Supporting Information

The Molecular Basis of Sulfosugar Selectivity in Sulfoglycolysis

Mahima Sharma,[1] Palika Abayakoon,[2] Ruwan Epa,[2] Yi Jin,[1] James P. Lingford,[3,4] Tomohiro Shimada,[5] Masahiro Nakano,[7] Janice W.-Y. Mui,[2] Akira Ishihama,[6] Ethan D. Goddard-Borger,*[,3,4] Gideon J. Davies,*[,1] Spencer J. Williams*[,2]

Contents

Supplementary Figures

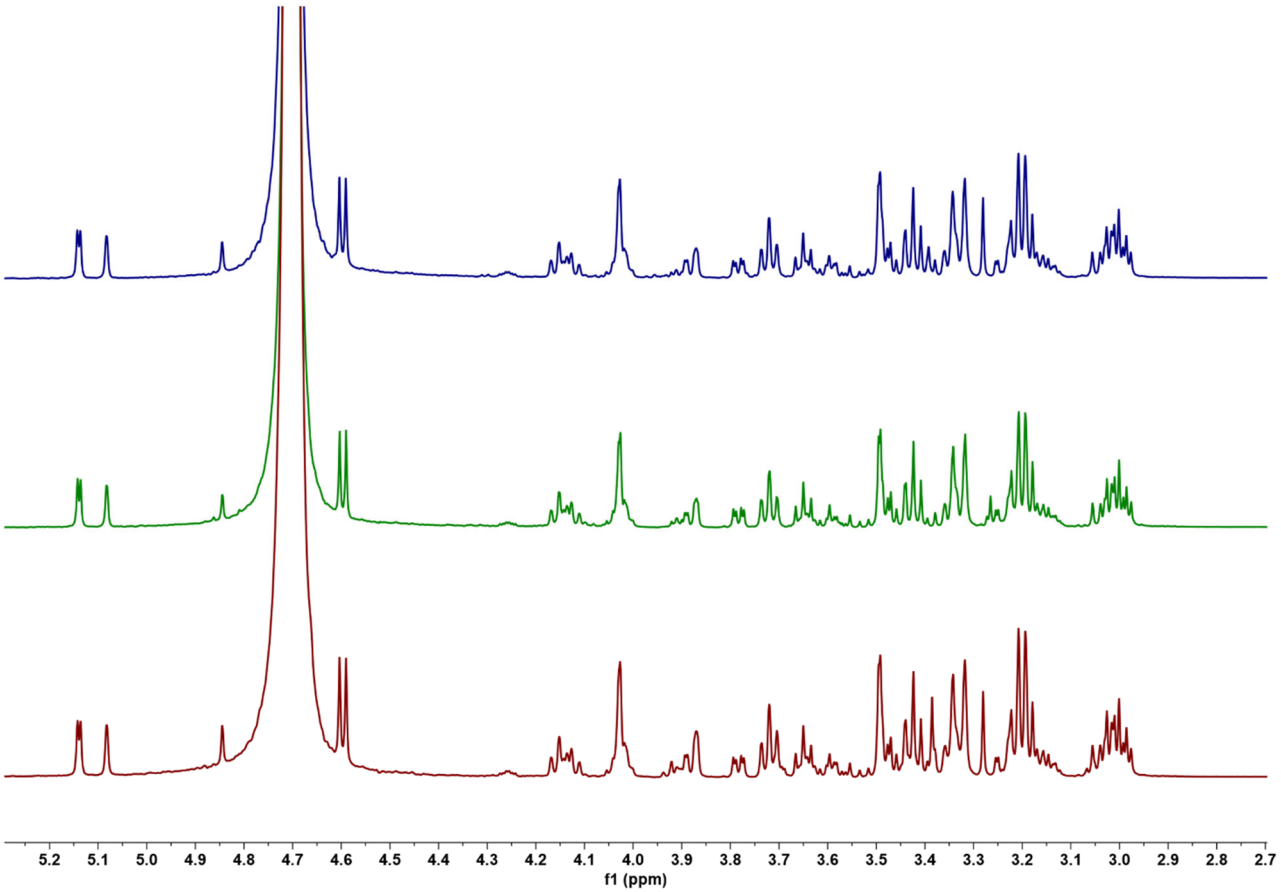


Figure S1. ^{1}H NMR spectra showing *Ec*YihS-catalyzed reactions of SQ (top), SF (middle) or SR (bottom). Reactions consisted of the sugar and *Ec*YihS in 50 mM sodium phosphate, 150 mM NaCl (pH 7.00) buffer at 37 °C for 24 h; then heat inactivated and solvent exchanged to D_2O.

