

## **SUPPORTING INFORMATION**

# **Biosynthesis of a Central Intermediate in Hydrogen Sulfide Metabolism by a Novel Human Sulfurtransferase and its Yeast Ortholog**

Scott L. Melideo<sup>§</sup>, Michael R. Jackson<sup>§</sup>,

and Marilyn Schuman Jorns\*

Department of Biochemistry and Molecular Biology

Drexel University College of Medicine

Philadelphia, PA 19102

\*To whom requests for reprints should be addressed. Phone: (215) 762-7495 FAX: (215) 762-4452; E-mail: [marilyn.jorns@drexelmed.edu](mailto:marilyn.jorns@drexelmed.edu)

Table S1 Primers used for PCR reactions

Reaction	Primers
Amplification of the <i>RDL1</i> gene from <i>S. cerevisiae</i> genomic DNA	5'- GCTAGCCATATGTGGAAGGCCGTGATGAATG-3' 5'- CCTAGAGTATACTCGAGTAAGTCAAGTTTATCACCC-3'
Ligation-independent cloning of an optimized synthetic gene encoding human TSTD1 isoform 1	5'-AGATTGGTGGCATGGCCGGGGCGCCG-3' 5'-GAGGAGAGTTTAGACTCACGACTCCTTCTCCAG-3'
Ligation-independent cloning of an optimized synthetic gene encoding human TSTD1 isoform 2	5'-AGATTGGTGGCATGGCAGGTGTTAGTG-3' 5'-GAGGAGAGTTTAGACTCAGGATTCTTTTTCTAACC-3'
Ligation-independent cloning of an optimized synthetic gene encoding human TSTD1 isoform 3	5'-AGATTGGTGGCATGGCCGGAGCCCC-3' 5'-GAGGAGAGTTTAGACTCAGCGACCTGCTAATAAC-3'
Mutation of Cys98 to Ala in yeast RDL1	5'-GGAGCTAATATTTTATGCTGCTTCTGGCAAACGCG-3' 5'-GCGGTTTGCCAGAAGCAGCATAAAAATATTAGCTCC-3'
Mutation of Cys98 to Ser in yeast RDL1	5'- GGAGCTAATATTTTATTGTGCTTCTGGCAAACGCG -3' 5'- CGCGYYYGCCAGAAGCACAATAAAAATATTAGCTCC-3'
Mutation of Cys79 to Ala in human TSTD1 isoform 1	5'- CATCTGGTTTTTTTCGCACAGATGGGTAAACGC-3' 5'-GCGTTTACCCATCTGTGCGAAAAAACCAGATG-3'
Mutation of Cys79 to Ser in human TSTD1 isoform 1	5'- CATCTGGTTTTTTTCTCTCAGATGGGTAAACGC-3' 5'- GCGTTTACCCATCTGAGAGAAAAAACCAGATG-3'

Table S2 BLASTp Search of the Bacterial Proteomes that Contain SDO-TSTD1 Fusion Proteins Using Human SQOR as the Query Sequence

	Organism (Taxonomy ID)	Highest Scoring Blast Hit				
		GenBank accession number (remarks)	Query Cover (%)	E Value	Identity (%)	Distance (nucleotides) <sup>a</sup>
1	<i>Nitrosococcus oceani</i> (1229)	YP_343998.1	88	4e <sup>-144</sup>	42	73
2	<i>Nitrosococcus watsonii</i> (473531)	YP_003760988.1	88	5e <sup>-111</sup>	42	73
3	<i>Methylocella silvestris</i> (199596)	(no hit obtained)	-	-	-	-
4	<i>Methylobacter tundripaludum</i> (173365)	WP_006892758.1	87	1e <sup>-114</sup>	42	21
5	<i>Methylomicrobium album</i> (39775)	WP_005371245.1	87	4e <sup>-111</sup>	41	45
6	<i>Methyloglobulus morosus</i> (1410681)	WP_023495229.1	86	5e <sup>-115</sup>	42	113
7	<i>Methylocystis rosea</i> (173366)	WP_018406351.1	51	1e <sup>-10</sup>	27	nd
8	<i>Methylosarcina fibrata</i> (105972)	WP_020565271.1	87	7e <sup>-112</sup>	41	246
9	<i>Bradyrhizobiaceae bacterium</i> (709797)	WP_009735658.1	67	0.001	22	nd
10	<i>Afipia broomeae</i> (56946)	WP_006019594.1	65	3e <sup>-6</sup>	22	nd
11	<i>γ-proteobacterium HTCC2148</i> (247634)	(bacterial proteome unavailable)	-	-	-	-
12	<i>Rhizobium giardinii</i> (56731)	WP_018327723.1	86	5e <sup>-90</sup>	35	nd
13	<i>Mesorhizobium amorphae</i> (71433)	WP_006203294.1	79	2e <sup>-23</sup>	26	nd

