTTLIAELNLN ERGCFDCLFIADVVGY EKGKFDALFLADVVG EDAGLDFVFLADAWGY ERGKFDLLFLPDGLAJ EKGLFSFIFIADGLFJ	EDS FGCKTEFWDHN DIKGVSAP YDVYRGSAEMALR YDVYRQSRDTAVR ADVNGERPDICDV EDSYGDNLETGVG SEKSIPHFLN
MsuD	α2	β3	TT 000000	x3	β4 α4a	0000.0
MsuD SsuD RutA 3RAO DmoA LadA RcaE BdsA CmoJ	GO GO WVIASA LESFILMAG SLEAWTTAAA DADQVP.VNDPFGAISA EAVQIP.VNDPFGAISA EAVQIP.VNDPFGAISA EAVQIP.VNDPFGAISA EAVQIP.VNDPFAIVAAA LGGQGAVALEPTSVIAT RFEPITILSA	70 LVPI TERIRYLVA I MIPVTORLKFLVAL LAAVTSRIQIYATA LAAVTDRLEIMTAV MAAVTEHVGFGVTA MAYVTKHLAFAVTF LIASTKKLGLVMTG MAAVTORLGLGATV LASVTKNIGLVGTF	* 80 RPGIISETVSAR RPSVTSPTVAAR ATLTLPPAIVAR RPGFHNPAVTAK AITFEOPYLLAR STTYEHPYGHAR STLLEOPYSFAR STTYYPPYHVAR STSFTEPFIISR	100 MAATLDRLSNGRL DAATLDRLSNGRA MAATIDSISGGRF MAANIDQLSNGRF RLSTLDHLTKGRV RMSTLDHLTKGRV RMSTLDHLTKGRI VFATLDNLSDGRI DLMSLDHISGGRA	110 LINVVTGGDPDENRGI LFNLVTGSDPQELAGI GVNLVTGVQKPEYEQN TLNVVSAWWEEEAKQY AWNVVSSYLNSAALNI AWNVVTSLNSAALNI GWNVVTGTAETASAA SWNVVTSLNDSEARNI GWNLVTSLNDSEARNI	120 GS. FLSHSER.Y GV.FLDHSER.Y GC.VFTAHDER.Y GG.VFTAHDER.Y GMDQQLAHDER.Y FGVPMVAHDDR.Y GVDEHLEHDIR.Y GVDEHLEHDIR.Y SKSNLPEHTER.Y
MsuD	α4 <u>0000000000000</u> 130 140		β4a	β4b 50 160	170 ^{β5}	α5 <u>000000000</u> 180 190
MsuD SsuD RutA 3RAO DmoA LadA RcaE BdsA CmoJ	EVTDEFLKIWRRVIQ. EASAEFTQVWRRLIQ. DYLTEYVQVLRDLWG. DRTEEFVTILKGLWK. EMADEYMEVMYKLWEGS DLADEYLEVCYKLWEGS DMADDFMELVYKLWEGA DRADEFLEAVKKLWS. EIAQEHLDVVRGLWN.S	WEDDAVKRDKKSGV WEDDAVIRDIENNI WEPDALERD.KQGR WSEDALLLDKVGGR WEHDAFIHNKKTGQ	GEAVDF RETVDF GESSDF EEFSY FTDGSKVHPINH YTDPSKVHEINH YADPAKVHRIDH FADPKKVQYVNH FFDQAKLHRINH	ECKHLKVQNAKAL NGKHIHVRGAKLL KGDFFTMNDCRVS QGKYYKVPGFHIC SCKYFEVPGPHIC EGPYFRSNGYGNT KGRWLSVRGPLQV KGKYFQVEGPLNI	YPPVQKPYPPLYFGG FPAIQQPYPPLYFGG PQPSVPMKVICAG PKPVQKQGIKLYAG EPSPQR.TPVIFQAG SYSPQG.TPVIFQAG SYSPQG.TPVIFQAG SRSKQG.EPVIFQAG	SDAAHD LAAE QVD SDVAQE LAAE QVD SDAGMAF SARYAD SKRGKEVIVNHAD SGRGSKFAASNAE SERGREFAAKHAE SERGROFGGRHGE SPRGRRFAGRWAE SETGROFAAKNAD
		_	-			
MsuD		<u> </u>	β7 230	α7a 222200000000 240	α7b <u>000000000000</u> 250 _ 260	α ⁷ c 2000000000 270
MsuD SsuD RutA 3RAO DmoA LadA RcaE BdsA CmoJ	VYLTWGEPPAAVAEKLA LYLTWGEPPAUVAEKLE FNFCFGKGVNTPTAFAP AYVMHGGTVEEVSVKIE GMFILTTSVEQARQITT CVFLGGKDVETLKFFVD CIFLGGAPIPKLAEQVR AVFSVSPNLDIMRAVYQ AIFTHSNSLEETKAFYA	DVRERAARHGRKVK QVRAKAAAHGRKIR TAARMKQAAEQTGR DMKNRRKKVTEEPL DIRNQAEAAGRSRD DIRKRAKYGRNPD AIRAEAVAEGRAAD DIKAHVAAAGRDPE DVKSRAADEGRDPS	FGIRLHVIVRET FGIRLHVIVRET DVGSYVLFMVIA QSFGLAAYVICR SIKIFMLLTVIT HIKMFAGICVIV SIKLMAAFSCVI QTKVFTAVMPVL SVRIFPGISPIV.	AEEAWKAADKLIE NDEAWQAAERLIS DETDDAARAKWEH .HTEEEALEEWRR GDSDEAAEAKYQE GKTHDEAMEKLNS APTHEEAVQKYQE GETEQVARERLEY ADTEEEAEKKYRE	HISDETIEAAQKSFSI HLDDETIAKAQAAFAF YKAGADEEALSWLTE(ITDVKDDALGYAGYQI YLSYANPEGMLALYG VLSYANPEGMLALYG VLDSQTPEVAVASYA LNSLVHPEVGLSTLS FAELIPIENAVTYLAF	KFDSEGQRRMAALH TTDSVGQQRMAALH SQKDTRSGTDTNV FVSKSQLEQQVKL GWTGIDFAKLDPD GGTGYDLSKYSSN WFTGIDLSSYDPS SHSGLNLSKYPLD FFDDYDLSVYPLD
MsuD	$\begin{array}{c} \beta7a & \beta7b & \eta^{7}\\ 280 & TT & 290 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 &$	7d QQ GLVBGGS			α7 <u>2222222222</u> 10 320 000 ARTKEYADIC T	
Raub	NCKDDNIEICDNIWACV					CTUI COVDUI EEN