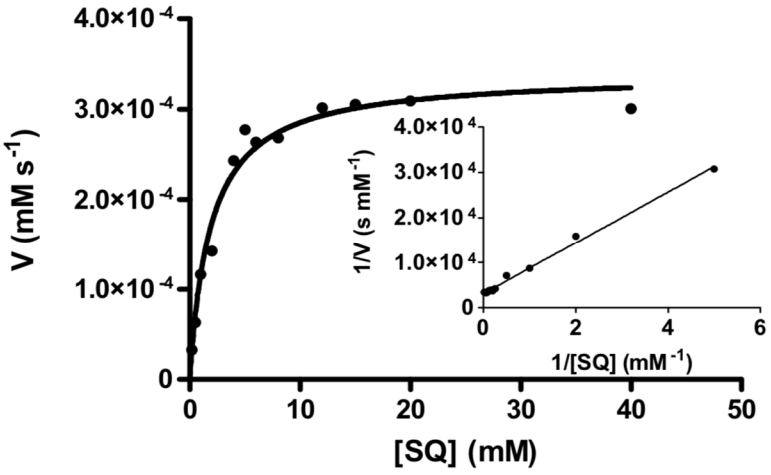


Figure S2. Kinetic characterization of SQ isomerase *Ec*YihS. Michaelis-Menten plot for *Ec*YihS catalyzed isomerization of SQ, inset, Lineweaver-Burk plot.

