Supporting Information Appendix:
Sulfatase-Activated Fluorophores for Rapid Discrimination of Mycobacterial Species and Strains

Kimberly E. Beatty, Monique Williams, Brian L. Carlson, Benjamin M. Swarts, Robin M. Warren, Paul D. van Helden, and Carolyn R. Bertozzi

Materials and Methods

Synthesis and Purification of DDAO-sulfate

All chemicals were purchased from Sigma-Aldrich unless otherwise indicated. DDAO (61 mg, 198 μmol, Synchem OHG) was dissolved in anhydrous DMF (2.8 mL) in a flask on ice under inert atmosphere (nitrogen gas). Sodium hydride (60% dispersion in mineral oil, 22.5 mg, 563 μmoles) was added to the stirring solution of DDAO. The solution changed from a deep red to blue. After 15 min, the flask was removed from the ice bath and allowed to warm to room temperature (15 min). Sulfur trioxide trimethylamine (SO_3-TMA, 135 mg, 970 μmol) was added, and the reaction was heated to 55 °C for 4 hours. The reaction was then cooled and quenched with methanol. The solvent was evaporated, and the resulting solid was immediately purified by silica column chromatography using SiliaFlash P60 (SiliCycle) in ethyl acetate:acetone (8:1 to 2:1 gradient) with 0.05% triethylamine to give a red oil, which was assumed to be DDAO-sulfate with a triethylammonium counterion (79 mg, 161 μmol, 81.5% yield).

DDAO-sulfate was further purified by reversed phase HPLC using a Varian Pro Star system on a Dynamax 100 Å C18 semi-preparative column. A linear gradient of water (solvent A) and methanol (solvent B) was used (100% to 0% water). A single peak corresponding to DDAO-sulfate was collected, and solvent was evaporated to give pure DDAO-sulfate as a dark orange solid (47 mg, 96 μmol, 48.5% overall yield). Stock solutions of DDAO-sulfate (20 mM in DMSO) were stored at -20 °C.

^{1}H NMR (CD_3OD, 400 MHz): δ 7.636 (s, 1 H), 7.56 (d, *J* = 5.6 Hz, 1 H), 7.52 (s, 1 H), 7.27 (d, *J* = 5.2 Hz, 1 H), 1.84 (t, *J* = 4.8 Hz, 6 H). Spectrum contains the triethylammonium counterion: δ 3.15 (dd, *J* = 4.8 Hz, 5 H), 1.24 (t, *J* = 4.8 Hz, 7.4H) (**Figure S8**). ^{13}C NMR (CD_3OD, 600 MHz): δ 174.59, 157.73, 150.39, 142.71, 141.25, 140.88, 139.01, 138.07, 136.09, 134.22, 121.75, 120.36, 40.67, 27.05 (**Figure S9**). High resolution mass spectrometry (ESI, negative mode) was performed at the UC-Berkeley Mass Spectrometry Laboratory. A peak corresponding to $C_{15}H_{10}Cl_2NO_5S^-$ was observed at 385.9649 m/z (calculated: 385.9662 m/z).

Characterization of Fluorophores

Fluorescence properties of DDAO and DDAO-sulfate were measured on a PTI fluorimeter. Excitation and emission spectra for DDAO and DDAO-sulfate were measured in PBS (pH 10) and normalized (**Figure S10**). The fluorescence quantum yields for DDAO and DDAO-sulfate were determined using the method described by Fery-Forgues and Lavabre (1). Fluorescein in 0.1 N NaOH (Quantum yield: 0.92, excitation at 430 nm, emission integrated from 440 to 720 nm) was used as a standard to find the quantum yield of DDAO-sulfate in basic PBS (pH 10). DDAO-sulfate was subject to hydrolysis in 0.1 N NaOH. The quantum yield of DDAO-sulfate was

calculated to be 0.028. Sulforhodamine 101 in ethanol (Quantum Yield: 0.95, excitation at 490 nm, emission integrated from 500 to 800 nm) was used as a standard to find the quantum yield of DDAO in ethanol and basic PBS (pH 10). The calculated quantum yield was 0.64 in ethanol and 0.40 in PBS.

Enzyme Activation of DDAO-sulfate with Commercial Sulfatases

A. aerogenes sulfatase (Sigma, S1629) was evaluated with DDAO-sulfate. This enzyme was used at a final concentration of 0.246 U/mL (enzyme units "U" are as defined by Sigma). DDAO-sulfate concentrations ranged from 25 to 800 μM. Reactions were performed in triplicate in buffer (50 mM Tris, 100 mM NaCl, 2.5 mM $MgCl_2$, 250 μM $CaCl_2$, pH = 7.5) at 37°C. The reaction was monitored by measuring DDAO fluorescence (ex. 635 nm, em. 670 bp 665) on a fluorescence microplate reader (Gemini XL, Molecular Devices). Kinetic parameters were determined using the method of initial rates(2) and plotted using Kaleidagraph. A general equation for Michaelis-Menton kinetics, velocity = (v_{max} * [substrate] / {Km + [substrate]}, was used to fit data (**Figure S11**). *A. aerogenes* sulfatase was evaluated with the commercial substrates 4MUS and *p*-NPS as described(3).

Patella vulgata sulfatase ("Keyhole Limpet," Sigma, S8629) was evaluated with DDAO-sulfate. This enzyme was used at a final concentration of 0.31 U/mL. DDAO-sulfate concentrations ranged from 5 to 200 μM. Reactions were performed in triplicate in acidic buffer (100 mM Potassium Acetate Buffer, pH = 5) at 37°C. The reaction was monitored by measuring DDAO fluorescence on a fluorescence microplate reader. Due to the substrate inhibition observed, a general equation for substrate inhibition, velocity = {v_{max} * [substrate]} / {Km + {[substrate]*(1+[substrate]/K_i}}, was used to fit kinetic data.

Similarly, *Helix pomatia* sulfatase (Sigma, S9626) was evaluated with DDAO-sulfate. This enzyme was used at a final concentration of 0.17 U/mL. DDAO-sulfate concentrations ranged from 10 to 100 μM. Reactions were performed in triplicate in acidic buffer (100 mM Potassium Acetate Buffer, pH = 5) at 37°C before analysis as described above.

Lastly, sulfatase type VIII from abalone entrails (Sigma, S9754) was evaluated. This enzyme was used at a final concentration of 242 U/mL, with DDAO-sulfate [1-100 μM]. For both *H. pomatia* and abalone entrails, Michaelis-Menton kinetics, velocity = (v_{max} * [substrate] / {Km + [substrate]}, was used to fit kinetic data. A summary of results are shown in **Table S1**.

Table S1: Kinetic Parameters for DDAO-sulfate with various enzymes

	*Aerobacter aerogenes*ϕ	Abalone Entrails	*Patella vulgata*†	*Helix pomatia*
K_m (μM)	164 ± 14	29 ± 3	54 ± .00008	13 ± 3
v_{max} (au/sec)	16.88 ± 0.48	1.04 ± .03	0.29 ± 0.04	2.05 ± 0.13
		pH 7.5	pH 5	

ϕ The K_m values are higher for commercial substrates *p*NPS (1800 ± 120 μM) and 4-MUS (700 160 μM).

† Substrate Inhibition observed.

Bacterial Strains and Culture

M. smegmatis (mc^{2}155, ATCC 70084), *M. marinum* (ATCC BAA-535), *M. flavescens* (ATCC 14474), *M. intracellulare* (ATCC 35847), *M. nonchromogenicum* (ATCC 19530), *M. peregrinum* (ATCC 14467), *M. scrofulaceum* (ATCC 19981, strain L2238), and *M. septicum* (ATCC 700731) were purchased from American Type Tissue Culture Collection (ATCC). These species were all treated as BioSafety Level 2 pathogens.

M. africanum, *M. bovis* (BCG), *M. avium*, *M. microti*, M. simiae, *M. kansasii*, and *M. tuberculosis* (*M. tb.* strains: H37Ra, H37Rv, CDC1551, Erdman) were obtained from Prof. Lee Riley (UC Berkeley). Lineage-specific strains of *M. tuberculosis* were obtained from Prof. Midori Kato-Maeda (UC San Francisco), and are listed in **Table S2**. Clinically-derived strains were from University of Stellenbosch in South Africa, and are listed in **Table S3**. The strains *M. tb* H37Rv-ΔFGE and H37Rv-Complement (ΔFGE[Rv0712] + pMV306.kan-FGE) have been previously described(4). These species were all treated as BioSafety Level 3 pathogens.

M. tuberculosis strains, *M. africanum*, *M. avium*, *M. bovis*, and *M. kansasii* were grown at 37°C in Middlebrook 7H9 broth or on 7H11 agar supplemented with 0.5% glycerol, 0.05% Tween-80, and 10% OADC. Other species were grown at 30 °C (*M. marinum*, *M. septicum*, *M. intracellulare*) or 37°C (*M. flavescens*, *M. nonchromogenicum*, *M. peregrinum*, *M. scrofulaceum*, *M. smegmatis*) in 7H9 broth or 7H10 agar supplemented with 0.5% glycerol, 0.05% Tween-80, and 10% OADC. Cultures were grown to log phase ($0D_{600}$ = 0.5 to 1.5) before harvesting by centrifugation at 4°C. The supernatant was discarded and the pellet was frozen (-20°C).

Table S2: *M. tb.* Strains from UCSF

UID	TBID	Octal Code	Lineage
X005761	TB1100056	777777777700771	EuAm
X005772	TB1100354	757777777720771	EuAm
X005752	TB1003453	777777763760771	EuAm
X005748	TB1003253	777777377560771	EuAm
X005711	TB1002014	000000000003771	East Asian
X005751	TB1003244	000000000003771	East Asian
X005705	TB1002173	000000000003771	East Asian
X005692	TB1001718	674777477411771	IndoOceanic
X005696	TB1001833	677777477413771	IndoOceanic
X005720	TB1002311	674000003413771	IndoOceanic

Table S3: *M. tb.* Strains from University of Stellenbosch

Strain Family	**strain #**	**Lineage**
EAI	1659	Indo-Oceanic
CAS	974	India-East Africa
Atypical Beijing	2701	East Asian
Typical Beijing	1125	East Asian
F15 (KZN LAM strain)	2576	Euro-American
F11 (LAM)	3380	Euro-American
Haarlem	405	Euro-American
Haarlem - like	1127	Euro-American
LCC	198	Euro-American
EAI	2493	Indo-Oceanic
EAI	1146	Indo-Oceanic
CAS	630	India-East Africa
CAS	2666	India-East Africa
LAM	2000	Euro-American
Haarlem	5395	Euro-American
LCC - 2 bander	2073	Euro-American
F12	300	Euro-American
Typical Beijing	4097	East Asian
Typical Beijing	1116	East Asian

Tissue Culture

Mammalian cells were purchased from ATCC and maintained in a 37°C, 5% CO_2 humidified incubator chamber. Cells were grown in standard medium [e.g., Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% (v/v) fetal bovine serum

(Invitrogen), 50 U/mL penicillin, and 50 μg/mL streptomycin (DMEM++). Near-confluent cells were passaged with 0.05% trypsin in 0.52 mM EDTA (Invitrogen).

Preparation of Clarified Lysates

Mycobacterial cell pellets were put in Lysis Buffer [50 mM Tris-Cl (pH 7 at 4°C), 100 mM NaCl, 0.5 mM $CaCl_2$, 0.5 mM $MgCl_2$, plus an EDTA-free Protease Inhibitor Tablet (Roche)]. Cells were lysed by mechanical disruption using 0.1 mm zirconia/silica beads (BioSpec Products) on a FastPrep (MP Biomedicals) bead-beating instrument. Crude lysates was clarified by centrifugation (10 min, 16,000*g*, 4°C). Mammalian cells were processed using the same procedure. Mycobacterial cell pellets from BioSafety Level 3 pathogens were lysed in Lysis Buffer supplemented with 0.1% Triton X-100 and sterilized by filtration twice through a 0.2 μm filter (Acrodisc PVDF 13 mm syringe filters).

Total protein concentrations of clarified lysates were determined using a Pierce BCA protein assay kit (Thermo Scientific). Lysates were stored in Lysis Buffer supplemented with 5% glycerol at -80°C until use.