		200	300	319	529	339
MsuD	DGRRDNLEIAP	PNLWAGVGLVRGGS	GTALV	GDPQQV<mark>A</mark>ARIK I	EYADLG. IESF	IFSGYPHLEEA
SsuD	NGKRDNLEISP	PNLWAGVGLVRGGA	GTALV	GDGPTV <mark>A</mark> ARINI	EYAALG. <mark>ID</mark> SF	VLSGYPHLEEA
RutA	RQMADPTSAVN	NIN	MGTLV	GSYASVA RMLDI	EVASVPG <mark>AE</mark> GV	LLTFDDFLSGI
3RAO	NDYSVSNRGLE	R	PNL <mark>I</mark>	GTPEQIAERILI	AFEKVG. <mark>VT</mark> LL	LLQFSPQLEEM
DmoA	EPLQAMENDSI	LRTTLESLTHGENAKKWTVRDVIRE	RCIGGLGPVL <mark>V</mark>	GGPQKVADELEI	RWVDEGG <mark>VD</mark> GF	NLAYAVTPGSV
LadA	DYIGSISVGEI	IINNMSKLDG	KWFKLSV	GTPKKVADEMQ	YLVEEAG <mark>ID</mark> G F :	NLVQYVSPGTF
RcaE	TPMSELHTELS	SQTQVARFAGLTVGDVLADW	HAHGVRTKPV <mark>V</mark>	GTPEEV <mark>A</mark> DAIVI	ELAEGAD <mark>LD</mark> G <mark>F</mark>	LL TPVIQPGST
BdsA	TKFSDIVADLG	GDRHVPTMLQMFSAVAGGGADLTLAELGRRY	GTNVGFVPQW <mark>A</mark>	GTAEQIADQLIS	SHFEAGA <mark>AD</mark> GF	IISPAYLPGIY
CmoJ	EPFPDIGDVGK	KNAFQSTTDRIKREAKARNLTLREVAQEM	AFPRTLF <mark>I</mark>	GTPERVASLIE	FWFNAEA <mark>AD</mark> G <mark>F</mark>	IV <mark>G.SDIPGTL</mark>