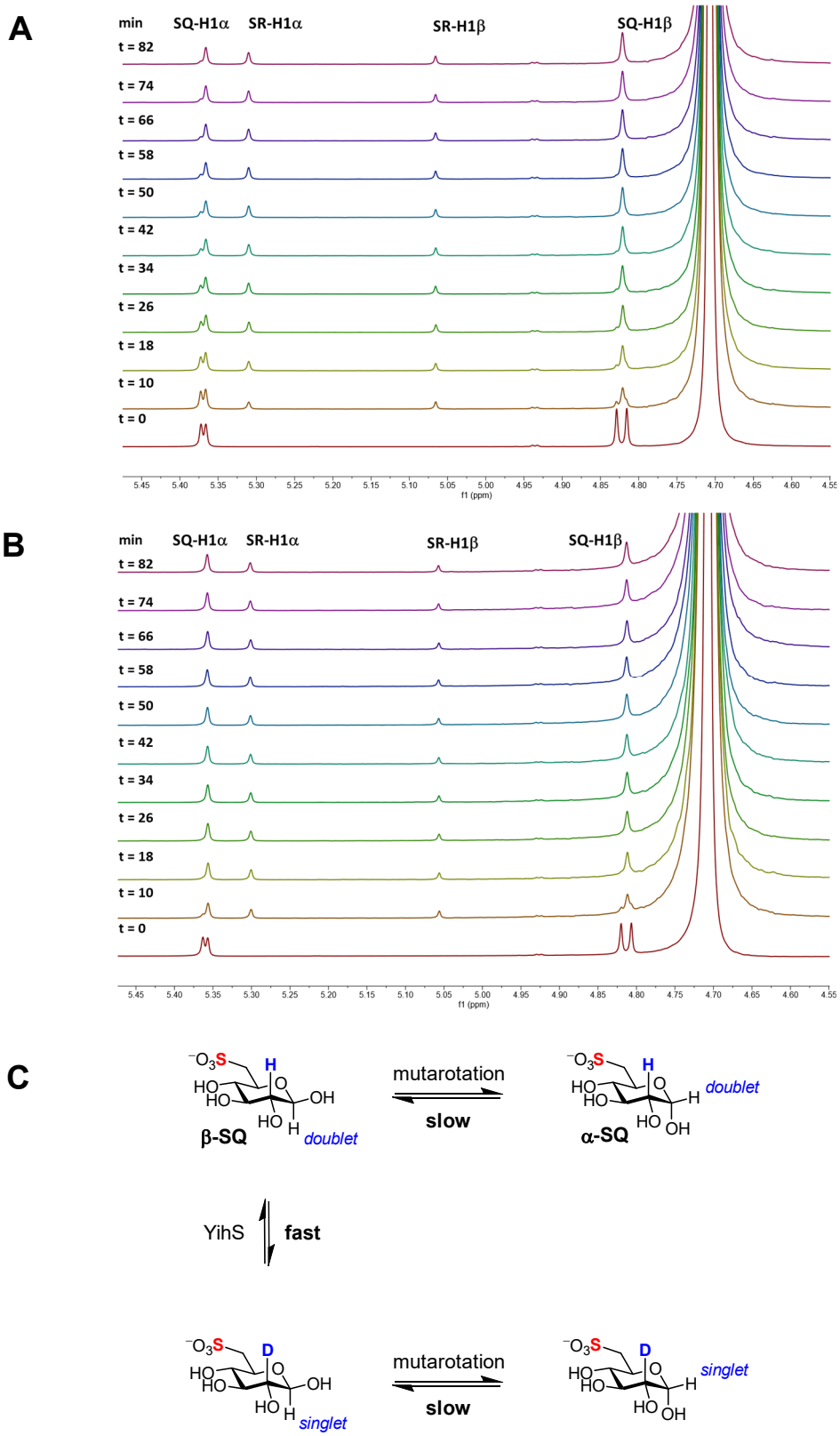


Figure S3. H/D exchange of *Ec*YihS-catalysed conversion of SQ in D_2O. *A,* Time-course ^1H NMR spectra (in D_2O) of SQ incubated with *Ec*YihS; SQ reference spectrum (t = 0). *B,* Time-course ^1H NMR spectra of SQ (in D_2O) incubated with *Ec*YihS and *Hs*SQM mutarotase; SQ reference spectrum (t = 0). *C*, Scheme showing H/D exchange at C2 catalyzed by YihS isomerase.