96-well Plate Sulfatase Assay

Mycobacterial lysates (5 μg) were analyzed for sulfatase activity with either 25 μM DDAO-sulfate or 1 mM 4-MUS in Lysate Reaction Buffer [LR Buffer: 50 mM Tris (pH 7.5 at 37°C), 100 mM NaCl, 250 μM $MnCl_2$, $MgCl_2$, and $CaCl_2$]. Samples were prepared in triplicate, and reactions were initiated by addition of sulfated probe. Species analyzed included *M. tuberculosis* (H37Rv, Erdman, and CDC1551), *M. bovis* (BCG), *M. kansasii*, *M. avium*, *M. intracellulare*, *M. africanum*, *M. smegmatis*, *M. marinum*, *M. nonchromogenicum*, and *M. septicum*. Bovine serum albumin (Pierce) was included as a sulfatase-free control.

Fluorescence of hydrolyzed probe was measured at various time points, including 10 m, 30 m, 60 m, and 25 h. For reactions with DDAO-sulfate, fluorescence of DDAO was measured (excitation 635 nm, emission 675 nm bp 665). After 10 m, all lysates had measurable DDAO formation compared to solutions containing BSA (**Figure S3**). For reactions with 4-MUS, fluorescence of 4MU was measured (ex. 390 nm, em. 450 nm bp 435). After 30 m, all lysates had measurable 4MU formation compared to solutions containing BSA. Although 4MU fluorescence is enhanced by basification (> pH 7.8), fluorescence was directly monitored at pH 7.5.

For most applications, it is important that DDAO-sulfate remain stable over time in aqueous buffer at 37°C. The production of DDAO from DDAO-sulfate (100 μM) was detected after 23 h incubation with mycobacterial lysates (2 μg) in LR Buffer. Samples

were prepared in triplicate, and species analyzed included *M. marinum*, *M. flavescens*, *M. nonchromogenicum*, *M. tb,* (H37Rv), *M. tb.* (CDC1551), *M. africanum*, *M. kansasii*, and *M. bovis*. A lysate-free (buffer) control was included to measure background hydrolysis of DDAO-sulfate. After 24 h, the amount of DDAO produced from enzyme-free hydrolytic cleavage was less than 2% of the amount produced in the presence of mycobacterial lysates (**Figure S12**).

Lysates from wildtype *M. tb.* H37Rv, H37Rv-ΔFGE, and H37Rv-Complement (ΔFGE[Rv0712] + pMV306.kan-FGE) were also analyzed. For these samples, 10 μg of total protein lysates were used for each reaction. As observed previously(4), the FGE knockout strain has diminished, but measurable, sulfatase activity. Using an *in vitro* assay, we confirmed that the knockout (ΔFGE) strain does have residual activity following overnight incubation (25 h) with 25 μM DDAO-sulfate or 1 mM 4-MUS (**Figure S13**). This activity most likely results from Type III sulfatases. In this experiment, the amount of DDAO produced in samples containing BSA, but lacking mycobacterial sulfatases, was 1% of that produced by samples containing wildtype *M. tb.* H37Rv.

In-gel Activity Assay

Mycobacterial lysates (1-20 μg) were resolved by native gel electrophoresis (4-15% Tris-Cl Criterion gel, Bio-Rad) in 1X Tris-Glycine buffer. Protein loading dye did not contain DTT or SDS. Both gradient (4-15%) and 7.5% gels were used to resolve lysates; the latter type provides better resolution of higher bands, including the AtsB bands observed in *M. tb.* Gels were run at 180V for 55 m on top of ice before soaking in LR Buffer with 10 μM DDAO-sulfate. Fluorescence of DDAO was detected on a fluorescence scanner (Typhoon 9410 Variable Mode Imager, GE Healthcare; excitation 633 nm, emission 670 bp 30, 100 μm resolution.) Images were analyzed and false colored (Green-Fire-Blue) in ImageJ(5).

As mentioned in the article, additional lysates from *M. tb.* strains were evaluated. The fluorescence image of this gel is included as **Figure S7**.

We prepared protein lysates from wild-type (WT), knockout (ΔFGE), and the complemented (+FGE) strains. Fluorescent bands were observed for the WT and complemented (+FGE) lysates, but were absent in the ΔFGE strain-derived lysate (**Figure S2**). From this experiment, we conclude that the sulfatase bands we observe in gel-resolved *M. tb.* H37Rv lysates are FGE-dependent.

In-gel Activity Assay with Mammalian Lysates

Lysates from mammalian cells were analyzed using the in-gel assay (**Figure S4**). For this experiment, we resolved lysates from 7 different cell lines (i.e., COS-7, Rat-1

fibroblasts, BT549, CHO, HEK293, RAW macrophages, and Jurkats). For comparison, mycobacterial lysates were also resolved on the same native protein gel. The gel was exposed to 10 μM DDAO-sulfate before imaging. None of the mammalian cell lysates produced fluorescent bands.

Furthermore, we examined sulfatase activity in *M. tb*-infected RAW macrophages. Briefly, macrophages (4 x 10^7 cells/flask) in RPMI were infected with *M. tb.* (4 x 10^8 bacteria/flask; MOI = 10) for 4 h. *M. tb.* H37Rv, *M. tb.* Erdman, and *M. tb.* CDC1551 were each used for infections. For mock-infection, a small volume of PBS (vehicle) was added to macrophages for 4 h. After infection, cells were washed twice with warm PBS. For an initial time point, cells were collected immediately following infection. For later time points (1d, 3d, or 7d), cells were placed in fresh RPMI medium supplemented with horse serum (10%), Hepes (10 mM) and 25 μg/mL amikacin. At each time point, cells were collected in PBS by scraping the flask surface. Cells were pelleted by centrifugation and placed in LR Buffer supplemented with 0.1% Triton X-100. Each sample was frozen (- 20°C) prior to lysis by mechanical disruption. Sterile-filtered lysates were concentrated (Millipore Biomax, MWCO 5 kDa) and quantified using a BCA assay (Pierce). Complex lysates (e.g., containing mycobacterial and macrophage proteins) were resolved by gel electrophoresis, exposed to 25 μM DDAO-sulfate, and imaged (**Figures S4B**). Mock-infected macrophages did not produce fluorescent bands. In contrast, complex lysates from *M. tb.* infected macrophages did produce fluorescent bands corresponding to sulfatase activity.

DDAO-sulfate Comparison with *p*-NPS, X-Sulf and 4-MUS

In order to compare sulfatase-activated probes, we prepared four identical native gels (4-15% Tris-Cl) with different mycobacterial lysates (50-100 μg/lane). After separating the lysates by electrophoresis (180V, 55 m, on ice), gels were soaked in LR buffer with each probe. DDAO-sulfate was used at a final concentration of 25 μM. 4-MUS was used at a final concentration of 1 mM. The two chromogenic substrates, *p*-NPS and 5-bromo-4-chloro-3-indoyl sulfate (X-Sulf) were used at a final concentration of 10 mM. Gels were imaged at several time points, and all had distinct sulfatase bands after 15 m (**Figure S5**). DDAO was imaged on a Typhoon fluorescence scanner. The other three gels were imaged on a FluorChem M fluorescence gel scanner (ProteinSimple). The gel treated with X-Sulf was imaged using the trans-UV light source and a 593/40 nm emission filter; the hydrolysis product formed a tight blue band in the gel which was stable for days. 4-MU formation was imaged using the trans-UV light source and a 537/26 nm emission filter. The trans-UV light source was used for imaging the hydrolysis of *p*-NPS, but no emission filter was used. The hydrolysis of *p*-NPS gave a yellow product that rapidly diffused through and out of the gel.

Expression of AtsB in *M. smegmatis*

AtsB (Rv3299) was amplified from *M. tb.* (H37Rv) genomic DNA and ligated into a pET28b vector (Novagen) using NdeI and HindIII sites. The gene was cut from this vector and inserted into a pMV261 plasmid, which had been modified by site-directed mutagenesis (Qiagen) to contain an Nde1 site. The finished plasmid, pMV261-AtsB, was transformed into *M. smegmatis* (mc^2155, ATCC 70084) by electroporation to give "AtsB-*M.smeg*".

Cultures of wild-type *M. smegmatis* or AtsB-*M. smeg*. were grown in 7H9 broth for 17 h, or until the $0D_{600}$ = 1.4 to 2.1, and harvested by centrifugation at 4°C. The supernatant was discarded and the pellet was frozen. The thawed pellet was subsequently lysed in Lysis Buffer by mechanical disruption and clarified by centrifugation. Total protein lysates from two clones of AtsB-*M. smeg.*, wild-type *M. smegmatis*, and *M. tb.* (CDC1551, Erdman, and H37Rv) were separated by native gel electrophoresis (7.5% Tris-Cl Criterion gel, Bio-Rad) in 1X Tris-Glycine buffer. The gel was run at 180V for 60 min before soaking in LR Buffer with 25 μM DDAO-sulfate. Fluorescence of DDAO was detected on a fluorescence scanner as described above. The additional sulfatase bands observed in lysates derived from transformed *M. smegmatis* are similar to the banding observed in *M. tb.*, providing support that the enzyme AtsB produces more than one fluorescent band in the in-gel assay(**Figure S6**).

In-gel assay of *M. tb.* CDC1551 Transposon Mutants

We used *M. tb.* CDC1551 transposon mutants in order to determine which sulfatases produce fluorescent bands in our in-gel assay. We received transposon mutants as part of NIAID Contract No. HHSN266200400091C, entitled "Tuberculosis Vaccine Testing and Research Materials," which was awarded to Colorado State University. We obtained four strains, including *M. tb.* CDC1551 transposon mutant 2379 [MT0738, Rv0711 (AtsA)], 333 [MT3398, Rv3299c (AtsB)], 3299 [MT3162, Rv3077 (AtsF)], and 370 [MT0310, Rv0296c (AtsG)]. Each mutant was grown, lysed, and analyzed on a native gel (7.5% Tris-Cl) as described above for other samples.

The lysates obtained from wild-type CDC1551 and three of the transposon mutants (AtsA, AtsF, and AtsG) each produced similar banding patterns (**Figure S6**). In contrast, the AtsB transposon mutant lacks all bands associated with *M. tb*. sulfatase activity. This observation is consistent with AtsB producing the group of fluorescent bands observed for *M. tb.* strains H37Rv, CDC1551, and Erdman.

Evaluation of Sulfatase Inhibitors

We evaluated a variety of enzyme inhibitors with mycobacterial lysates, including a set of arylsulfamates. (Structures are shown in **Figure S14**). Arylsulfamates were all obtained from Sigma. Each arylsulfamate (500 μM) was added to clarified lysates (10 μg). Reaction buffer contained 50 mM Tris (pH 7.5 at 37°C) with 50 mM NaCl, with no metals added. Phosphatase inhibitor cocktails 1 and 2 (Sigma) were used *in combination* at 4% (volume:volume). Two metal chelators were also evaluated, TPEN and 1,10-Phenanthroline monohydrate, both from Sigma. TPEN was used at 2 mM and 1,10-Phenanthroline monohydrate was used at 4 mM final concentration. Lysates were incubated with each compound for 1 hour at 37°C. Protein loading dye was added to each reaction and samples were loaded onto a native gel (4-15% Tris-Cl) and resolved (180V, 60 min, room temperature). Gels were reacted in 100 mM Tris Buffer (pH 7.5 at 37°C), 100 mM NaCl with 10 μM DDAO-sulfate. The formation of DDAO was imaged on a fluorescence gel scanner as described. Treatment of lysates with phosphatase inhibitor cocktails 1 and 2 did not affect sulfatase banding patterns.

Identification of Proteins by Mass Spectrometry

Crude lysates from 100 to 200 mL cultures of *M. kansasii*, *M. avium*, *M. bovis* (BCG) or *M. tuberculosis* strains (CDC1551, Erdman, and H37Rv) were prepared as described above. The lysates was concentrated (MWCO 10 KDa Amicon Ultra, Millipore) and exchanged into NaCl-free buffer [50 mM Tris-Cl (pH 7 at 4°C), 1 mM $MgCl_2$, 0.5 mM $CaCl_2$, 0.5 mM DTT]. Sulfatases were partially enriched using a strong anion exchange column (HiTrap Q FF 5 mL, GE Healthcare). Concentrated lysates were loaded onto the column and proteins were eluted using a step gradient from 0 M NaCl to 1 M NaCl. Sulfatases eluted between 200 and 400 mM NaCl. The stained protein gel from partially enriched lysate fractions from *M. avium* and *M. kansasii* is shown in **Figure S15**.