	α8	α8a		α8b																										
MsuD	222222	222		2222	2																									
	340		350			360	2		3	70			3	80																
MsuD	YRFAEL	FPLI	LPEP	YASL	AGI	RGV	CNL:	FGP	FGE	MI	AND	VLI	PAK	AN	Α.															
SsuD	YRVGELI	FPLI	L D V A	IPEI	Р.	QI	? <mark>Q</mark> P:	LNP	QGE	AV	AND	FII	RK	VA	QS															
RutA	ETFGERI	[QPL]	QCR	AHLP	AL	ΓQĒ	/AG	LCG	Ř																					
3RAO	KRFSEK	/MPLV	VEAK	RKEL	IT																									
DmoA	TDFIDY	VPEI	LRKR	GRAÇ	DS	KPO	G <mark>S</mark> LI	RRK	LIG	TN	DGR	VE:	БΤН	PA	AQ	YRD	AY	VG	ĸΕ	sv	AD	RI	QE	SI	?F:	AN.	AK.	ΑP	VA	E
LadA	VDFIEL	VPE1	LQKR	GLYR	VD	YEEO	G T Y]	REK	LFG	KG	NYR	LPI	DH	IA	AR	YRN	IIS	SN	v.											
RcaE	IDFIEH	/LPII	LRER	GVAA	SG	YDAI	PTLI	RER	LLG	TE	ΤPV	LRI	EDH	PG	AG	YRA	lQS	GA	гv	G.										
BdsA	EEFVDQ	7VPL1	LQQR	GVFR	RTE	YEG?	C <mark>T</mark> L]	REH	LGL	AH	PEV	GS	IR.																	
CmoJ	DAFVEK	TPII	LQER	GLYR	QD	RGG	F T LI	REN	LGL	GI	РОН	٥s	νLΗ	SS	нн								• •	• •						

Figure S2. Sequence alignment of MsuD structural homologs. Homologous structures identified using the DALI server with a minimum sequence identity cut-off of 19% are displayed. The protein secondary structural elements from the MsuD structure are indicated above the MsuD sequence. Blue boxes indicate conserved and similar residues, with strictly conserved residues highlighted in red and similar residues in red text. The alignment figure was prepared using the Easy Sequencing in PostScript (ESPript) online program v3.0 (1). MsuD (GenID:ABA75653.1) was aligned with SsuD (GenID: CAB40391.1, seqID: 67%), RutA (GenID:AAC74097.1, seqID: 25%), PDBID: 3RAO (GenID:AAS39998.1, seqID: 32%), DmoA (GenID:ADU77278.1, seqID: 23%), LadA (GenID:ABO68832.1, seqID: 25%), RcaE (UniProt ID: A0A3B6UEK8, seqID: 23%), BdsA (GenID:BAC20180.1, seqID: 22%), and CmoJ (GenID:CAB14891.1, seqID: 24%).



Figure S3. Comparison of class C flavin-dependent monooxygenases that crystallize as tetramers. Tetramers are displayed from left to right are MsuD (PDB ID 7JW9), SsuD (PDB ID 1NQK), and BdsA (PDB ID 5XKD). Protein chains are colored green, blue, yellow, and gray, respectively.