<sup>a</sup>The indicated distance is determined based on the number of intervening nucleotides between the human SQOR homolog gene and the SDO-TSTD1 fusion protein gene in the bacterial chromosome. Complete genomic data are available only for *N. oceani* and *N. watsonii*. Genomic distances in other bacteria, where whole-genome shotgun contigs are available, could be determined when the genes for the SQOR homolog and the SDO-TSTD1 fusion protein were found in the same contig but not when the genes were located in different contigs (nd).

```

M A G A P T V S L P E L R S L L A S G R 20
atggctggagcgcaccacggctctcgcttctctgaactccggttcaactcctagcctccggacgg
|||||.|||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.
ATGGCCGGGGCGCCGACGGTGTCCCTCCCGAGCTTCGCTCTCTGCTCGCCTCTGGCCGT

A R L F D V R S R E E A A A G T I P G A 40
gccccgctcttcgacgtgcgctctcgcgagggagggcagctgggaccatcccaggggcg
||.|||||.|||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.
GCTCGGTTGTTTGATGTGCGTAGCCGGAAGAAGCGGCGGCAGGAACAATCCGGGTGCT

L N I P V S E L E S A L O M E P A A F Q 60
ctcaacatcccgggtgtccgagttggagagtgctctgcagatggagccagctgccttccag
|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.
CTCAACATCCCAGGTAAGTGAGCTCGAATCTGCTCTCCAATGGAACCGGCCGCGTTTTAG

A L Y S A E K P K L E D E H L V F F C Q 80
gcttatattctgctgagaagccaaagctggaagatgagcatctcgttttcttctgtcag
||.|||||.|||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.
GCATTGTAICTAGCGGAGAAACCGAAACTGGAAGACGAACATCTGGTTTTTTCTGCCAG

M G K R G L Q A T Q L A R S L G Y T G A 100
atgggcaagcggggcctccaggccacgcagctggccccggagtcttgatacaactggggct
|||||.|||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.
ATGGGTAAACCGGACTGCAAGCCACGCAGCTGGCAGCTCCCTGGTTATACCGGTGCT

R N Y A G A Y R E W L E K E S 115
cgcaactacgctggagcctatagagaatggttgagaaagagagt
||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.
CGTAACTATGCGGGTGCCTATCGCGAATGGCTGGAGAAGGAGTCG

```

Figure S1 The nucleotide sequence of the synthetic gene used to express recombinant TSTD1 isoform 1 (upper case) is compared with the original gene sequence (lower case). The top line shows the amino acid sequence of isoform 1.

```

M A G V S E L E S A L Q M E P A A F Q A      20
atggctggagtgccgagttggagagtgctctgcagatggagccagctgccttccaggct
|||||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
ATGGCAGGTGTTAGTGAATTAGAATCCGCATTACAAATGGAACCGGCAGCATTTC AAGCT

L Y S A E K P K L E D E H L V F F C Q M      40
ttatattctgctgagaagccaaagctggaagatgagcatctcgttttcttctgtcagatg
|||||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
TTATACTCAGCAGAAAAACCGAAATTAGAAGATGAACACTTAGTATTTTTTTGTCAAATG

G K R G L Q A T Q L A R S L G Y T G A R      60
ggcaagcggggcctccaggccacgcagctggcccggagtcttggatacaactggggctcgc
|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
GGAAAACGTGGTTTACAAGCAACTCAATTAGCACGCTCATTAGGATACACAGGCGCTCGG

N Y A G A Y R E W L E K E S      74
aactacgctggagcctatagagaatggttggagaaagagagt
|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
AATTATGCAGGGGCTTATCGCGAATGGTTAGAAAAAGAATCC

```

Figure S2 The nucleotide sequence of the synthetic gene used to express recombinant TSTD1 isoform 2 (upper case) is compared with the original gene sequence (lower case). The top line shows the amino acid sequence of isoform 2.

```

M A G A P T V S L P E L R S L L A S G R      20
atggctggagcgcccacgggtctcgcttccctgaactccggttccactcctagcctccggaagg
|||||.|||||.|||.|||||.|||.|||||.|||.|||||.|||.|||||.|||.|||||.
ATGGCCGGAGCCCCTACAGTCTCTTTACCAGAATTACGCTCATTATTAGCATCAGGCCGT

A R L F D V R S R E E A A A G T I P G A      40
gcccggtctctcgacgtgogctctcgcgaggaggcggcagctgggaccatcccagggggcg
|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.
GCCCGCCTTTTTGACGTACGTTCCGCGAAGAAGCCGCAGCAGGAACCATCCCAGGAGCT

L N I P V S E L E S A L O M E P A A F Q      60
ctcaacatcccgggtgtccgagttggagagtgctctgcagatggagccagctgccttccag
.|.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.
TTAAACATCCCAGTCTCAGAATTAGAATCAGCATTACAAATGGAACCAGCAGCATTTCAA

A L Y S A E K P K L E D E H L V F F C Q      80
gctttatattctgctgagaagccaaagctggaagatgagcatctcgttttcttctgtcag
||..|.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.
GCACCTTACTCCGCCGAAAAACCAAACTTGAAGATGAACACTTAGTATTTTTTTGTCAA

M G K R G L Q A T Q L A R S L G Y T G Y      100
atgggcaagcggggcctccaggccacgcagctggcccggagtcttgatacactgggtac
|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.|||||.
ATGGGAAACGTGGTTTACAAGCAACACAATTAGCTCGCAGCTTAGGTTATACAGGATAT

G E V W L L A G R      109
ggggaggtgtggctgctagctgggagg
||.|||||.|||||.|||||.|||||.
GGAGAAGTTTGGTTATTAGCAGGTCGC