Sulfatase-active fractions were concentrated (MWCO 10 KDa) and resolved by native gel electrophoresis (7.5% Tris-Cl Criterion gel). The gel was soaked in 10 μM DDAO-sulfate in LR Buffer and imaged. Sulfatase bands were excised with a clean blade and submitted for extraction, tryptic digest, and LC-MS/MS analysis at the UC Berkeley QB3 Vincent J. Coates Proteomics and Mass Spectrometry Laboratory. Data collection was by CID fragmentation on the top five peaks, and processing was done using SEQUEST. Matches required two or more peptides per protein. Each identified sulfatase is indicated in **Figure 4**, and the identified peptides, with XCorr and DeltaCN values, are provided in **Table S4**.

Table S4: Identification of Sulfatases by Mass Spectrometry

Species	Sulfatase ID: Percent Coverage	Peptide Sequence	XCorr	DeltaCN
M. tuberculosis (H37Rv)	AtsB (Rv3299): 3.7%	R.LAQNGLIYNR.F	3.1854	0.4716
		R.DNGYVTGAFGK.W	3.1771	0.5711
		R.QVSSEPLPTGDVTVR.M	3.7077	0.6253
	AtsB (Rv3299): 3.4%	K.DAEGSPVLVDTGNR.L	4.5392	0.696
		R.DNGYVTGAFGK.W	3.2285	0.6628
		R.VTLWADDR.L	2.557	0.5096
M. tuberculosis (Erdman)	ZnSulfB (Rv2407): 8.9%	K.DIVTQIPQQR.V	2.1861	0.4395
		K.DIVTQIPQQR.V	4.0058	0.427
		R.ALAATEFSGR.I	1.949	0.4011
	AtsG (Rv0296c): 13.3%	R.YLGVYHHPDVYSPR.L	3.638	0.5768
		R.YRPADSAAVELPDYLPDTPEVR.Q	5.1201	0.518
		R.QDVAEFYGSIATADEAVGR.L	5.0788	0.6861
		R.YTETYLR.I	2.3748	0.561
M. bovis (BCG)	AtsB (Rv3299): 7.3%	R.HDFQTLVDAAK.D	4.1944	0.6929
		K.IGQALAAGR.A	2.6829	0.4847
		R.LEAGQALASSPMK.S	3.1322	0.6252
		R.VLPLLGGLAVMFGDLPPLPTTAR.F	4.5609	0.6408
		R.M*LFDSHQPVAASGGR.V	2.2382	0.4529
	AtsD(Rv0663): 13%	R.TLKLNGYNTAQFGK.C	2.737	0.3548
		K.LNGYNTAQFGK.C	2.863	0.6727
		R.TPEEGYHFMADMTDK.A	4.3143	0.7346
		R.LWLIEATR.Y	2.5958	0.5051
		R.YNVLPLDDDTASR.I	2.9947	0.4687
		R.GNTQVLFSNMGR.L	3.0279	0.5056
		R.GNTQVLFSNM*GR.L	3.4603	05132
		R.LSENCVLNLK.N	2.0754	0.4538
		R.LSENCVLNLK.N	3.4942	0.523
		K.HFYAESADPLPAGAHQVR.M	5.6359	0.7879
		R.MEFAYAGGGLGK.G	2.4877	0.6198
		R.MEFAYAGGGLGK.G	4.2188	0.6087
		R.M*EFAYAGGGLGK.G	2.9429	0.5234
M. kansasii	ZnSulfB (Rv2407): 6.6%	-.MLEITLLGTGSPIPDPDR.A	4.2264	0.6155
		-.M*LEITLLGTGSPIPDPDR.A	3.6867	0.5575
	AtsA (Rv0711): 10%	R.NATTVGMATIEEFTDGFPNCNGR.I	4.2071	0.7356
		R.MAEVFAGFLSYTDAQIGR.I	2.4536	0.5035
		K.LFDHLGGPQTYNHYPIGWAMAFNTPYK.L	6.1876	0.7166
		K.LFDHLGGPQTYNHYPIGWAM*AFNTPYK.L	4.4769	0.6035
		R.LHYVYNFLGER.Q	3.2923	0.6412
		R.LHYVYNFLGER.Q	3.2937	0.5694
M. avium	AtsA (Rv0711): 8%	R.ASLLTGR.N	2.2059	0.5019
		K.DIADKTIEFIR.D	3.4666	0.5667
		R.FDMGYER.Y	2.3449	0.4479
		R.MAEVFAGFLSYTDAQIGR.I	4.7602	0.6681
		K.ALWFSEAAK.Y	1.937	0.3858
		K.ALWFSEAAK.Y	3.008	0.496
		R.LHYVYNFLGER.Q	3.2438	0.603

Genomic Analysis

NCBI BlastP (protein-protein BLAST) searches were performed using the amino acid sequence for each sulfatase encoded in *M. tb.* H37Rv. The search against non-redundant protein sequences (nr) was restricted to organisms designated "mycobacteria". After identifying potential sulfatases in mycobacterial genomes, protein sequence alignments were done using NCBI's Cobalt (Constraint-based multiple protein alignment tool). Fasta files were viewed using Jalview (6). Alignments of a portion of each sulfatase, including the regions near active site residues, are provided as images at the end of the SI Appendix. The Type I sulfatase alignments are shown on pages 25-26. The Type II sulfatase alignments for Rv3406 (FeSulf-A) are shown on page 27. The Type III sulfatases are shown on 28. Fasta files for each alignment are available upon request from the corresponding author (crb@berkeley.edu).

Accession codes for the sulfatases identified from each species are shown in **Table S5**.

Table S5: Accession Codes for Sulfatases

	AtsA	AtsB	AtsD	AtsF	AtsG	FeSulf-A	ZnSulf-A	ZnSulf-B	ZnSulf-C
M.tb. H37Rv (Rv Number)	Rv0711	Rv3299	Rv0663	Rv3077	Rv0296c	Rv3406	Rv3796	Rv2407	Rv3762C
M.tb. H37Rv	NP_215225.1	NP_217816.1	NP_215177.1	YP_177923.1	YP_177712.1	NP_217923.1	YP_178016.1	NP_216923.1	NP_218279.1
M.tb. CDC1551	NP_335155.1	NP_337927.1	NP_335103	NP_337684.1	NP_334719.1	NP_338038	NP_338455.1	NP_336961	NP_338419.1
M. smegmatis	YP_885833.1			YP_885676.1	YP_885620.1	YP_888161.1		YP_888836.1	YP_890557.1
M. marinum	YP_001849354.1	YP_001849545.1	YP_001853209.1	YP_001849884.1	YP_001848873.1	YP_001849453	YP_001853617	YP_001851998.1	YP_001849914.1
M. bovis (BCG)	YP_976856	NP_856972.1	NP_854340.1	YP_979186.1	NP_853968.1	NP_857080	YP_979937	YP_978512.1	YP_979900.1
M. canetti	YP_004744184.1	YP_004746733.1	YP_004744133.1	YP_004746520.1	YP_004743786.1	YP_004746836.1		YP_004745870.1	YP_004747185.1
M. kansassii	ZP_04746372.1	ZP_04750005.1	ZP_04746586.1	ZP_04747335.1	ZP_04751396.1	ZP_04746502.1		ZP_04748616.1	ZP_04747185.1
M. avium	ZP_05218413.1					ZP_05218298.1		ZP_05216109.1	ZP_05214833.1
M. intracellulare	ZP_05225468.1				ZP_05226598.1	YP_005339752.1		ZP_05223758.1	ZP_05227463.1
M. africanum		YP_004724949.1	YP_004722375.1	YP_004724727.1	YP_004722013.1	YP_004725053.1		YP_004724076.1	YP_004725398.1

Supporting Figures:

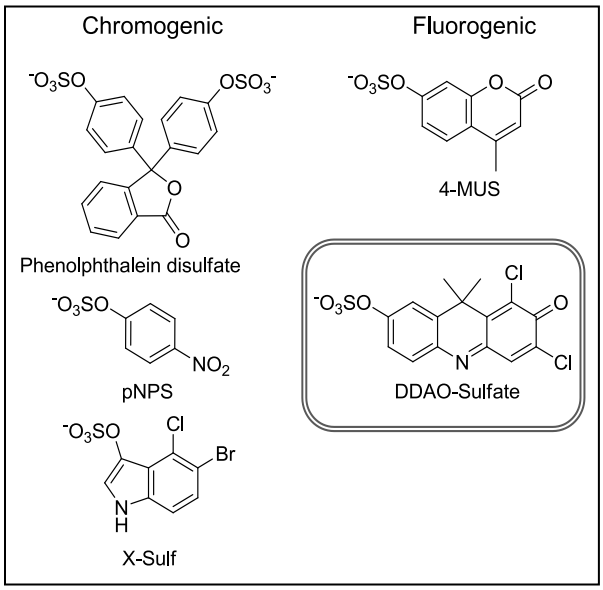


Figure S1. Structures of chromogenic and fluorogenic sulfatase substrates.

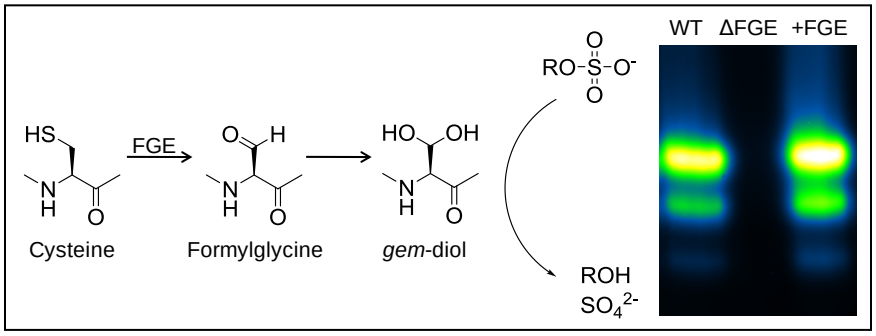
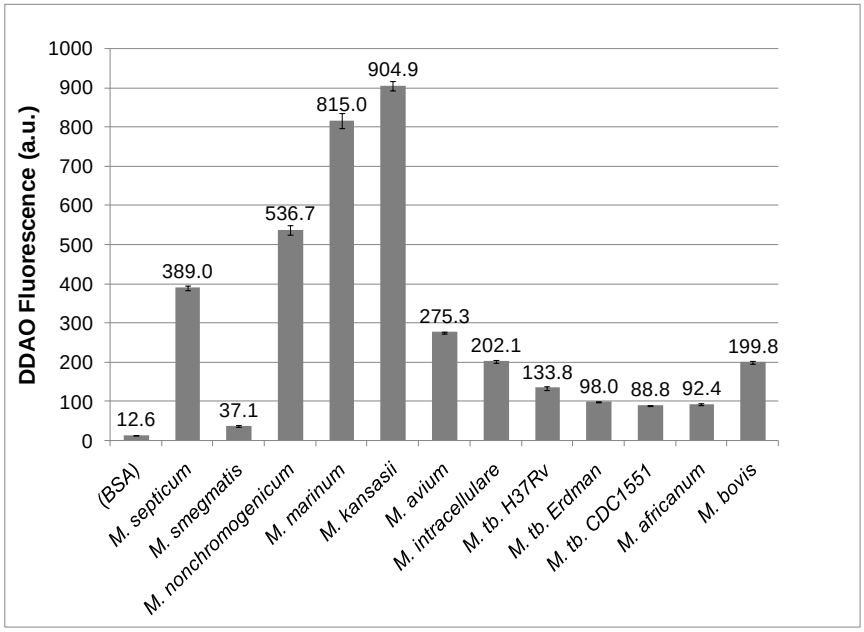


Figure S2. The formylglycine generating enzyme (FGE) is necessary for detecting Type 1 sulfatase activity in *M. tb.* H37Rv lysates. The active site of Type I sulfatases contains a catalytic formylglycine residue, which is formed by a formylglycine generating enzyme (FGE). Hydration to the *gem*-diol produces a hydroxyl that acts as a nucleophile to hydrolyze a sulfate ester. The *M. tb.* H37Rv strain containing a knock-out of the gene encoding FGE (*fge*) is known to have decreased sulfatase activity (4). Lysates from the wild-type strain (WT), the *fge* knock-out strain (ΔFGE), and the complemented strain (+FGE) were resolved by native gel electrophoresis. The protein gel was soaked in a solution of DDAO-sulfate. Fluorescence imaging revealed bands of DDAO fluorescence, which are associated with sulfatase activity found only in the WT and +FGE strains.