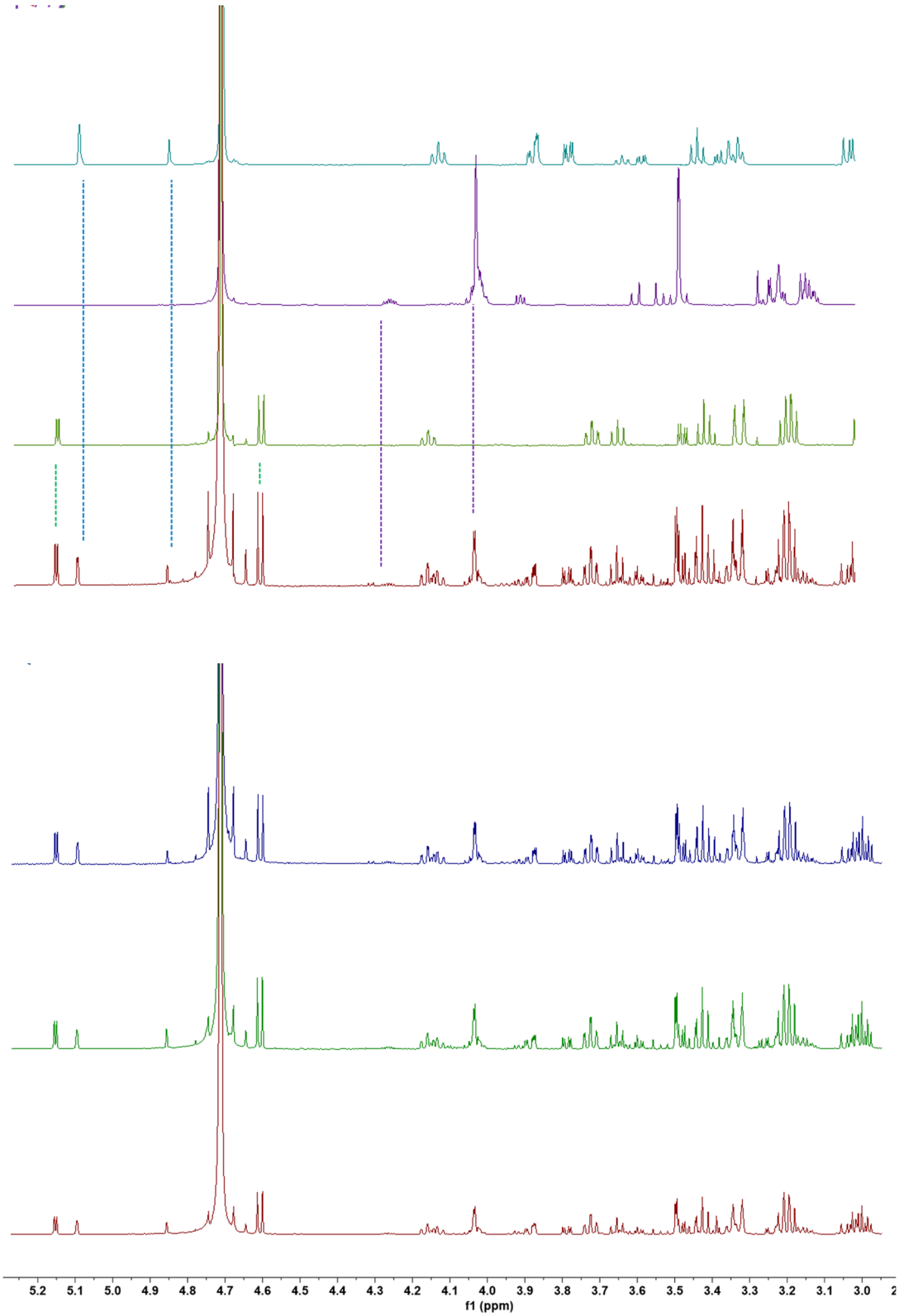
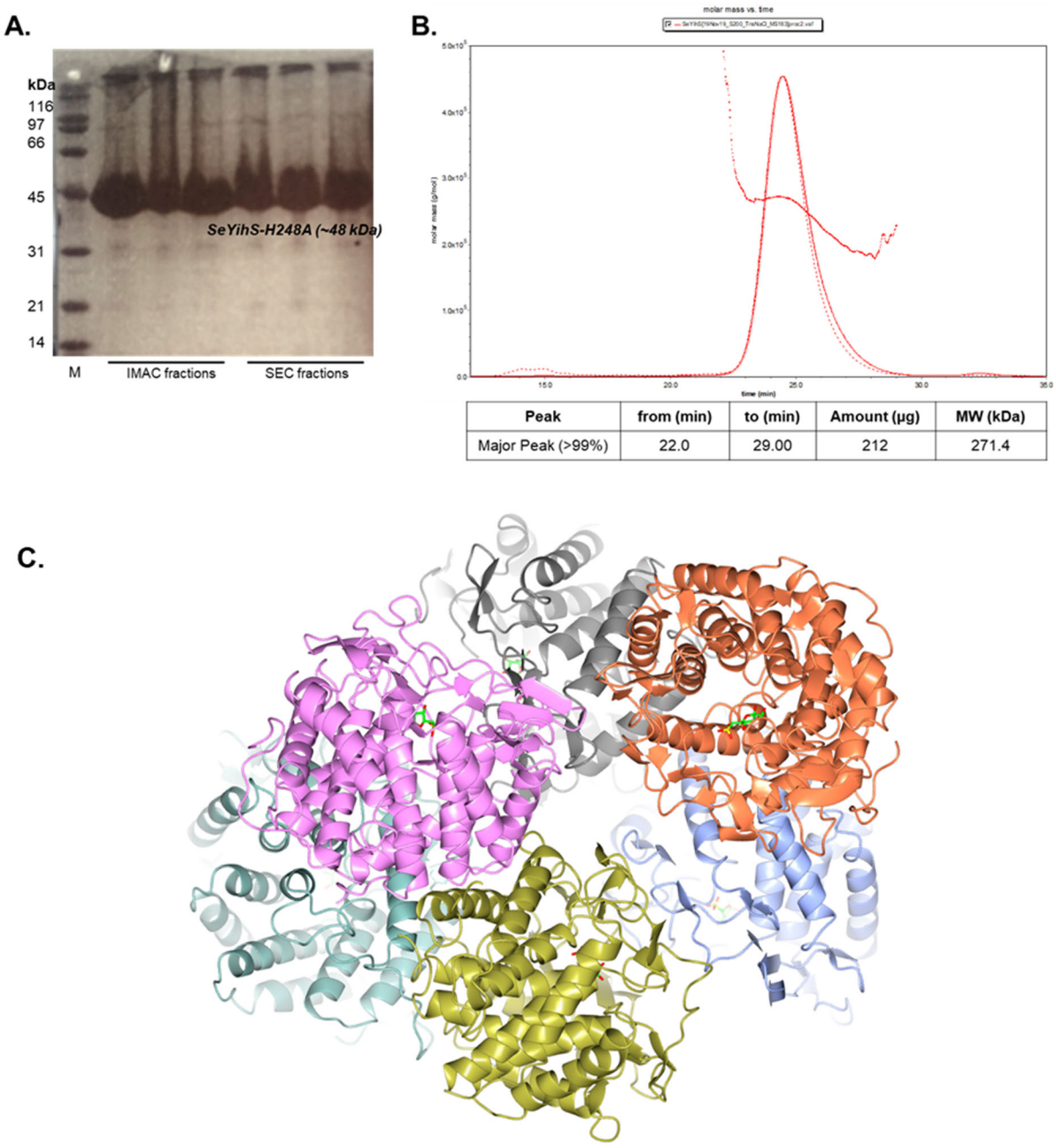