```

Figure S3 The nucleotide sequence of the synthetic gene used to express recombinant TSTD1 isoform 3 (upper case) is compared with the original gene sequence (lower case). The top line shows the amino acid sequence of isoform 3.

```

1  MAGAPTVSLPELRSLLASGRARLFDVRSREEAAAGTIPGALNIPVSELESALQMEPAAFQ 60
3  MAGAPTVSLPELRSLLASGRARLFDVRSREEAAAGTIPGALNIPVSELESALQMEPAAFQ 60
2  MAG-----VSELESALQMEPAAFQ 19
***                                     *****

1  ALYSAEKPKLEDEHLVFFCOMGKRGLQATQLARSLGYTGGARNYAGAYREWLEKES 115
3  ALYSAEKPKLEDEHLVFFCOMGKRGLQATQLARSLGYTG-----YGEVWLLAGR 109
2  ALYSAEKPKLEDEHLVFFCOMGKRGLQATQLARSLGYTGGARNYAGAYREWLEKES 74
*****                                     **

```

Figure S4 Alignment of human TSTD1 isoforms. Amino acids encoded across a splice junction and cysteines are highlighted in yellow and red, respectively. All isoforms contain a tripeptide at the N-terminus and a central core of 55 amino acids that are encoded by exons 1 and 3, respectively. The N-terminal tripeptide in isoforms 1 and 3 is followed by a stretch of 41 amino acids (encoded by exon 2) that is absent in isoform 2. The 16 amino acids at the C-terminus of isoforms 1 and 2 (encoded by exon 4) are replaced by a decapeptide in isoform 3 (encoded by retained intron 3).

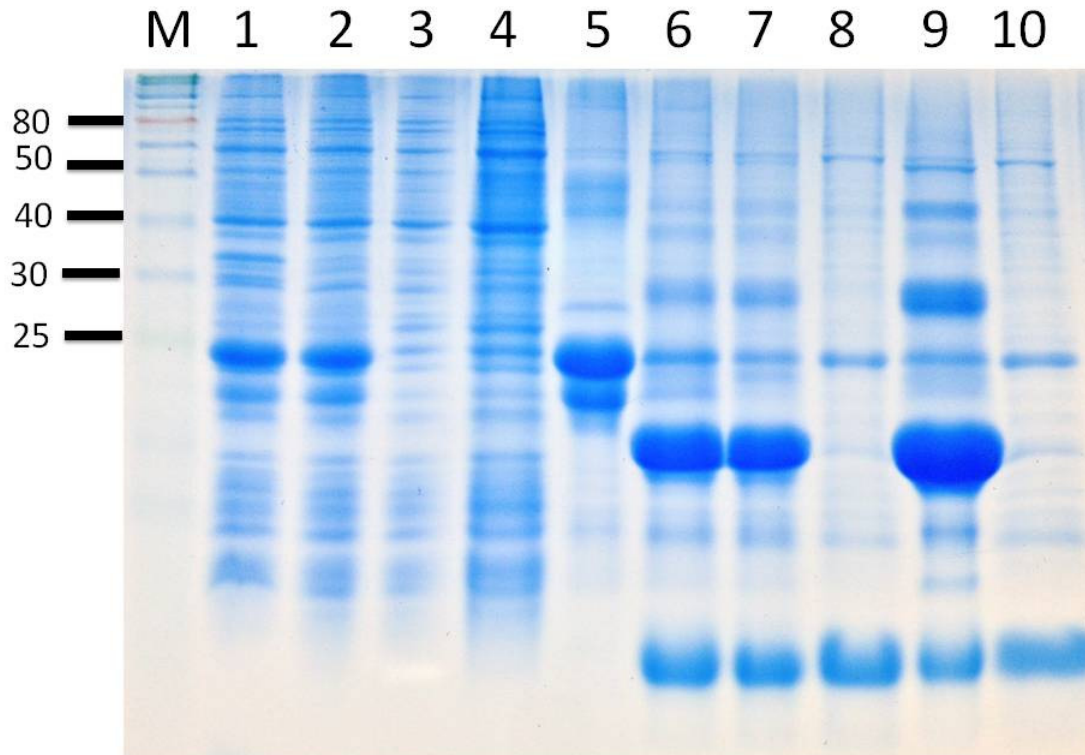


Figure S5 Purification of recombinant human TSTD1 isoform 2 with a cleavable N-terminal His-SUMO tag. The SDS 15% polyacrylamide gel was stained for protein with ProSieve Blue Protein Staining Solution (Lonza); lane M, molecular markers; lane 1, whole cell lysate; lane 2, low speed supernatant; lane 3, Ni affinity column flow-through; lane 4, Ni affinity column wash; lane 5, Ni affinity column eluate; lane 6, Ni affinity column eluate after cleavage with SUMO hydrolase; lane 7, SUMO hydrolase-treated sample after dialysis; lane 8, second Ni affinity column flow-through; lane 9, second Ni affinity column eluate; lane 10, dialyzed sample from lane 8.