A.



B.

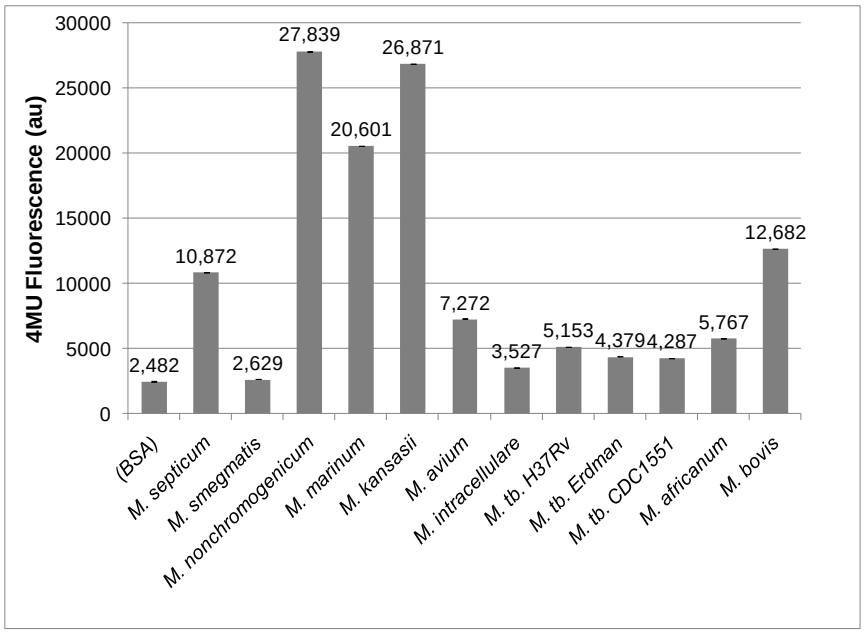


Figure S3. Sulfatase activity in mycobacterial lysates. Mycobacterial lysates (5 μg) were analyzed for sulfatase activity with either 25 μM DDAO-sulfate (A) or 1 mM 4-MUS (B) in LR Buffer. DDAO fluorescence (ex. 635 nm, em. 675 nm) was measured after 10 m, and 4-MU fluorescence (ex. 360 nm, em. 450 nm) was measured after 30 m. Samples were prepared in triplicate and error bars represent one standard deviation.

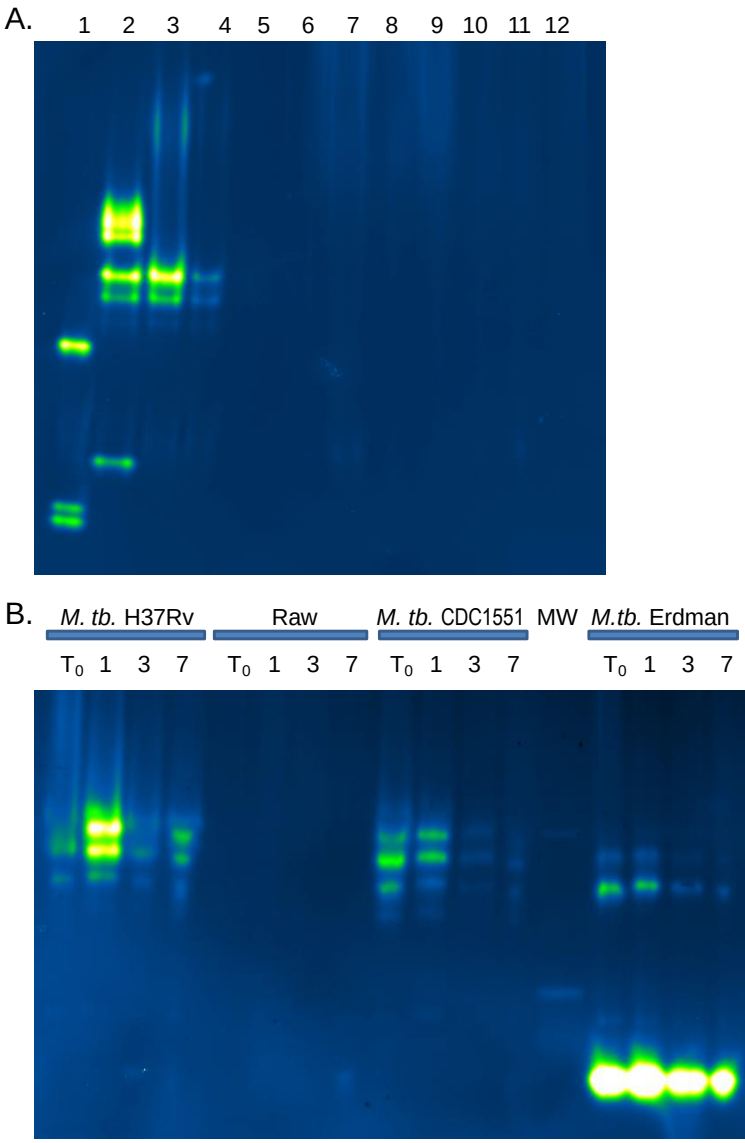


Figure S4. In-gel assay of lysates from mammalian cells. A. Mammalian cells were grown to near confluence before collection and lysis. Samples were loaded onto the gel (Tris-Cl 4-15%) and resolved by electrophoresis. The gel was soaked in a solution of LR Buffer with 10 μM DDAO-sulfate before imaging. Four mycobacterial lysates were loaded onto the gel (lanes 1-4) for comparison. Lane 1: *M. avium*. Lane 2: *M. bovis* (BCG). Lane 3: *M. tb.* H37Rv. Lane 4: *M. tb.* CDC1551. Lane 5 does not contain any sample. Lysates from mammalian cells are in lanes 6-12. Lane 6: COS-7 (83 μg). Lane 7: Rat-1 fibroblasts (105 μg). Lane 8: BT549 (480 μg). Lane 9: CHO (140 μg). Lane 10: HEK293 (410 μg). Lane 11: RAW macrophages (160 μg). Lane 12: Jurkats (240 μg).

B. RAW macrophages were infected with *M. tb.* (MOI = 10). Infected cells were collected, washed, and lysed at different time points (T_o = immediately after 4 h infection; 1 day, 3 days, 7 days). Mock-infected macrophages were also lysed at each time point. For each sample, 28 μg of complex lysate was resolved on a native protein gel before exposure to 25 μM DDAO-sulfate (15 m) and imaging.

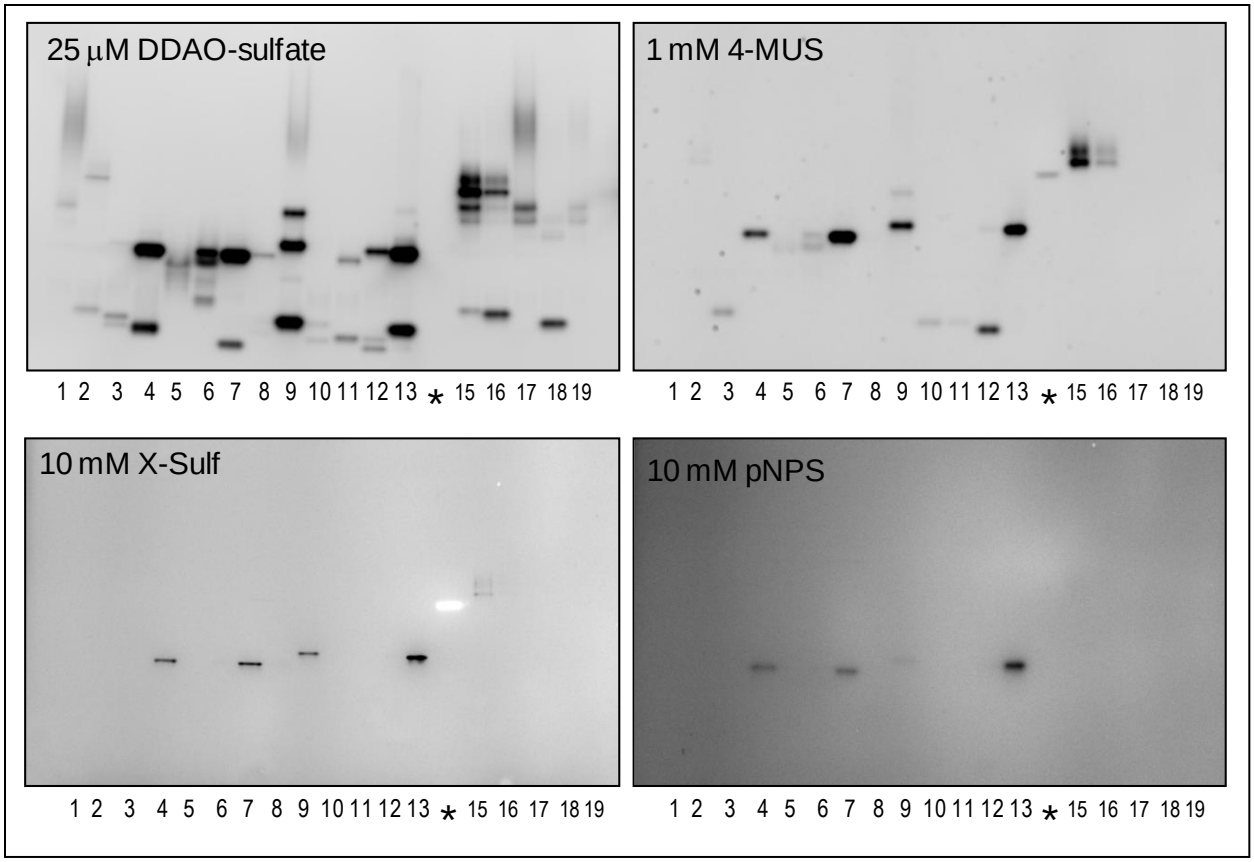


Figure S5. Comparison of activatable sulfatase probes. Mycobacterial lysates from a variety of species were resolved by native protein gel electrophoresis. Each gel was incubated for 15 m with a sulfatase probe (e.g., DDAO-sulfate, 4-MUS, X-Sulf, or *p*-NPS). Gels were imaged to capture signal from each hydrolyzed probe. The lane containing the molecular weight ladder is indicated with an asterisk. In order, the other lanes are **1**: *M. tuberculosis* (H37Ra), **2**: *M. microti*, **3**: *M. simiae*, **4**: *M. flavescens*, **5**: *M. peregrinum*, **6**: *M. septicum*, **7**: *M. nonchromogenicum*, **8**: *M. smegmatis*, **9**: *M. marinum*, **10**: *M. scrofulaceum*, **11**: *M. intracellulare*, **12**: *M. avium*, **13**: *M. kansasii*, (* = MW ladder), **15**: *M. bovis (BCG)*, **16**: *M. africanum*, **17-19**: *M. tuberculosis* (**17**: H37Rv, **18**: Erdman, **19**: CDC1551).