Figure S4. Representative polder omit electron density maps for structures of MsuD demonstrating the presence of the active site lid, protein C-terminus, and the ligands FMN, MS⁻,

and succinate. (A) and (B) Polder omit electron density for the best ordered lid region and Cterminus (chain A) from ternary-MsuD (cocrystal) are displayed. Protein chains A-D are colored green, blue, yellow, and gray, respectively. Polder mFo-DFc omit maps were calculated for segments of five residues at a time and are contoured at 3σ level, carved to 2 Å about the lids. Polder omit maps for ligands (FMN, MS⁻, and succinate) are shown from chains A/B/C/D on top and E/F/G/H on the bottom of the MsuD tetramer in (C) ternary-MsuD (cocrystal), (D) binarysoak MsuD, (E) binary-titrated MsuD, and (F) ternary-soak MsuD. Omit maps for ligands are at 5σ level and Fo-Fc electron density maps at $\pm 3\sigma$ level, colored green for positive peaks and red for negative peaks, carved up to 2 Å about the ligands.



Figure S5. Analysis of residues and surrounding environment of (A) FMN and (B) MS⁻ in chain A of ternary-MsuD, showing interactions within hydrogen bonding distance and residues within van der Waals contact distances. Figure generated using LigPlot+ (2).



Figure S6. (A) Overlay of unliganded MsuD (black), ternary-MsuD (green), SsuD (pink), CmoJ (white), and PDBID: 3RAO (orange), demonstrating the position of aromatic residues analogous to W195 in the absence of FMN. FMN is displayed with cyan carbons from the ternary-MsuD cocrystal structure. (B) Overlay of FMN from different homolog structures demonstrating differences in ribityl tail positioning. The isoalloxazine rings of FMN were superimposed for MsuD (green, PDB ID 7JW9) with RutA (cyan, PDB ID 5WAN), LadA (purple, PDB ID 3B9O), RcaE (yellow, PDB ID 5W4Y) and BdsA (pink, PDB ID 5XKD).



Figure S7. Recombinantly expressed MsuD variants were purified and analyzed by SDS-PAGE. Approximately 1 µg samples of fully purified MsuD including WT and variants were loaded onto a BioRad AnyKD Protein Gel and stained with SimplyBlue SafeStain. The MW sizes of the NEB Protein Ladder are labeled.



Figure S8. Structural superimpositions of MsuD homologs reveal similarities and differences in their lids and C-termini. (A) An overlay of several lids from MsuD homologs is displayed. Protein overlays are displayed as ribbons. MsuD is colored green with FMN and MS⁻ in cyan and magenta sticks, respectively. A variable number of helices are present in the lid region of different homologs, and an overlay demonstrates how the first helix is consistently placed. The homologs shown are RcaE (orange), LadA (magenta) and CmoJ (gray). (B) A comparison of the C-termini in MsuD and its homologs is displayed. The C-termini of the three tetrameric monooxygenase structures are displayed for MsuD (blue), SsuD (red), and BdsA (magenta), and the dimeric homologs (DmoA, EmoA, LadA, RcaE, CmoJ, and PDB ID 3SDO) are shown in white. The C-terminus in MsuD is in a conformation not previously observed.



Figure S9. The proposed reactions of RutA, DszA, and HcbA1 demonstrate differences in substrate oxidation state with MsuD. Recent work by Matthew's et al have proposed the use of an N5-(hydro)peroxy (FI_{N500})intermediate in the reactions of three class C two-component flavin-dependent monooxygenases, RutA, DszA, and HcbA1, and a flavin N5-oxide is formed in the reaction cycle (3). Considering that the starting substrates of RutA, DszA, and HcbA1 are uracil, dibenzothiophene sulfone, and hexachlorobenzene, respectively, these substrates are in a more oxidized state at the carbon center proposed for attack by the FI_{N500} than an alkanesulfonate substrate for MsuD. Therefore, the mechanisms of RutA, DszA, and HcbA1 would generate FI_{N500} and must invoke an NADH molecule for the final water reduction step to regenerate oxidized flavin. Oxidation states for carbon in the MsuD reaction are shown.