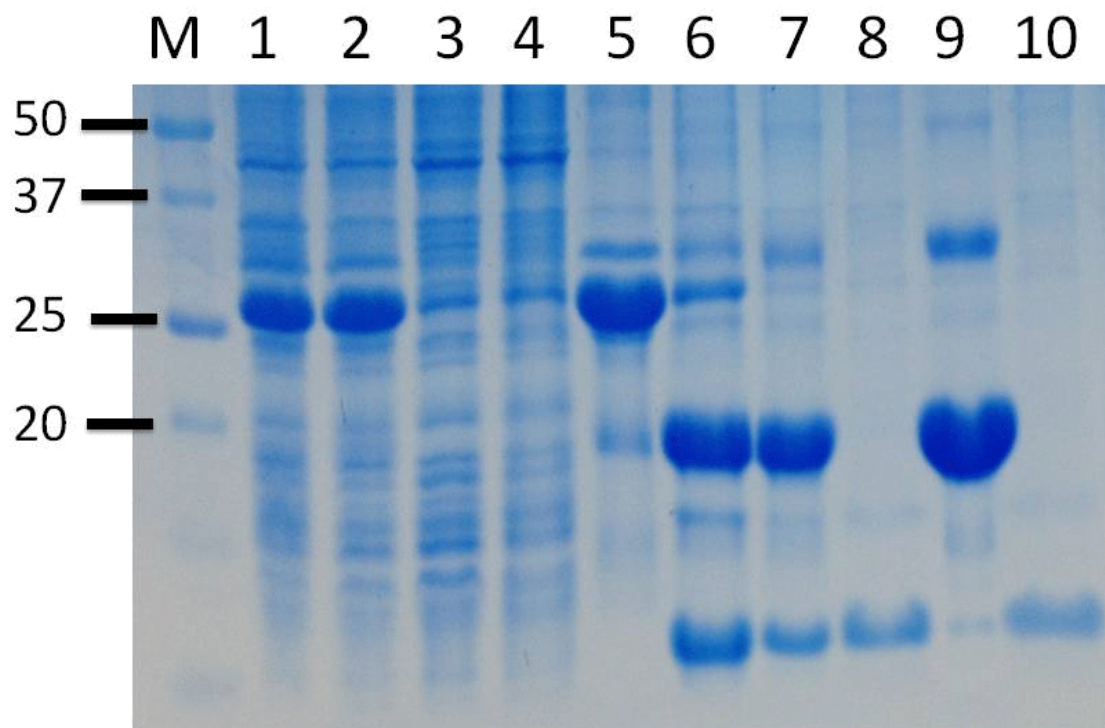


Figure S6 Purification of recombinant human TSTD1 isoform 3 with a cleavable N-terminal His-SUMO tag. The SDS 15% polyacrylamide gel was stained for protein with ProSieve Blue Protein Staining Solution (Lonza); lane M, molecular markers; lane 1, whole cell lysate; lane 2, low speed supernatant; lane 3, Ni affinity column flow-through; lane 4, Ni affinity column wash; lane 5, Ni affinity column eluate; lane 6, Ni affinity column eluate after cleavage with SUMO hydrolase; lane 7, SUMO hydrolase-treated sample after dialysis; lane 8, second Ni affinity column flow-through; lane 9, second Ni affinity column eluate; lane 10, dialyzed sample from lane 8.