Figure S4. (Top panel) ^{1}H NMR spectrum of SQ incubated with *Se*YihS for 24 h in 50 mM sodium phosphate, 150 mM NaCl (pH 7.00) buffer at 37 °C, then heat inactivated and solvent exchanged to D_2O, along with reference spectra of SQ, SF and SR in D_2O. (Bottom panel) ^{1}H NMR spectra showing *Se*YihS-catalyzed reactions of SQ (upper), SF (middle) or SR (lower). Reactions consisted of the sugar and *Ec*YihS in 50 mM sodium phosphate, 150 mM NaCl (pH 7.00) buffer at 37 °C for 24 h; then heat inactivated and solvent exchanged to D_2O.



Peak	from (min)	to (min)	Amount (µg)	MW (kDa)
Major Peak (>99%)	22.0	29.00	212	271.4

Figure S5. Purification, properties and molecular assembly of *Se*YihS-H248A. *A,* SDS-PAGE analysis of purified *Se*YihS after IMAC and size exclusion chromatography (SEC) showing expected Mw of 48,000 Da. *B,* SEC-MALS plot reveals oligomeric state of *Se*YihS-WT in solution. UV-trace and an average molecular weight trace (red), calculated from the refractive index and light scattering signal gave mass estimate of 271 kDa, which corresponds to a hexamer (with some dissociation under the experimental conditions) and comprises >99% of the eluted material confirming homogeneity of the sample. *C*, Ribbon diagram of *Se*YihS hexamer.