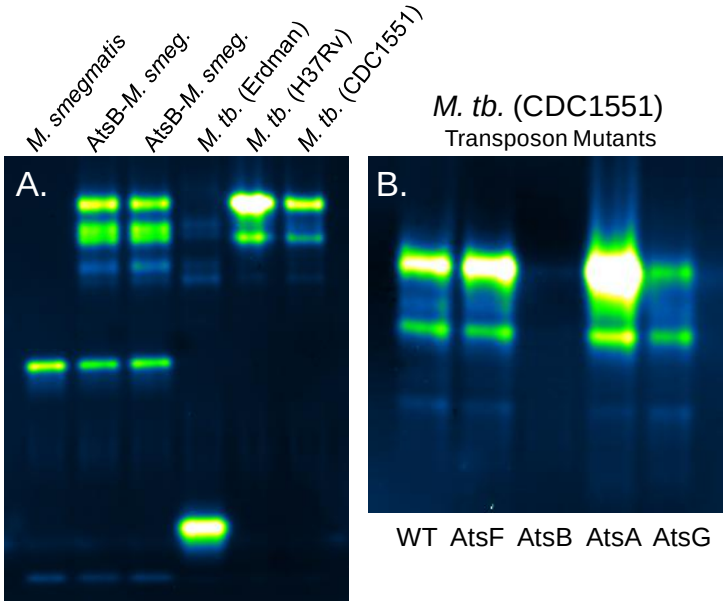


Figure S6. ***M. tb.*** **AtsB activity.** **A.** Plasmid-encoded AtsB (pMV261-AtsB) was expressed heterologously in *M. smegmatis.* We examined lysates derived from two separate clones (AtsB-*M.smeg.*), wildtype *M. smegmatis*, and three *M. tb.* strains (H37Rv, Erdman, and CDC1551). The expression of AtsB results in the appearance of new bands in lysates from AtsB-*M. smeg*. **B.** Protein lysates from wildtype *M.tb.* CDC1551 (WT) and four transposon mutants (AtsA, AtsB, AtsF, and AtsG) were analyzed. Sulfatase banding patterns were similar for all but the AtsB transposon mutant, which lacked sulfatase activity in the in-gel assay.

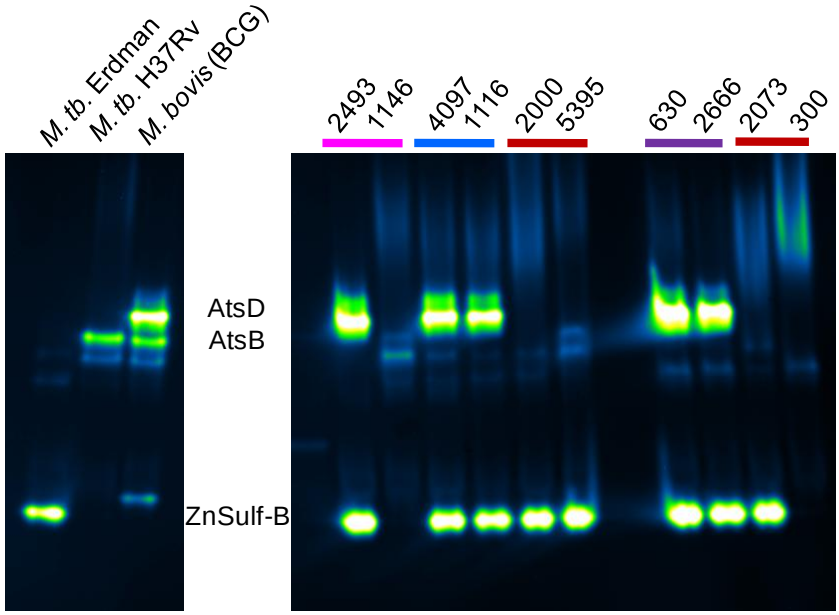


Figure S7. Clinically-derived isolates of *M. tb.* have varying patterns of sulfatase activity. The name of each lineage is color coded: Euro-American (red, 2000, 5395, 2073, 300), Indo-Oceanic (pink, 2493, 1146), East Asian (4097, 1116), and India and East Africa (purple, 630, 2666). *M. tb.* lab strains (Erdman and H37Rv) and *M. bovis* (BCG) were loaded on the same gel for comparison.

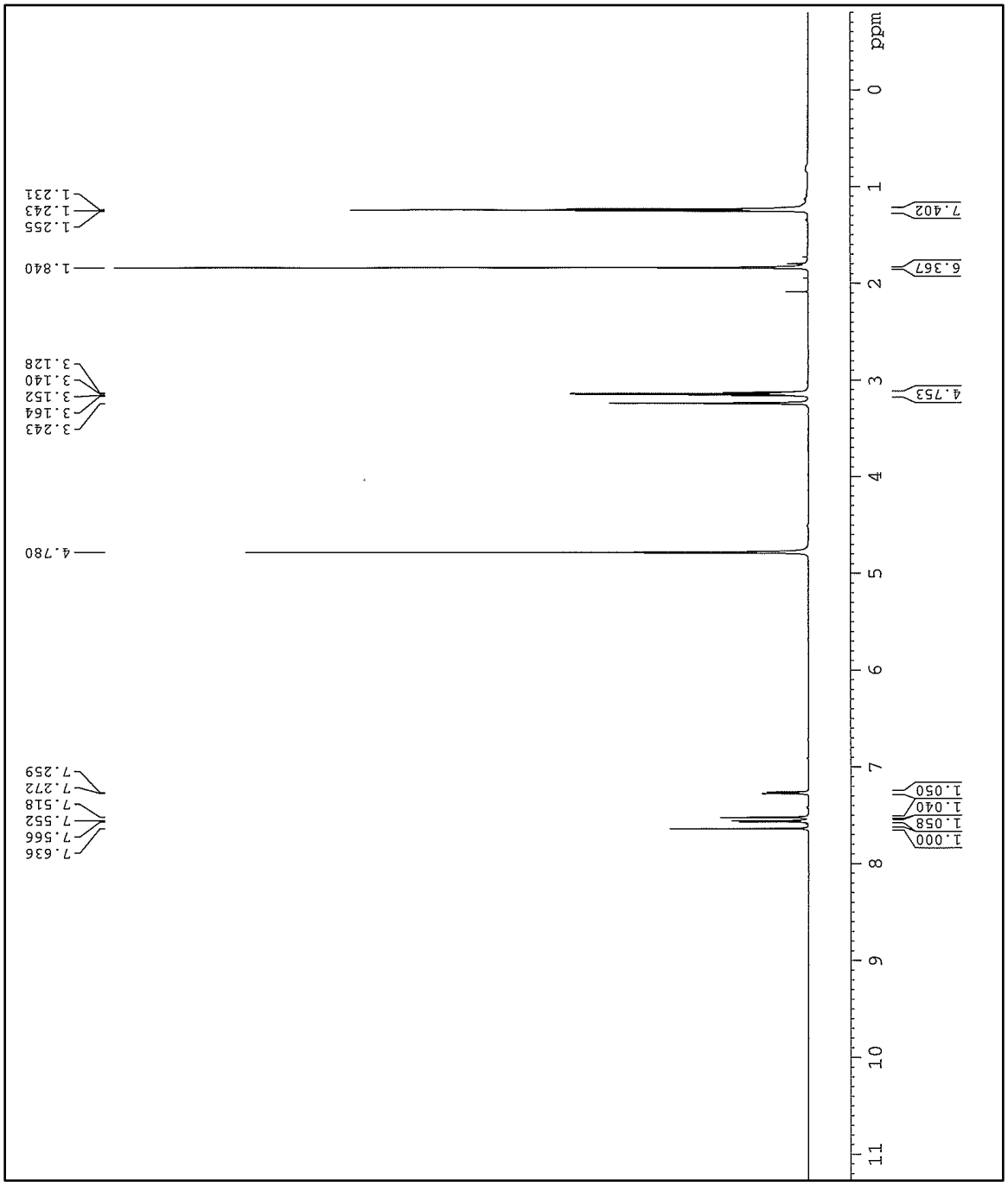


Figure S8. ^{1}H NMR spectrum of DDAO-sulfate.

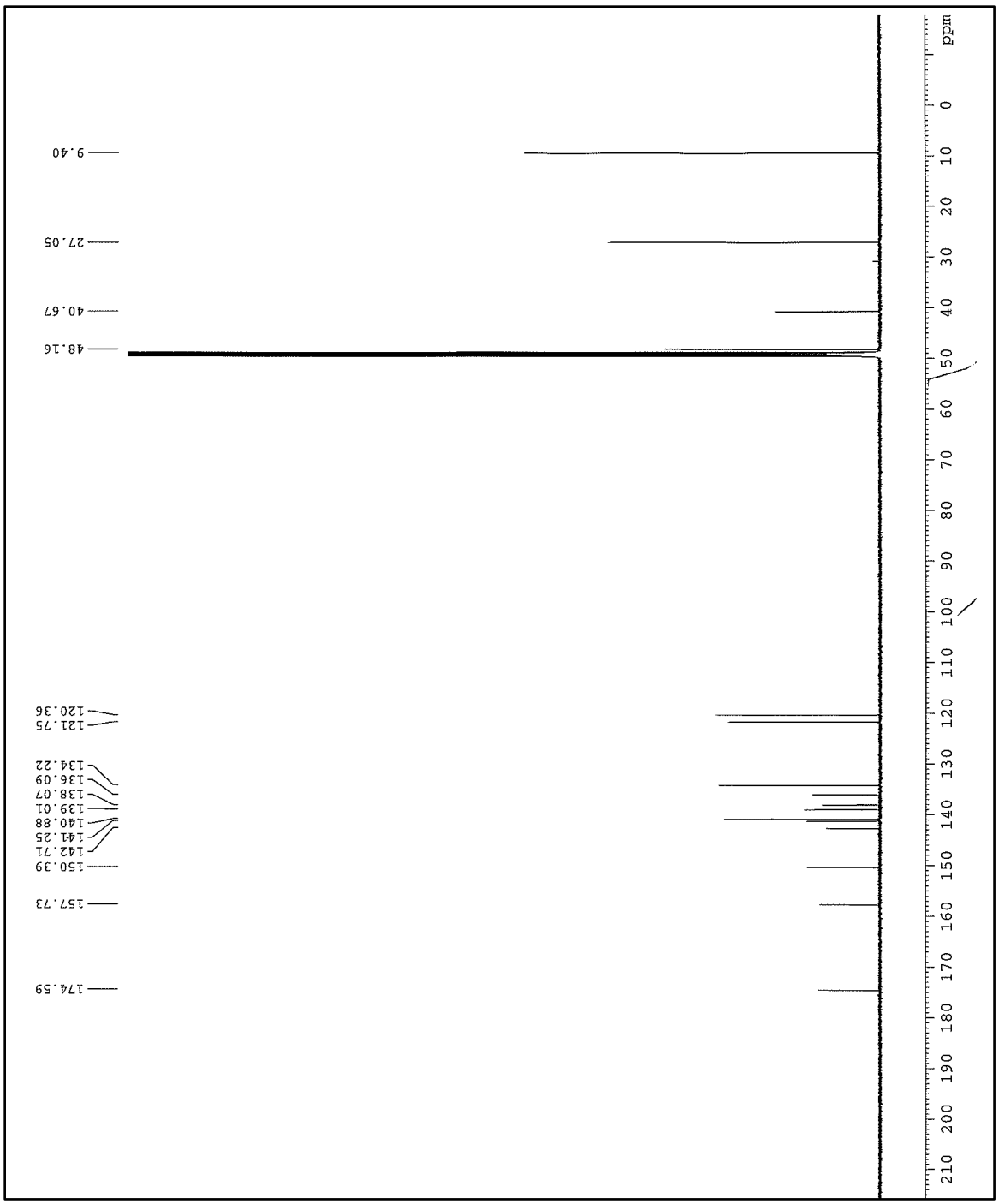


Figure S9. ^{13}C NMR spectrum of DDAO-sulfate.