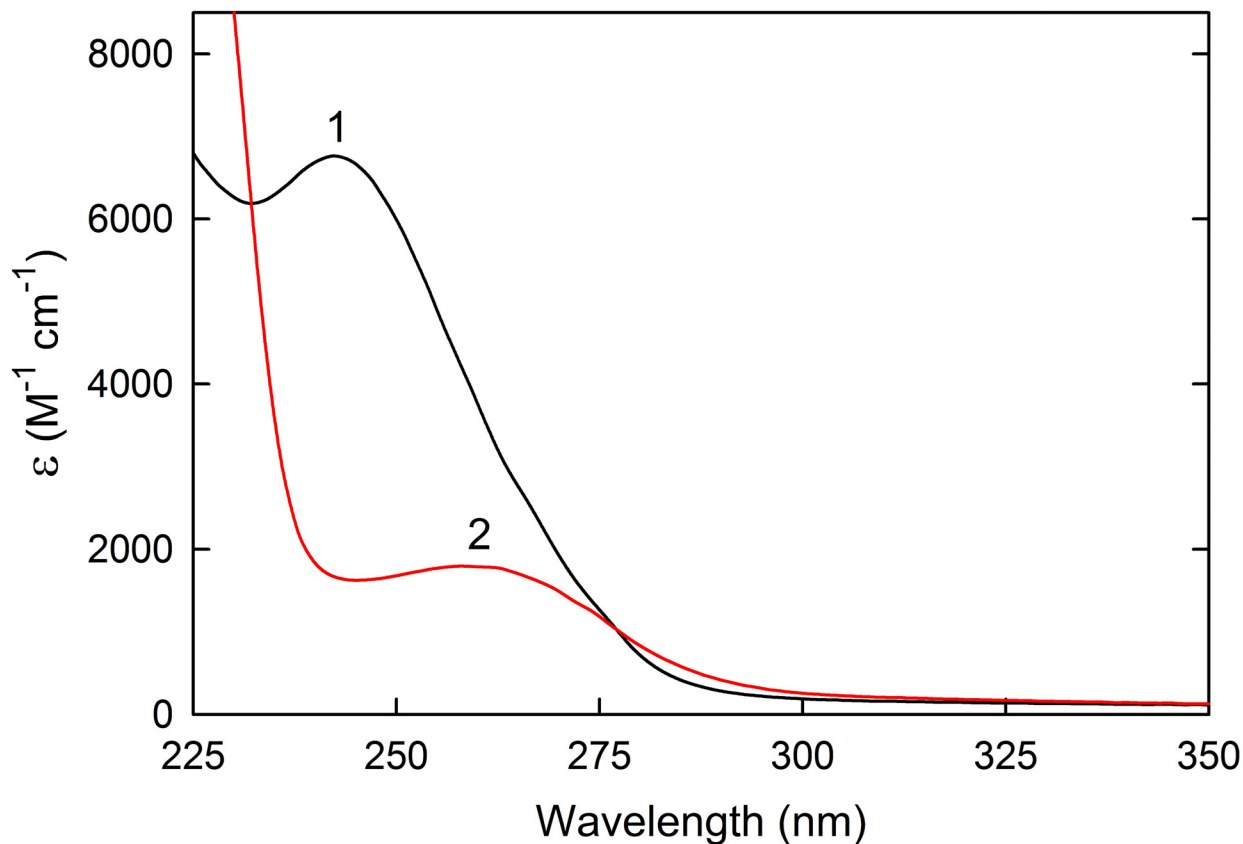


Figure S7 Spectral properties of *p*-toluenethiosulfonate ( $p\text{-Tol-SO}_2\text{S}^-$ ) and *p*-toluenesulfonite ( $p\text{-Tol-SO}_2^-$ ). Spectra were recorded at 25 °C using a 2 mm cuvette. Curve 1 is the absorption spectrum of  $p\text{-Tol-SO}_2\text{S}^-$  in 50 mM potassium/sodium phosphate buffer, pH 8.0, recorded 2 s after addition of 1.6 mM potassium cyanide. Curve 2 is the absorption spectrum of  $p\text{-Tol-SO}_2^-$  in the same buffer, recorded 190 s after addition of 60 nM human rhodanese, an enzyme previously shown to catalyze the transfer of the sulfane sulfur from  $p\text{-Tol-SO}_2\text{S}^-$  to cyanide (Sorbo, B. (1953) Acta Chem. Scand. 7, 32-37).