Supporting Tables.

Table S1. Genome and gene information for *msu* and *sfn* genes from *P. fluorescens* P0-1, *P. aeruginosa* PAO1, and *P. putida* KT2440.

Sequence Name	Protein Name	Length	locus_tag	old_locus_tag	gene product annotation
	SfnR1	1131	PA2354		transcriptional regulator
	MsuC	1185	PA2355		FMNH ₂ -dependent monooxygenase
P. aeruginosa PAO1	MsuD	1146	PA2356		methanesulfonate monooxygenase
NC_002516	MsuE	561	PA2357		FMN reductase MsuE
	SfnR2	1086	PA2359		transcriptional regulator
	SfnG	1095	PA3954		hypothetical protein
	MsuE	564	PFL01_RS19665	Pfl01_3915	FMN reductase
	MsuD	1146	PFL01_RS19670	Pfl01_3916	FMNH ₂ -dependent alkanesulfonate monooxygenase
P. fluorescens Pf0-1 NC 007492	MsuC	1188	PFL01_RS19675	Pfl01_3917	acyl-CoA dehydrogenase family protein
_	SfnR	1104	PFL01_RS19680	Pfl01_3918	sigma-54-dependent Fis family transcriptional regulator
	SfnG	1095	PFL01_RS14520	Pfl01_2879	dimethyl sulfone monooxygenase SfnG
	SfnR	1104	PP_RS14410	PP_2771	sigma-54-dependent Fis family transcriptional regulator
	MsuC	1188	PP_RS14415	PP_2772	acyl-CoA dehydrogenase family protein
	hypothetical protein CDS	159	PP_RS14420	PP_2773	hypothetical protein
	SfnB	1233	PP_RS14335	PP_2755	SfnB family sulfur acquisition oxidoreductase
	hypothetical protein CDS	387	PP_RS14340	PP_2756	hypothetical protein
	sugar ABC transporter substrate- binding protein CDS	972	PP RS14345	PP 2757	sugar ABC transporter substrate-binding protein
P. putida NC 002947	sugar ABC transporter substrate- binding protein CDS	1029	PP RS14350	PP 2758	sugar ABC transporter substrate-binding protein
	sugar ABC transporter ATP- binding protein CDS	1539	PP_RS14355	PP_2759	sugar ABC transporter ATP-binding protein
	ABC transporter permease CDS	978	PP_RS14360	PP_2760	ABC transporter permease
	ABC transporter permease CDS	987	PP_RS14365	PP_2761	ABC transporter permease
	SfnA	1239	PP_RS14370	PP_2762	acyl-CoA dehydrogenase family protein
	SfnF	561	PP_RS14375	PP_2764	FMN reductase
	SfnG	1092	PP_RS14380	PP_2765	dimethyl sulfone monooxygenase SfnG

Table S2. Structural alignment results for MsuD from the DALI server (4), containing matches above or equal to 19% sequence identity. With the exceptions of RcaE, 1LUC, and 5LXE, these proteins are predicted to be class C flavin-dependent monooxygenases.

Chain	RMSD	% id	Number of aligned Cα	Name	Description	References
1NQK	1.3	69	328	SsuD	Alkanesulfonate monooxygenase	(5)
3RAO	2.0	32	320	N/A	Luciferase-like monooxygenase	(6)
5WAN	2.1	25	312	RutA	Pyrimidine monooxygenase	(7)
5W4Z	2.1	23	312	RcaE	Riboflavin lyase	(8)
3B9O	2.5	25	328	LadA	Long-chain alkane monooxygenase	(9)
5XKD	2.6	22	336	BdsA	Dibenzothiophene monooxygenase	(10)
3SDO	2.3	21	248	N/A	Nitrilotriacetate monooxyge nase	(11)
6ASL	2.7	24	248	CmoJ	Cysteine salvage pathway	(8)
5DQP	2.8	19	320	EmoA	Ethylenediaminetetraacetic acid monooxygenase	(12)
6AK1	2.8	23	328	DmoA	Dimethylsulfide monooxygenase	(13)
1LUC	2.7	19	296	N/A	Bacterial luciferase	(14)
3FGC	2.7	18	304	N/A	Bacterial luciferase	(15)
5LXE	2.4	21	288	Rh- fgd1	Glucose-6-phosphate dehydrogenase	(16)