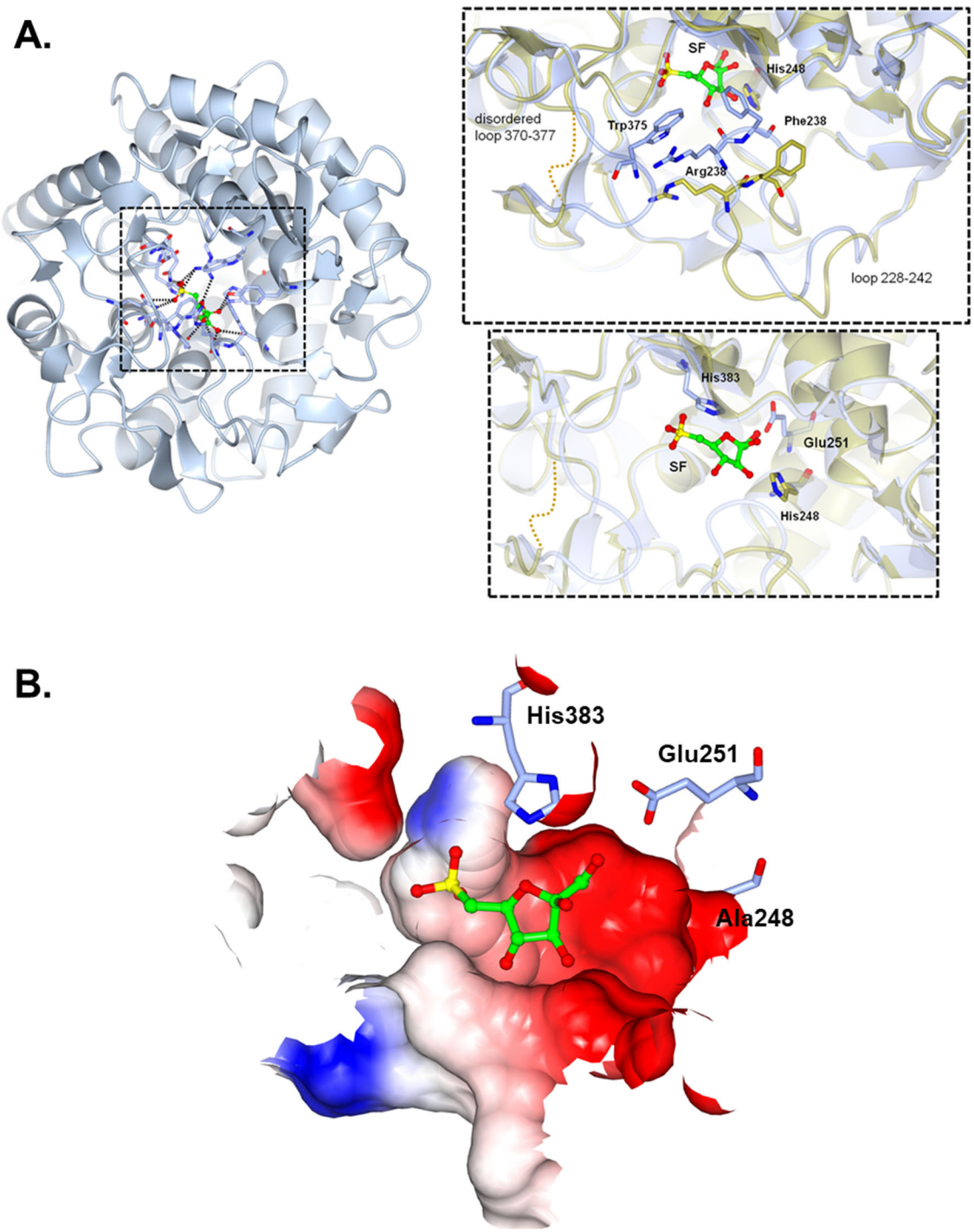


Figure S6. Overlay of wild-type *Se*YihS and active site mutant *Se*YihS-H248A•SF showing loop rearrangements upon binding sulfofructose. *A,* Active sites of wild-type (2AFA.pdb) and ligand bound structures (in gold and blue respectively) show re-ordering of loops, cation-pi stacking interaction of Arg238 and Trp375 and different side-chain conformations of Phe239 observed in crystal structures. *B*, Electrostatic potential depiction of YihS•SF showing product bound in a compact, polar active site. SF and catalytic residues (248, 383, 251) are shown as cylinders.