N. oceani	MIFRQLFDPESSSTYTYLIGDPATKEAVFIDPVNTRVDEYLNLLNKYNLKLKYSLETHAHA	60
M. silvestris	MIFRQLFDHVSGETYTYLLASRPGGALIIDPVLEKVDRYLQLFDELVDKLVKAVDTHLHA	60
M. tundripaludum	MLFKQLFDQETWYTYTYLIADPVSKDAILIDPVNTHIDEYIELLAHGLQLKYSLETHVHA	60
M. album	MIFKQLFDPEETWYTYTYFIADTDAKEAVLIDPVKSHIDEYIALLDEHGLKLYTLETHVHA	60
M. morosus	MLFKQLFDQETWYTYTYLIADPANKEAVLIDPVNTHIDYLAAMLTEQGLQLKYLETHVHA	60
	*:*:**** : ****:.. :*:**** ::* *: : : :*: :*:** **	
N. oceani	DHITASGLLRQHTGAKTGIGQACGAQYADYQLKDGVVLAFFGQEEIKVLATPGHTPGSIS	120
M. silvestris	DHVTGLGALRDRTHCITVMGEQTKADVVS MRVADGDRIDIE-GLSLEALFTPGHTDSDYS	119
M. tundripaludum	DHITASGLLRQRLGAQTAVSGLCGAESADIQIQDGDIFKFAGDEQIKVIATPGHTRGSIS	120
M. album	DHITASGLLRQKLGAEETGVQLCGAIGADLQDGDVLEFGNGEKIKVIATPGHTQGSVS	120
M. morosus	DHITASGLLRQKLGSGTGVGALCGAETADFQMKDGLFSFTNGEQIKVIATPGHTKGSMS	120
	**:* . * **:: . * .. * .. : : ** : : : : : : : ***** **	
N. oceani	YLWRDRVFTGDALLINGCGRITDFQGGDPTLYDSVTQKLFPLPGETIVYPGHYNGRWVS	180
M. silvestris	FLEGRVFTGDTLLIRGTGRITDFQNGDPRAYHSIFDRLLKLPDETLVYPAHDYKGDTVS	179
M. tundripaludum	FLWRDLFTGDSLLIGGCGRTDFQGGDAGALYDCITQRLFTLPDETLVYPGHYQQRWVS	180
M. album	FLWRDRVFTGDSLFIGGCGRTDFQSGDAGALYDCITRKLFTLPDETLVYPGHYQQRWVS	180
M. morosus	FLWRDRVFTGDSLFIGGCGRTDFQGGDAGALYDCITQRLFTLPDDTLVYPGHYQQRWVS	180
	:* . *:****:*:* * ***** ** :*.. :*:..** :*:****.***: **	
N. oceani	SVEQERTGNRLAGKTRAFFIEIMNNLNLPKPRLIDEAVPANRRCGLTEEE-----IRQD	238
M. silvestris	TIGERLRFNPRKLVRSVDEYVDLMNNLNLPNPKMMDVAVPANMRVGFHQDE-----IARKG	239
M. tundripaludum	SIMQERTTNPRLAGKTRQFFIEIMNNLNLPKPRLIDEAVPANRYCGLDENERQDAVALRD	240
M. album	SIMQERITNPRLAGKTRREEFFIEIMNNLNLPRPRLIDEAVPANRYCGLDEDERQDAVARRD	240
M. morosus	NIVQERTTNPRLAGKTRDEFITIMNNLNLPKPKLIDEAVPANRYCGLLENERQDAVAHRE	240
	.. :** * ** : : : : :*****.*:~* ***** * : :*	
N. oceani	TMMG-----EKRVSFQDLVQEARQVREIDVATVKQRLGDKGKTAIID	278
M. silvestris	WAAQ-----AAEALALFGRPDIAIVD	256
M. tundripaludum	A-----SLPVRATTRTEDMVAEAKQHIVEIDVAKSKQLLTKGNVVLVD	283
M. album	STR-----PERAMISAQDLVSAAKERITEIGIDKARQLLNQSGVAVVD	283
M. morosus	TVRESAKPGDVPVVNCGGDMPKSGGMTVQDLVAAAKQOITEVNVKAKQLIVQGSITVID	300
	: : : *	
N. oceani	VREPEEFAA-GHLPGAINVPRGVLEFRLG--NTAELADPNIPITILYCQTGGRAALAAWS	334
M. silvestris	LREKREKQGIIPGSLHAPYDLRESIGPGGALHELAAAGKTIILFYCAFGERSAMAVQA	316
M. tundripaludum	TREESEYAA-GHVDGALLVPRGMLEFKIG--NMPELADKSKAVLIYCR LGNRSALAAQT	339
M. album	VREESEYAA-GHIDNALPIPRGVLEFKVG--ATPELADKSKTVVVYCR TGGRAALAAQT	339
M. morosus	TREESEYAA-GHIDNAVLLPRGVLEFRIN--TIPELADKAKPVLIIYCR TGGRAALAAQS	356
	** * * : : : * * . : : ** : : ** * * :*:* . :	
N. oceani	LKCLGYTDATLIAGGYDAWRAAKQNAD-	361
M. silvestris	AQDAGFKSARHIAGGFDAWKKENGAVAP	344
M. tundripaludum	LQQLGYTNVLSMAGGF EAWKKSQEAEE--	365
M. album	LQNLGYSNVLSIAGGYEAWQKSL-----	362
M. morosus	LKTLGYTNVLSIAGGYETWQKGN-----	379
	: * :... :****:~*:	

Figure S8 Multiple sequence alignment of selected SDO-TSTD1 bacterial fusion proteins. The black box marks a poorly conserved, variable length linker (9 to 52 amino acids) between the N-terminal SDO-like domain (230 or 231 amino acids) and the C-terminal TSTD1-like domain (96 to 105 amino acids). GenBank accession numbers: *Nitrosococcus oceani* (WP\_013220760.1); *Methylocella silvestris* (WP\_012591638.1); *Methylobacter tundripaludum* (WP\_006892756.1); *Methylobacterium album* (WP\_005371248.1); *Methyloglobulus morosus* (WP\_023495230.1).