Table S3. Summary of ligand occupancies and lid regions within MsuD crystal structures. The active site lid consists of residues D250–L282. In structures with an ordered C-terminus, the final four residues are unable to be built. For soaking experiments, FMN binding appears to be strongest in chains A/C and E/G of the two MsuD tetramers, whereas FMN binds strongly in all molecules of the cocrystal structure (highlighted yellow).

		Ternary cocrystal				Unliganded			Bina	ry-so	ak		Tern	ary-s	oak		Binary-titrated				
	Chain	FMN	-SM	Lid	C-terminus	FMN	MS ⁻	Lid	C-terminus	FMN	MS	Lid	C-terminus	FMN	MS ⁻	Lid	C-terminus	FMN	MS ⁻	Lid	C-terminus
	A	Fª	F	0 ^b	0	-	-	D	D	F	-	0	D	NF	Ρ	0	D	NF	_f	0	D
One	В	F	F	0	0	-	Ι	D	D	Ρ	Ι	D	D	Ρ	Ρ	0	D	- c	Ι	D	0
amer	С	F	F	0	0	-	Ι	D	D	F	Ι	0	D	NF	Ρ	0	D	NF	— f	0	D
Tetr	D	F	F	0	0	-	-	D	D	NF	-	D	D	Ρ	Ρ	O a	D	Ρ	-	D	0
		1								_		_	_				_			-	_
	E					-	-	D	D	F	_ [†]	0	D	NF	Р	0	D	NF	-'	0	D
wo	F					-	-	D	D	_ c	-	D	D	_ c	-	D	D	_ c	-	D	D
mer J	G					-	-	D	D	Ρ	-	D	D	Ρ	_ e	D ^h	D	Ρ	-	D	D
Tetra	Η					-	-	D	D	_ c	_	O d	0	Ρ	_ e	D	0	Ρ	-	D	0

^a Due to the moderate resolution of liganded structures, occupancies of ligand are reported as fully bound (F, 100%), nearly fully bound (NF, 90% - 99%), or partially bound (P, <90%).

^b O= ordered; D= disordered

^c Phosphate has been placed in these chains as a substitute for the phosphate head of FMN as the rest of the FMN density is not present

^d The lid of this chain was only partially built up to R262, denoting a potential alternate conformation

^e A chloride ion was placed in these chains of the ternary soak

^f Succinate was placed in these chains of the binary structures; no MS⁻ was used in the experiment

^g lacks electron density for R278 and R279

^h contains electron density for a7^c only

Table S4. The ribityl tail from different FMN-bound structures of the structural homologs of MsuD demonstrate altered torsion angles. The torsion angles of the ribityl tail range from nearly all 180° in the fully extended state of CmoJ to a more compact state within BdsA.

		Torsion angle of ri	orsion angle of ribityl tail (°)										
Structure	PDB ID	N10-C1'-C2'-C3'	C1'-C2'-C3'-C4'	C2'-C3'-C4'-C5'	C3'-C4'-C5'-O5'								
MsuD	7JW9	-131.8	-77.4	70.7	-161.1								
RutA	5WAN	178.6	-53.1	-179.9	60.8								
BdsA	5XKD	124.4	67.7	42.5	-84.0								
RcaE	5W4Y	-174.8	-58.3	-178.5	56.6								
LadA	3B9O	-125.7	-84.2	4.7	-91.0								
CmoJ	6ASL	-179.1	179.6	-177.8	-171.8								
LuxA	3FGC	169.0	-74.3	-99.3	-31.4								

Table S5. Comparison of activity of alkanesulfonate monooxygenases from different organisms and genes with different sulfonate group containing substrates.