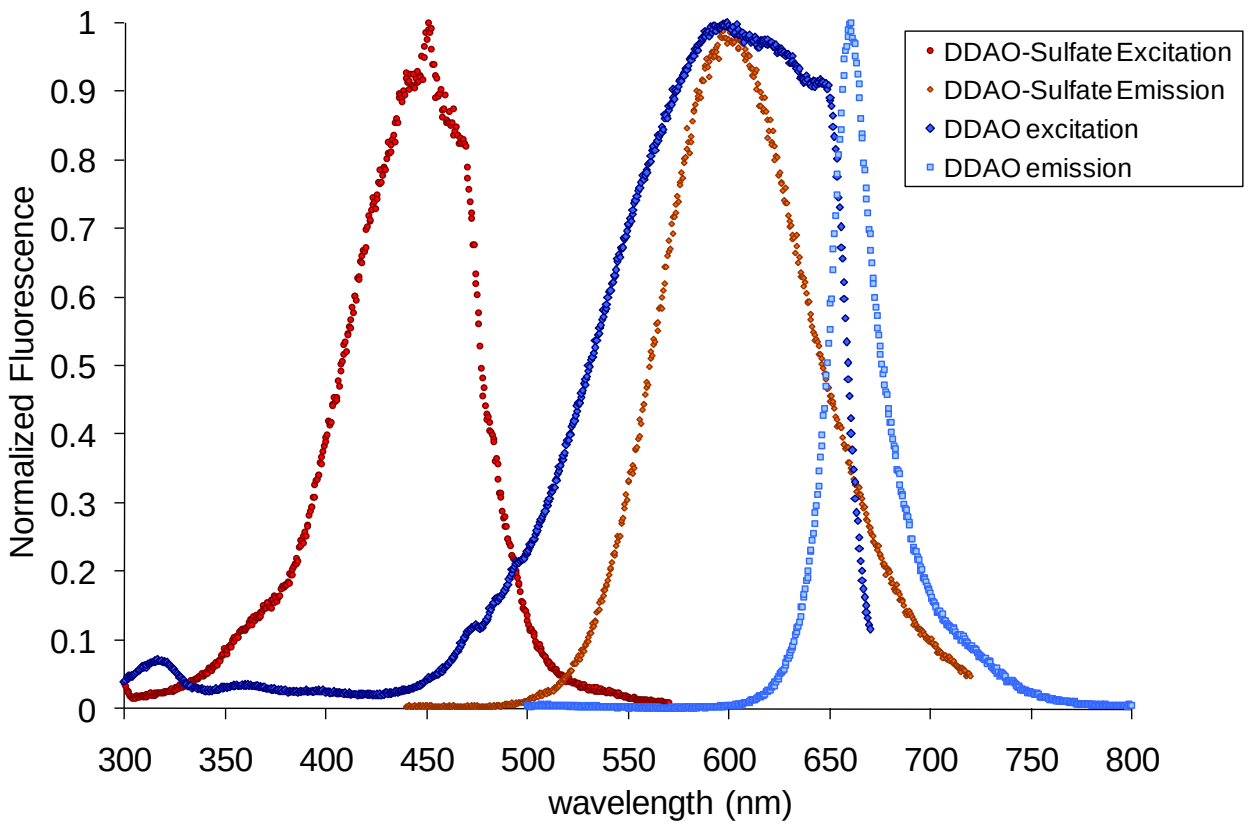


Figure S10. Normalized excitation and emission spectra for DDAO-sulfate and DDAO in aqueous buffer, pH 10.

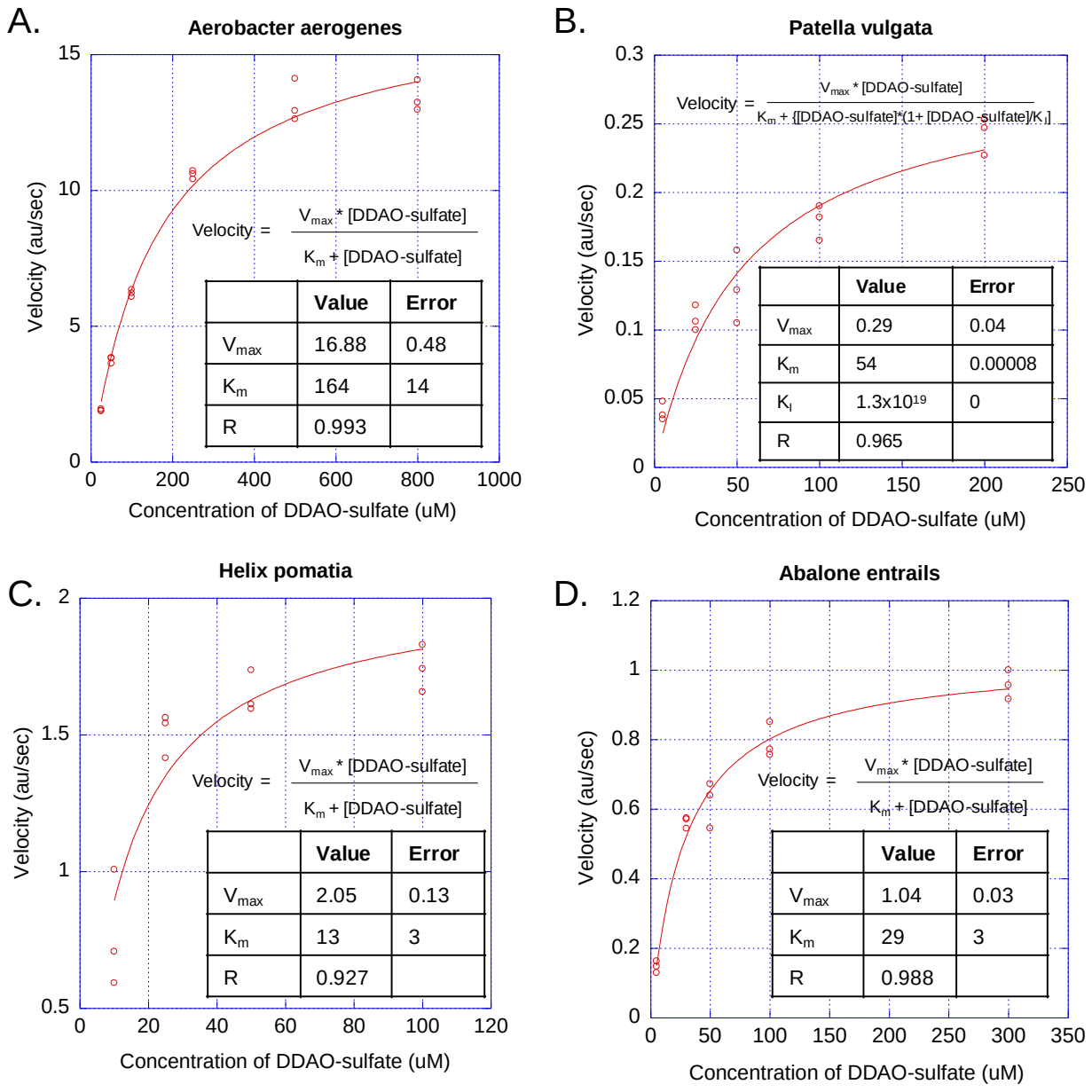


Figure S11. Representative kinetic data for DDAO-sulfate hydrolysis with sulfatases. Varying concentrations of DDAO-sulfate were incubated with each sulfatase, and the reaction progress was monitored by detecting formation of DDAO fluorescence (ex. 635 nm, em. 675 nm). Each plot shows the initial rate as a function of substrate concentration with data fitting. The enzymes sources evaluated were: **A.** *Aerobacter aerogenes*; **B.** *Patella vulgata*; **C.** *Helix pomatia*; **D.** Abalone entrails. Substrate inhibition was observed for *Patella vulgata*, but not for the other enzyme sources.

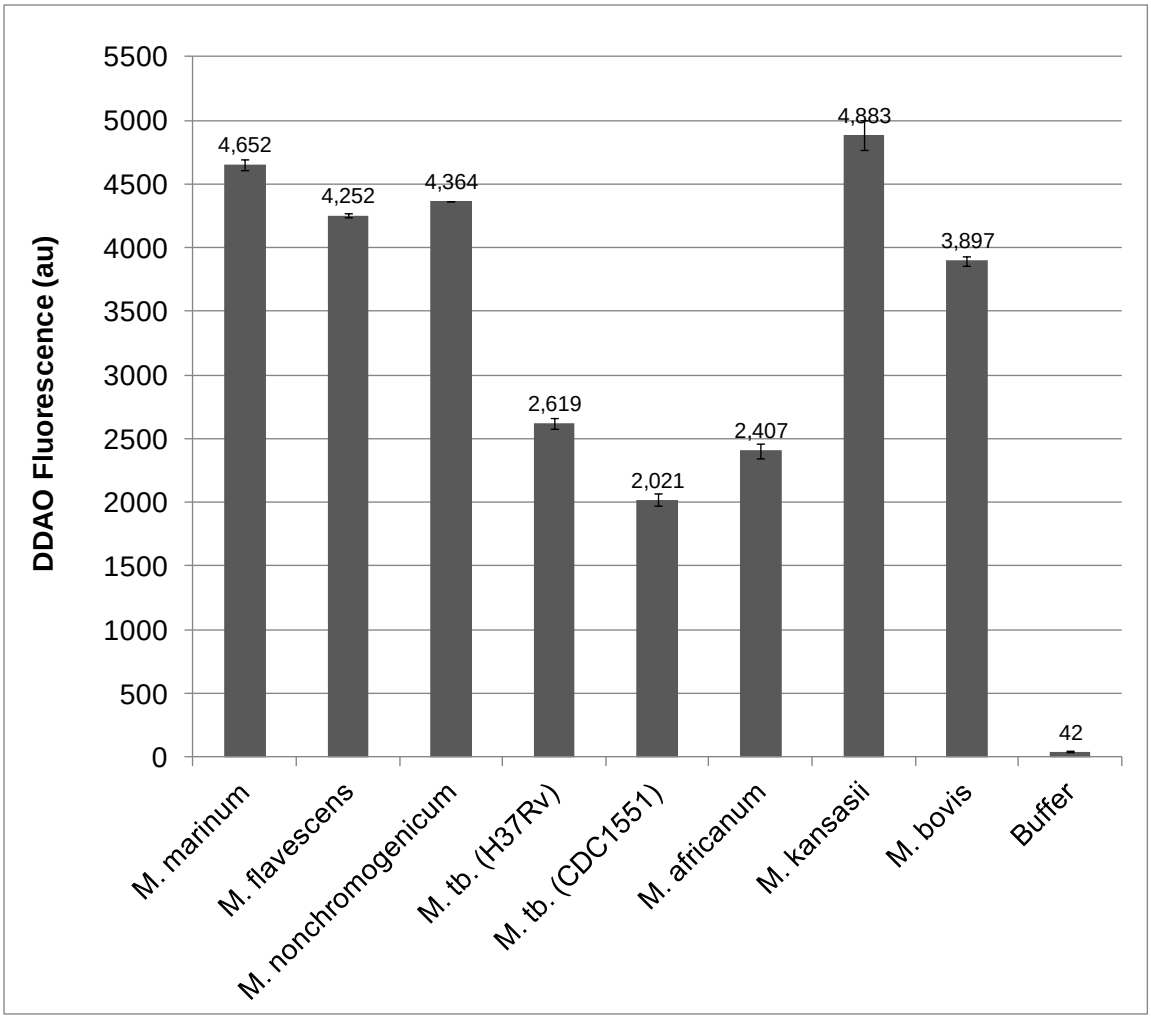
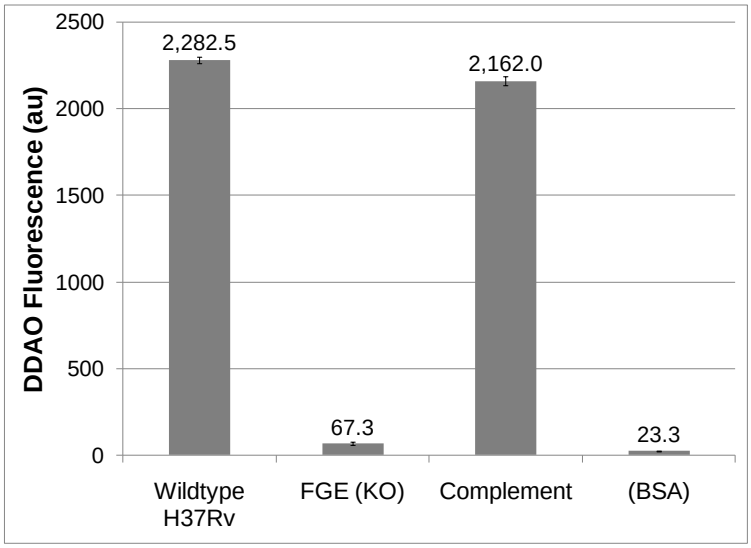


Figure S12. Production of DDAO from DDAO-sulfate. The amount of DDAO produced was detected after a 23 h incubation in LR Buffer at 37 °C with sulfatase-containing mycobacterial lysates (2 μg). This value was compared to the production of DDAO following incubation of DDAO-sulfate in buffer lacking sulfatases. Reactions were prepared in triplicate, and the error bars represent one standard deviation.

A.



B.