### A) N-terminal SDO-like Domain

N. oceani	-----MIFRQ <sup>Y</sup> LFDPESSTY <sup>T</sup> TYLLIGDPATKEAVFIDP <sup>N</sup> TRVDE	38
M. silvestris	-----MIFRQ <sup>L</sup> LDHVS <sup>G</sup> TYYTLLASRPGGEALIDP <sup>V</sup> LEKVDR	38
M. tundripaludum	-----MLFKQ <sup>L</sup> LDQETWTYYTYLIADPVSKDALLIDP <sup>V</sup> NTHIDE	38
M. album	-----MIFKQ <sup>L</sup> LDPE <sup>T</sup> WTYYTYFIADTAKEAVLIDP <sup>V</sup> KSHIDE	38
M. morosus	-----MLFKQ <sup>L</sup> LDQETWTYYTYLIADPANKEAVLIDP <sup>V</sup> NTHIDD	38
Human SDO	MAEAVL <sup>R</sup> VARRQLSQRGGSGAPILLRQMFEPVSC <sup>T</sup> FYYLLGDR <sup>S</sup> REAVLIDP <sup>V</sup> LETAPR	60
	: : : : * : : : * : : * : : : * : : *	
N. oceani	YLNLLN <sup>K</sup> Y <sup>N</sup> LKLYSL <sup>E</sup> THAHAD <sup>H</sup> ETASGLLRQHT-GAKTGI <sup>G</sup> QACGAQYADYQ <sup>L</sup> KDGVV	97
M. silvestris	YLQLFDEL <sup>D</sup> VKLKAVD <sup>T</sup> HLHAD <sup>H</sup> V <sup>T</sup> GLGALDRDT-HCITV <sup>M</sup> GEQTKADVSMRVADGDR	97
M. tundripaludum	YIELLAH <sup>G</sup> LQ <sup>L</sup> KYSL <sup>E</sup> THVHAD <sup>H</sup> ETASGLLRQRL-GAQTAVSGLCGAE <sup>S</sup> ADIQ <sup>I</sup> QD <sup>G</sup> DI	97
M. album	YIALLDEH <sup>G</sup> LKLKY <sup>T</sup> LE <sup>T</sup> THVHAD <sup>H</sup> ETASGLLRQKL-GAETG <sup>V</sup> GQLCGAIGADLQ <sup>L</sup> QD <sup>G</sup> DV	97
M. morosus	YLAMLTEQ <sup>G</sup> LQ <sup>L</sup> KY <sup>T</sup> LE <sup>T</sup> THVHAD <sup>H</sup> ETASGLLRQKL-GSQTG <sup>V</sup> GALCGAETAD <sup>F</sup> Q <sup>M</sup> KD <sup>G</sup> DL	97
Human SDO	DAQLIKELGLRLLYAVNT <sup>H</sup> CHAD <sup>H</sup> ETGSGLLR <sup>S</sup> LLPGCQSVIS <sup>R</sup> LSGA-QADL <sup>H</sup> IED <sup>G</sup> DS	119
	: : * : : * : : * : : * : : * : : *	
N. oceani	LAFGQGE <sup>E</sup> IKVLATPGHTPGSISYLWRDR--VFTGDALLINGCGR <sup>T</sup> DFQGGDPG <sup>T</sup> LYDSV	155
M. silvestris	IDI-EGL <sup>S</sup> LEALF <sup>T</sup> PGHTDDSY <sup>S</sup> FLEGR--VFTGDTLLIRGTG <sup>R</sup> TDFQNGDPRAQYHSI	154
M. tundripaludum	FKFAGDEQ <sup>I</sup> KVIATPGHTRG <sup>S</sup> ISF <sup>L</sup> WRDR--LFTGDSLLIGGCGR <sup>T</sup> DFQGGDAGALYDCI	155
M. album	LEFGN <sup>E</sup> E <sup>K</sup> IKVIATPGHTQGSVS <sup>F</sup> LRDR--VFTGDSLFIGGCGR <sup>T</sup> DFQSGDAGALYDCI	155
M. morosus	FSFTN <sup>G</sup> EQ <sup>I</sup> KVIATPGHTKGSMS <sup>F</sup> LRDR--VFTGDSLFIGGCGR <sup>T</sup> DFQGGDAGALYDCI	155
Human SDO	IRFGRFA-LETRAS <sup>P</sup> GHTPGCVTFVLNDSMAFTGDALLIRGCGR <sup>T</sup> DFQGGCAKTLYHSV	178
	: : : : * : : * : : : : * : : * : : * : : *	
N. oceani	TQKLFTLPGETIVYPGH <sup>D</sup> YNGRWVSSVEQERTGN <sup>R</sup> L <sup>G</sup> KTRAEFIEIMNNLNLPK <sup>R</sup> LI	215
M. silvestris	FDRL <sup>L</sup> KL <sup>P</sup> DELTVYPADYKGD <sup>T</sup> VSTIGEERLFNPRLK <sup>V</sup> RSVDEYVDLMNNLNLPK <sup>M</sup> M	214
M. tundripaludum	TQRLFTLPDELTVYPGH <sup>D</sup> YQQRWVSSIMQERTTNPRL <sup>G</sup> KTR <sup>E</sup> QFIEIMNNLNLPK <sup>R</sup> LI	215
M. album	TRK <sup>L</sup> FTLPDELTVYPGH <sup>D</sup> YQQRWVSSIMQERITNPR <sup>L</sup> GKTR <sup>E</sup> E <sup>F</sup> IEIMNNLNLPK <sup>R</sup> LI	215
M. morosus	TQRLFTLPDDTLVYPGH <sup>D</sup> YQQRWVSNIVQERTTNPRL <sup>G</sup> KTR <sup>D</sup> E <sup>F</sup> ITIMNNLNLPK <sup>P</sup> LI	215
Human SDO	HEKIFTLPGDCLIPADYHGFTVSTVEERTLN <sup>P</sup> RLTL-SCEE <sup>F</sup> VKIMGNLNLPK <sup>P</sup> QI	237
	: : * : : * : : * : : * : : * : : *	
N. oceani	DEAVPANRRCGLTEEE-	231
M. silvestris	DVAVPANMRVGFHQDE-	230
M. tundripaludum	DEAVPANRYCGLDENE-	231
M. album	DEAVPANRYCGLDEDE-	231
M. morosus	DEAVPANRYCGL <sup>E</sup> ENE-	231
Human SDO	DFAVPANMR <sup>C</sup> GVQTPTA	254
	* ***** *	