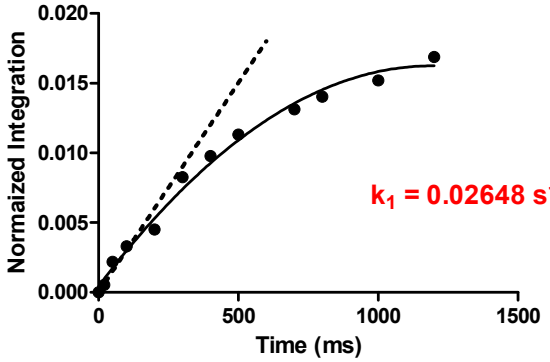
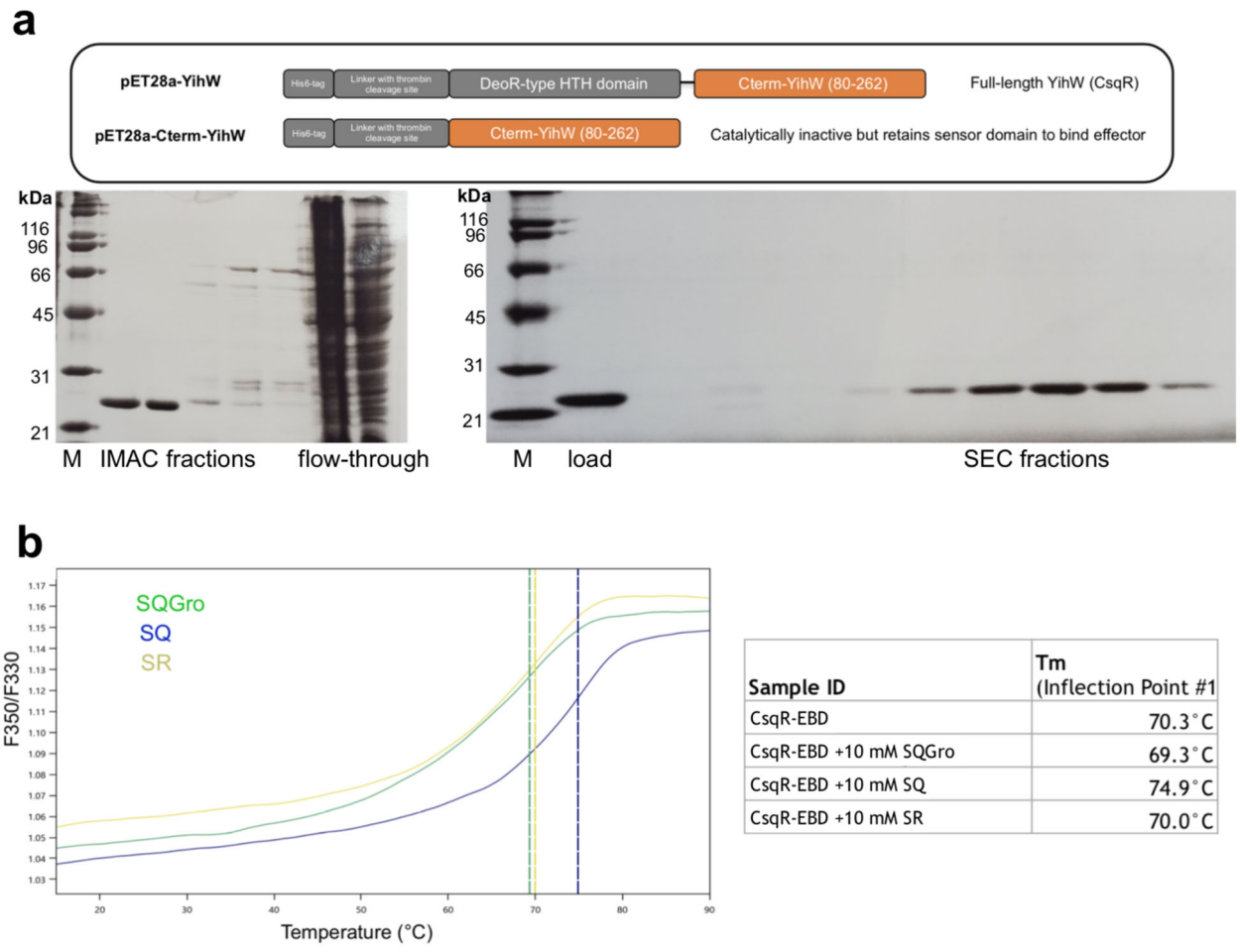
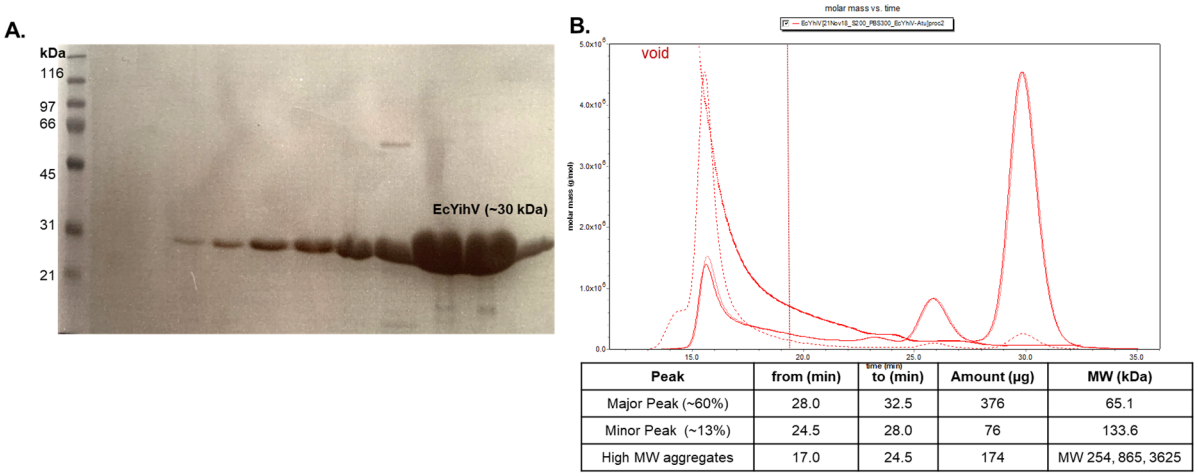


Figure S7. Kinetic analysis of SF mutarotation by inversion recovery 1D ^{1}H exchange spectroscopy. Inversion recovery curves for 5 mM β-SF at δ 4.25 ppm using a Gaussian-shaped pulse of 20 ms. Dashed line indicates tangent to the fitted curve at $t = 0$.