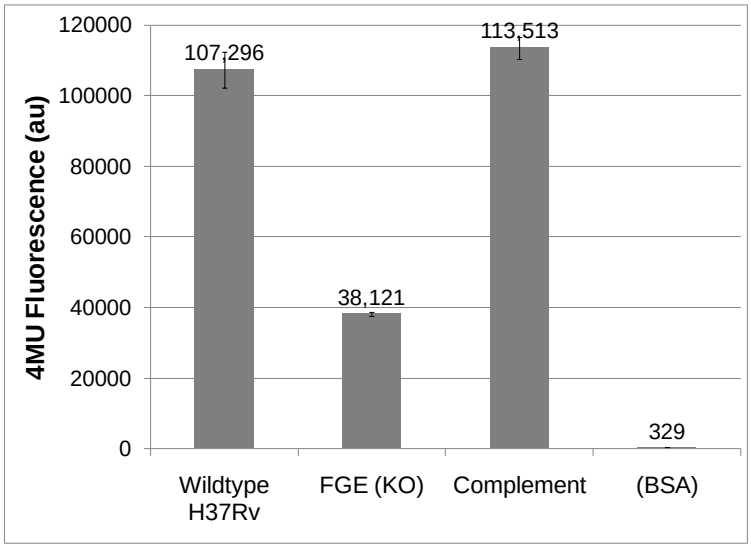


Figure S13. Sulfatase activity in *M. tb.* H37Rv strains. Lysates (10 μg/reaction) from wildtype *M. tb.* H37Rv, *M. tb* H37Rv-ΔFGE ("FGE KO"), and H37Rv-Complement ("Complement"; ΔFGE[Rv0712] + pMV306.kan-FGE) were reacted with either 25 μM DDAO-sulfate or 1 mM 4-MUS in LR Buffer. Fluorescence was measured at 25 h. Reactions were prepared in triplicate, and the error bars represent one standard deviation.

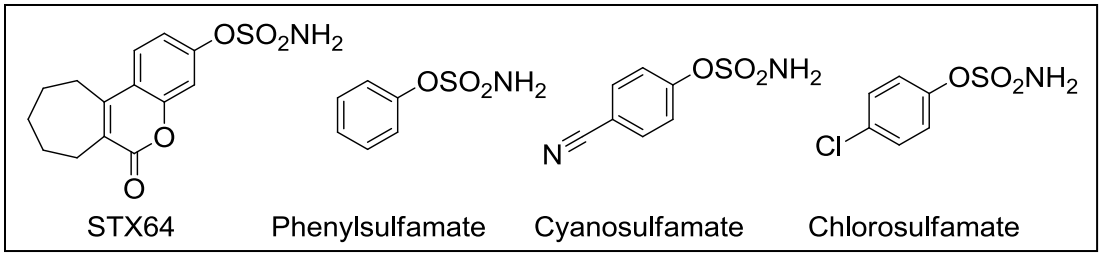


Figure S14: Structures of arylsulfamates.

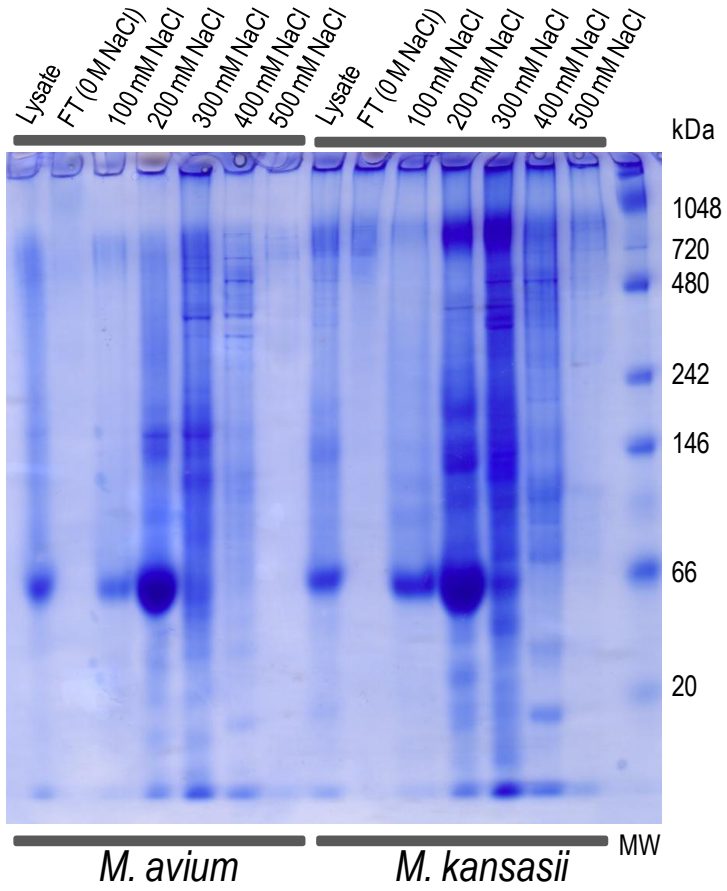
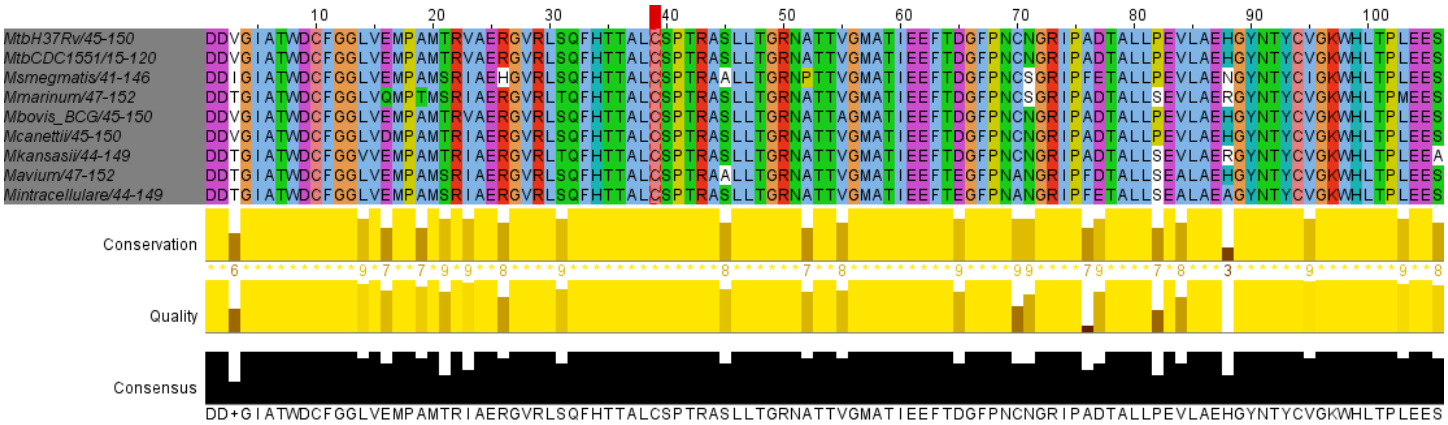
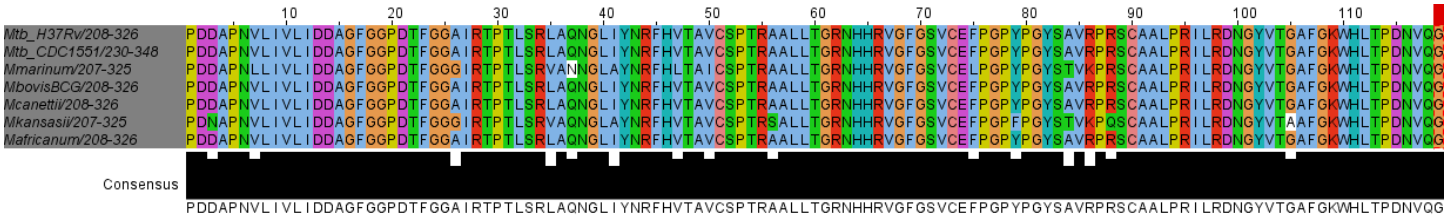


Figure S15. Coomassie stained protein gel of fractionated lysates from *M. avium* and *M. kansasii*. A strong anion exchange column was used to fractionate crude lysates from *M. avium* and *M. kansasii*. Proteins were eluted using a step gradient and then fractions were resolved by protein gel electrophoresis. Samples shown include lysates, flow-through (FT), and proteins eluted in with varying amounts of NaCl (100 mM to 500 mM) in 50 mM Tris-Cl buffer (pH 7.0), 1 mM MgCl2, 0.5 mM CaCl2. NativeMark molecular weight ladder (Life Technologies) is shown in the last lane.

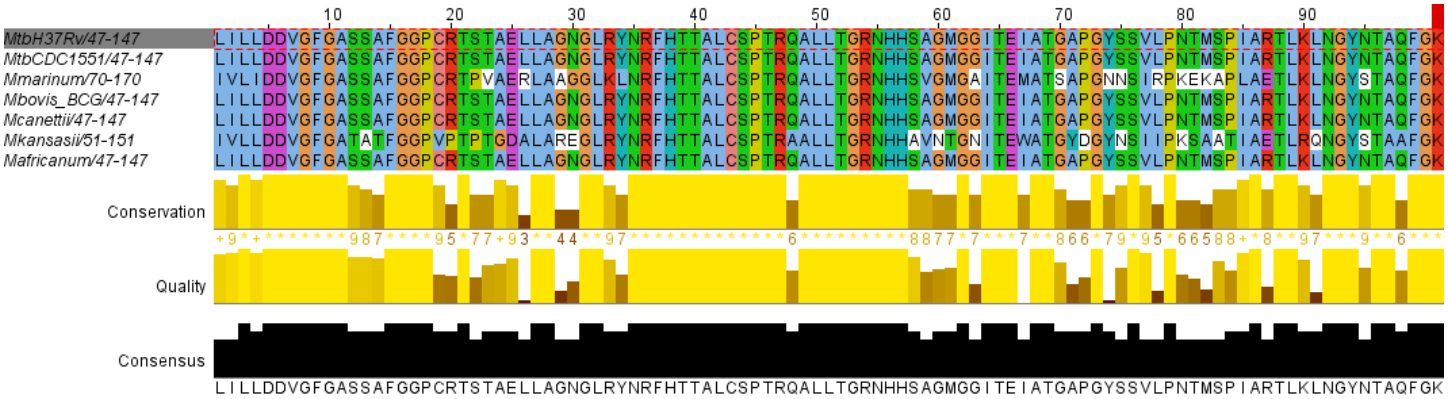
Type I Sulfatase Alignments



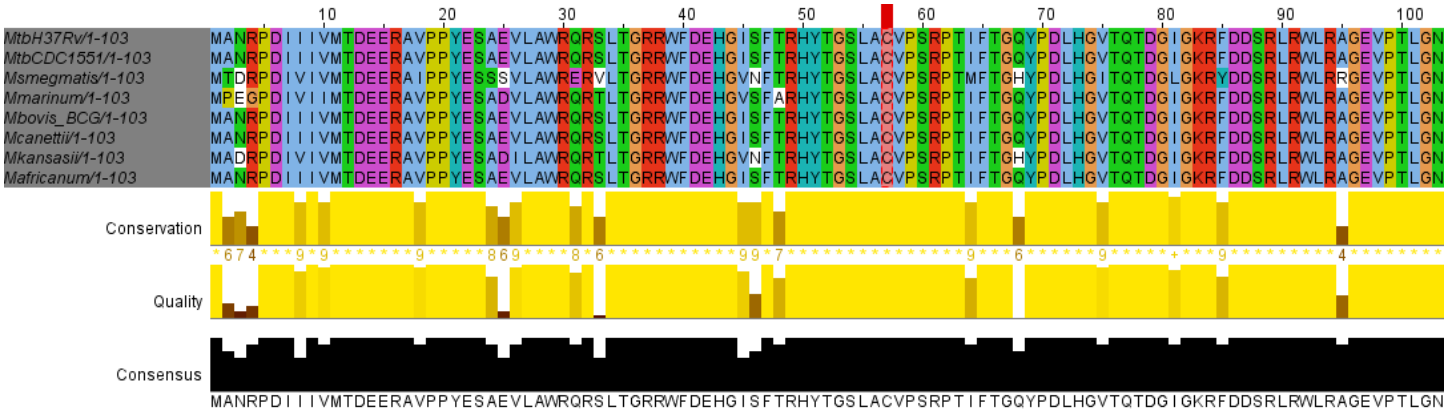
Sulfatase sequence alignment for a portion of Rv0711 (AtsA), including the region near the active site residues (CSPTR).