					Substrates			
			Methanesulfonate	Octanesulfonate	Pentanesulfonate	HEPES	PIPES	MOPS
	MsuD	% rolative activity	96 ± 7	100 ± 10	62 + 5	12 ± 4	11 + 7	10 ± 7
	P. fluorescens		00 ± 7	100 ± 10	02 ± 3	12 ± 4	14 ± 7	19 ± 7
Organism	MsuD	% relative activity	100	NI/A	12.0	0.0	21.0	17.0
and Gene	P. aeruginosa ^a		100	N/A	13.0	9.0	51.0	17.2
	SsuD		1 5	100	07.0	22.0	62.4	70.6
	E. coli ^b		1.5	100	07.3	22.9	03.1	10.0

^a (17)

^b (18)

^c Octanesulfonate activity was scaled from 46.3% of 100% (4.1 μmol min⁻¹ mg⁻¹) 1,3-Dioxo-2-isoindolineethanesulfonic acid activity for comparison of substrates tested in this work. Other substrates were similarly scaled.

Table S6. The *msuD* gene of *P. fluorescens* Pf0-1 was amplified by PCR using the forward and reverse DNA primers shown. The site for restriction enzyme Nhel is underlined in the forward primers and the site for restriction enzyme HindIII is underlined in the reverse primers. Mutagenesis was done with forward and reverse designed primers.

Name	Sequence
Pfl01_msuD_fw	5'- GAT ATA <u>GCT AGC</u> ATG GAT GTT TTC TGG TTC CTG CCG ACT CAC GGC
(Pfl01_3916_fw)	GAC GG -3'
Pfl01_msuD_rv	5'- GAT ATA <u>AAG CTT</u> TCA GGC GTT GGC TTT GGC GGG CAG TAC ATC GTT
(Pfl01_3916_rv)	GGC G -3'
MsuD-W195A-F	5- C GTG TAC CTG ACG <u>GCA</u> GGC GAA CCA CCG GC -3
IVISUD-VV 195A-R	5- TUG AUU TGU TUG GUG GUU -3
IVISUD-INZZJA-I	
MsuD-R225A-R	5'- TTC ACT TTG CGC CCG TGA -3'
MsuD-R296A-F	5'- C GGT CTG GTG <u>GCA</u> GGC GGT TCC GGC ACT G -3'
MsuD-R296A-R	5'- ACG CCG GCC CAC AGG TTG -3'
MsuD-Stop-F	5'- TC ACC GGG CC <u>A TGA GA</u> C GAA ATG ATC GCC AAC G -3'
MsuD-Stop-R	5'- GGT TGG TGA CGC CGC GTC -3'
1	

References

- 1. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320-W324
- 2. Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51**, 2778-2786
- 3. 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
- 4. Holm, L. (2020) DALI and the persistence of protein shape. *Protein Sci.* **29**, 128-140
- 5. 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.
- 6. 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. RCSB, doi: http://doi.org/10.2210/pdb3RAO/pdb
- 7. 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
- 8. 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. PhD Thesis, Texas A&M University: College Station, Texas
- 9. 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
- 10. 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
- 11. 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. RCSB, doi: 10.2210/pdb3SDO/pdb
- 12. 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

- 13. 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 Crystallographica Section F: Structural Biology Communications* **74**, 781-786
- 14. 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
- 15. 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
- 16. 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
- 17. Kertesz, M. A., Schmidt-Larbig, K., and Wuest, T. (1999) A novel reduced flavin mononucleotide-dependent methanesulfonate sulfonatase encoded by the sulfur-regulated *msu* operon of *Pseudomonas aeruginosa*. *J. Bacteriol.* **181**, 1464-1473
- 18. 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