### B) C-terminal TSTD1-like Domain

N. oceani	-----VATV <sup>K</sup> QR <sup>L</sup> LDG <sup>K</sup> TAIIDVREPEFAA-GHLPGA <sup>I</sup> NVPRGVLE <sup>F</sup> R <sup>L</sup> G---NT-	308
M. silvestris	-----AAEALALFR <sup>P</sup> DIATV <sup>D</sup> LREKREKQGIIPGSLHAPY <sup>P</sup> DLRESIG <sup>P</sup> GGALH	290
M. tundripaludum	-----VAKSKQL <sup>L</sup> TKGN <sup>V</sup> VLD <sup>T</sup> REESEYAA-GHVDGALLVPRG <sup>M</sup> LE <sup>F</sup> K <sup>I</sup> G---NM-	313
M. album	-----IDKARQLN <sup>Q</sup> SGVA <sup>V</sup> VDVRESEYAA-GHIDNALPIPRGV <sup>L</sup> E <sup>F</sup> KV <sup>G</sup> ---AT-	313
M. morosus	-----VEKAKQLIVQ <sup>S</sup> ITVID <sup>T</sup> REESEYAA-GHIDNAVLLPRGVLE <sup>F</sup> FRIN---TI-	330
Human TSTD1	MAGAPT <sup>V</sup> SLPEL <sup>R</sup> SLLASGRAR <sup>L</sup> FD <sup>V</sup> RSREAAA-GTIPGALNIPVSE <sup>L</sup> SAL <sup>Q</sup> MEPA <sup>A</sup> F	59
	: : * : : * : : * : : * : : *	
N. oceani	-----AELADPNIP <sup>I</sup> I <sup>L</sup> Y <sup>C</sup> QTGGR <sup>A</sup> ALA <sup>A</sup> WSL <sup>K</sup> CLGY <sup>T</sup> DAT <sup>L</sup> IAGGYDA <sup>W</sup> RAAKQ <sup>N</sup> AD-	361
M. silvestris	-----ELAAAGKTILFYCAFGR <sup>S</sup> AMAVQA <sup>A</sup> QDAG <sup>F</sup> K <sup>S</sup> ARHIAGGFDA <sup>W</sup> KKENGAV <sup>A</sup> P	344
M. tundripaludum	-----PELADKSKAVLIY <sup>C</sup> R <sup>L</sup> GN <sup>R</sup> SALAAQ <sup>T</sup> LQQLGY <sup>T</sup> NVLSMAGGF <sup>E</sup> AWK <sup>K</sup> SQ <sup>E</sup> EAE--	365
M. album	-----PELADKSK <sup>T</sup> VV <sup>V</sup> Y <sup>C</sup> RTGGR <sup>A</sup> ALAAQ <sup>T</sup> LQNLGYSNVLSIAGGYE <sup>A</sup> W <sup>K</sup> S <sup>L</sup> -----	362
M. morosus	-----PELADKAKPVI <sup>L</sup> Y <sup>C</sup> RTGGR <sup>A</sup> ALAAQ <sup>S</sup> LK <sup>T</sup> LG <sup>Y</sup> TN <sup>V</sup> LSIAGGYE <sup>T</sup> W <sup>K</sup> GN-----	379
Human TSTD1	QALYSAEK <sup>P</sup> K <sup>L</sup> E <sup>H</sup> LVEFF <sup>C</sup> QMK <sup>R</sup> GLQATQLARSLGY <sup>T</sup> GARNYAGAY <sup>R</sup> E <sup>W</sup> LEKES----	115
	: : : * : : * : : * : : * : : *	

Figure S9 Multiple sequence alignment of human SDO or human TSTD1 with the N-terminal SDO-like or the C-terminal TSTD1-like domain, respectively, in selected bacterial SDO-TSTD1 fusion proteins. Panel A: The black box marks a  $\beta$ -lactamase signature motif. Conserved metal binding and glutathione binding sites, identified in previous studies (Zhang et al. (2013) PLOS ONE, 8, e81885), are highlighted in yellow and magenta, respectively. Panel B: The black boxes mark a six amino acid active site loop and a rhodanese homology domain signature motif. GenBank accession numbers: *Nitrosococcus oceani* (WP\_013220760.1); *Methylocella silvestris* (WP\_012591638.1); *Methylobacter tundripaludum* (WP\_006892756.1); *Methylomicrobium album* (WP\_005371248.1); *Methyloglobulus morosus* (WP\_023495230.1).