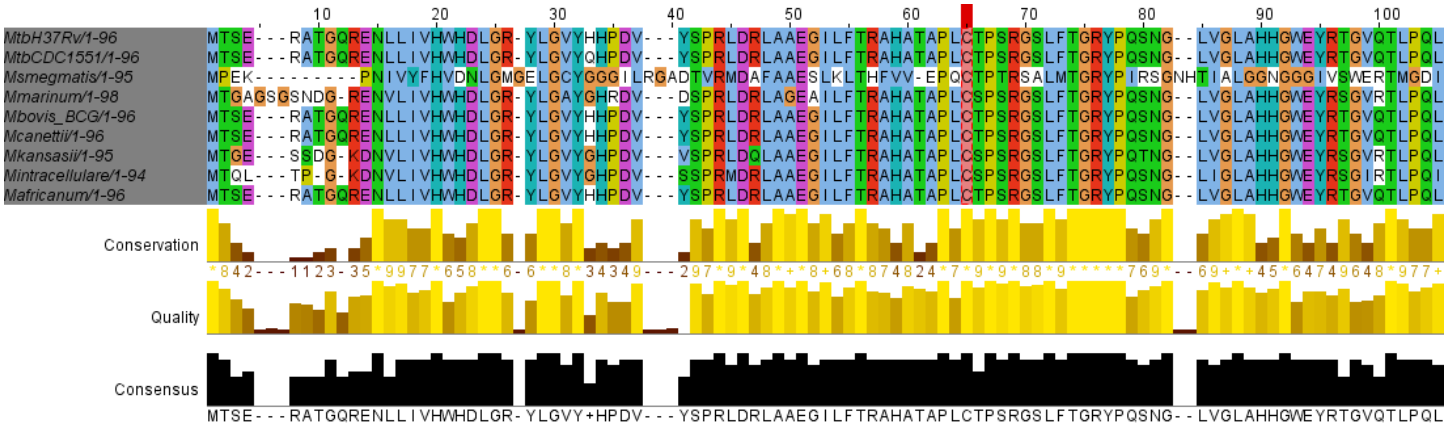
Sulfatase sequence alignment for a portion of Rv3299 (AtsB), including the region near the active site residues (CSPTR).



Sulfatase sequence alignment for a portion of Rv0663 (AtsD), including the region near the active site residues (CSPTR).

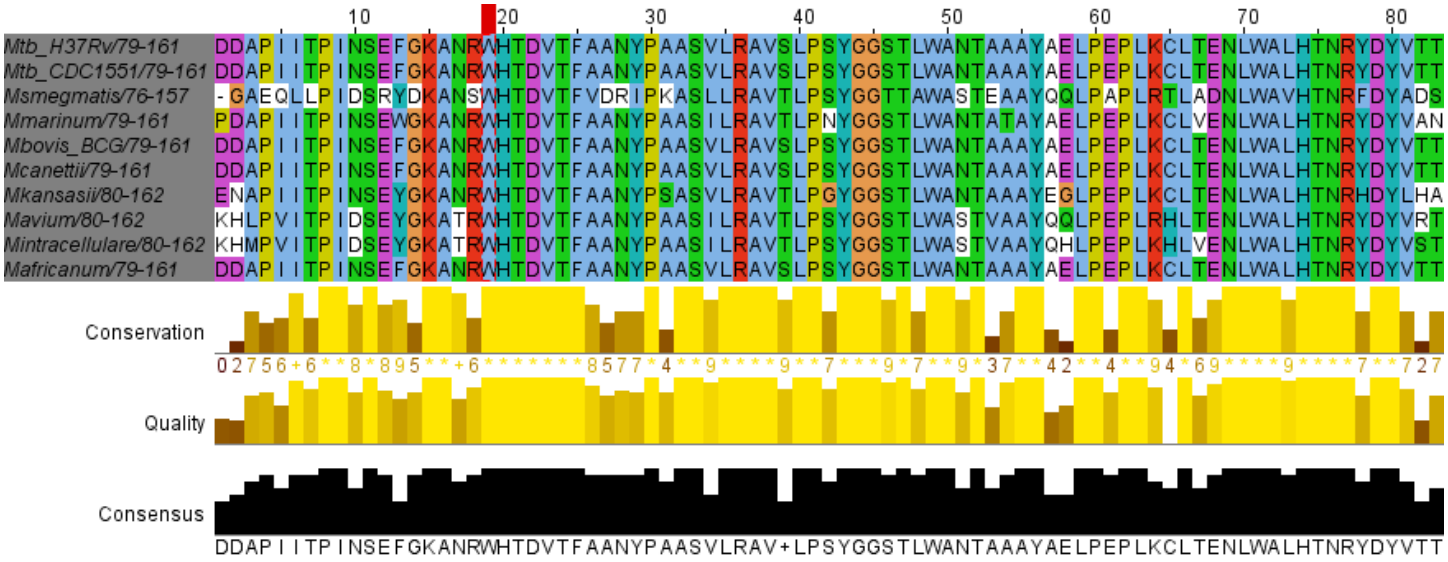


Sulfatase sequence alignment for a portion of Rv3077 (AtsF), including the region near the active site residues (CVPSR).



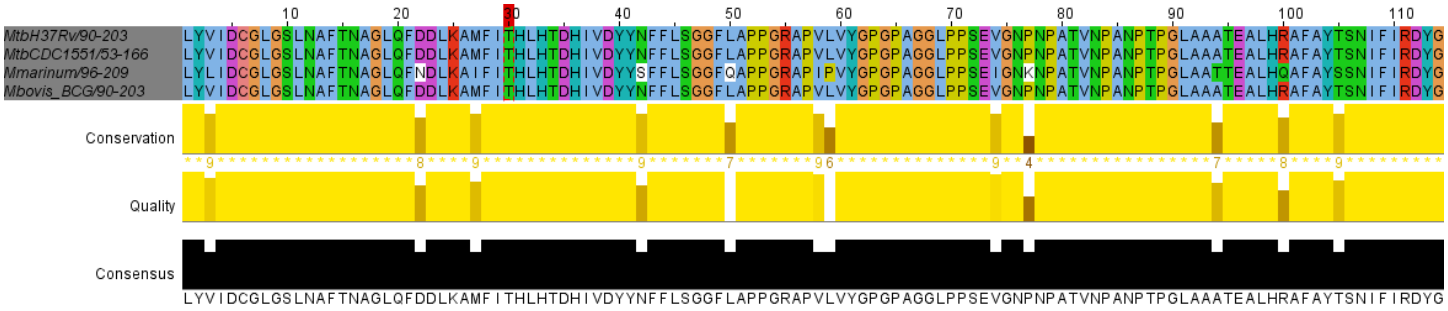
Sulfatase sequence alignment for a portion of Rv0296c (AtsG), including the region near the active site residues [C(S/T)PSR].

Type II Sulfatase Alignments

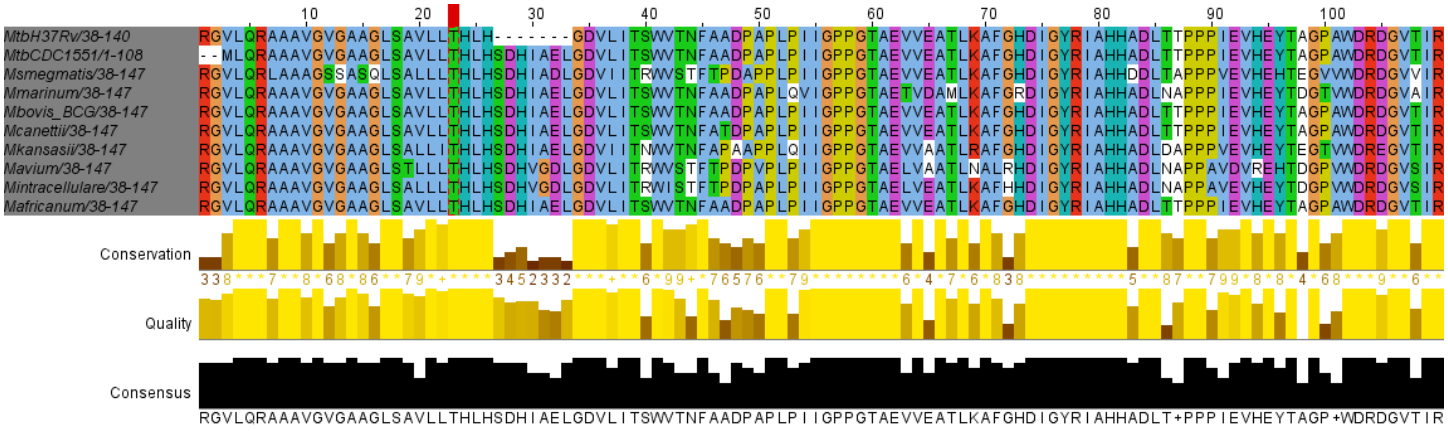


Sulfatase sequence alignment for a portion of Rv3407 (FeSulf-A).

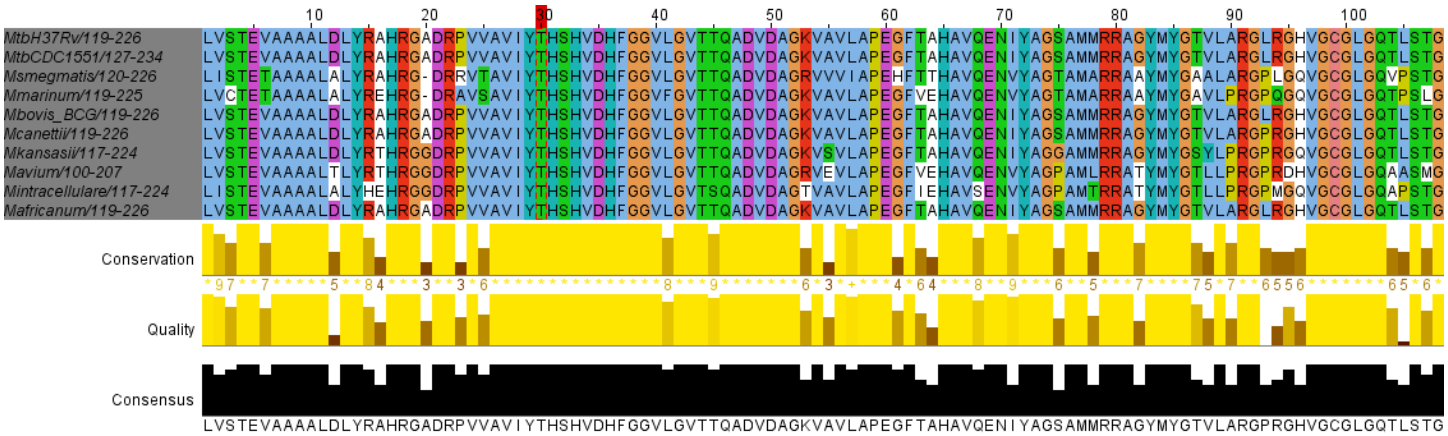
Type III Sulfatase Alignments



Sulfatase sequence alignment for a portion of Rv3796 (ZnSulf-A).



Sulfatase sequence alignment for a portion of Rv2407 (ZnSulf-B).



Sulfatase sequence alignment for a portion of Rv3762C (ZnSulf-C).

References:

1. Fery-Forgues S & Lavabre D (1999) Are Fluorescence Quantum Yields So Tricky to Measure? A Demonstration Using Familiar Stationery Products. *Journal of Chemical Education* 76(9):1260.
2. Allison RD PD (1979) Practical considerations in the design of initial velocity enzyme rate assays. *Methods Enzymology* 63:3-22.
3. Rush JS, Beatty KE, & Bertozzi CR (2010) Bioluminescent Probes of Sulfatase Activity. *ChemBioChem* 11(15):2096-2099.
4. Carlson BL*, et al.* (2008) Function and Structure of a Prokaryotic Formylglycine-generating Enzyme. *J. Biol. Chem.* 283(29):20117-20125.
5. Rasband WS (1997–2007) ImageJ (US National Institutes of Health, Bethesda). Available at http://imagej.nih.gov/ij/.
6. Waterhouse AM, Procter JB, Martin DMA, Clamp M, & Barton GJ (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25(9):1189-1191.