Sample ID	Tm (Inflection Point #1
CsqR-EBD	70.3°C
CsqR-EBD +10 mM SQGro	69.3°C
CsqR-EBD +10 mM SQ	74.9°C
CsqR-EBD +10 mM SR	70.0°C

Figure S8. Purification of effector binding domain of CsqR transcription regulator (CsqR-EBD) from *E. coli*. *A,* Schematic showing contructs of CsqR studied here. SDS-PAGE analysis of purified CsqR-EBD after two-step purification (Expected Mw: 22,000 Da). *B*, Temperature unfolding assay with different effector molecules (SQ, SQGro, SR) showing maximum T_m shift of 4.6 °C for SQ.



Peak	from (min)	to (min)	Amount (µg)	MW (kDa)
Major Peak (~60%)	28.0	32.5	376	65.1
Minor Peak (~13%)	24.5	28.0	76	133.6
High MW aggregates	17.0	24.5	174	MW 254, 865, 3625

Figure S9. Purification and properties of *Ec*YihV. *A,* SDS-PAGE analysis of purified *Ec*YihV after size exclusion chromatography (Expected Mw: 32,000 Da). *B,* SEC-MALS plot reveals oligomeric state of *Ec*YihV in solution. UV-trace and an average molecular weight trace (red), calculated from the refractive index and light scattering signal gave mass estimate of 65 kDa which corresponds to a dimer (along with a sub-population of tetrameric species with MW 134 kDa and higher MW aggregates).

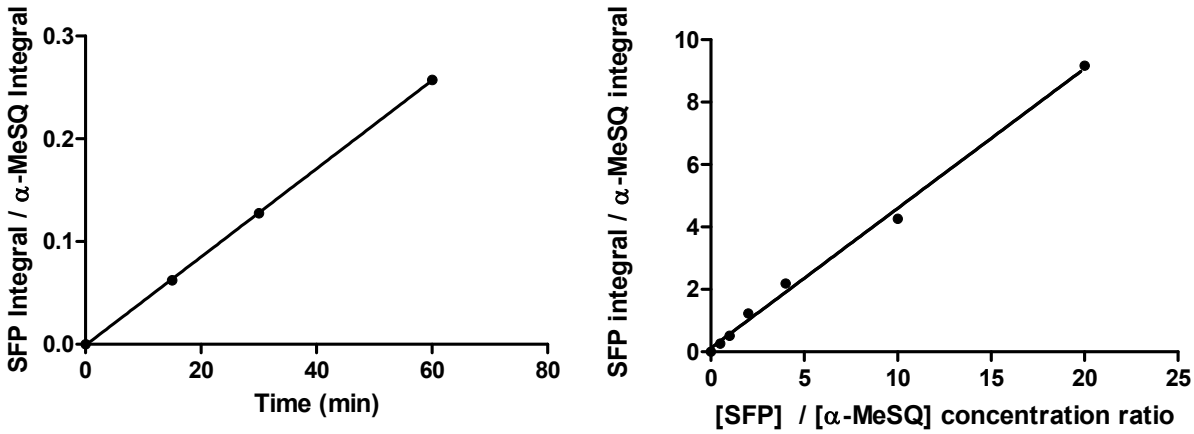


Figure S10. LC-MS/MS assay for SF kinase. (Left) Linearity of the *Ec*YihV kinase reaction with time. (Right) Calibration curve for detection of SFP using methyl α-sulfoquinovoside (α-MeSQ) as an internal standard. Reactions were carried out in 25 mM BTP (pH 7.5), 25 mM KCl, 5 mM $MgCl_2$, 0.1 mg/mL BSA, 1.0 mM ATP, 0.1 mM SF, 36.6 nM YihV, 30 °C, 60 min.

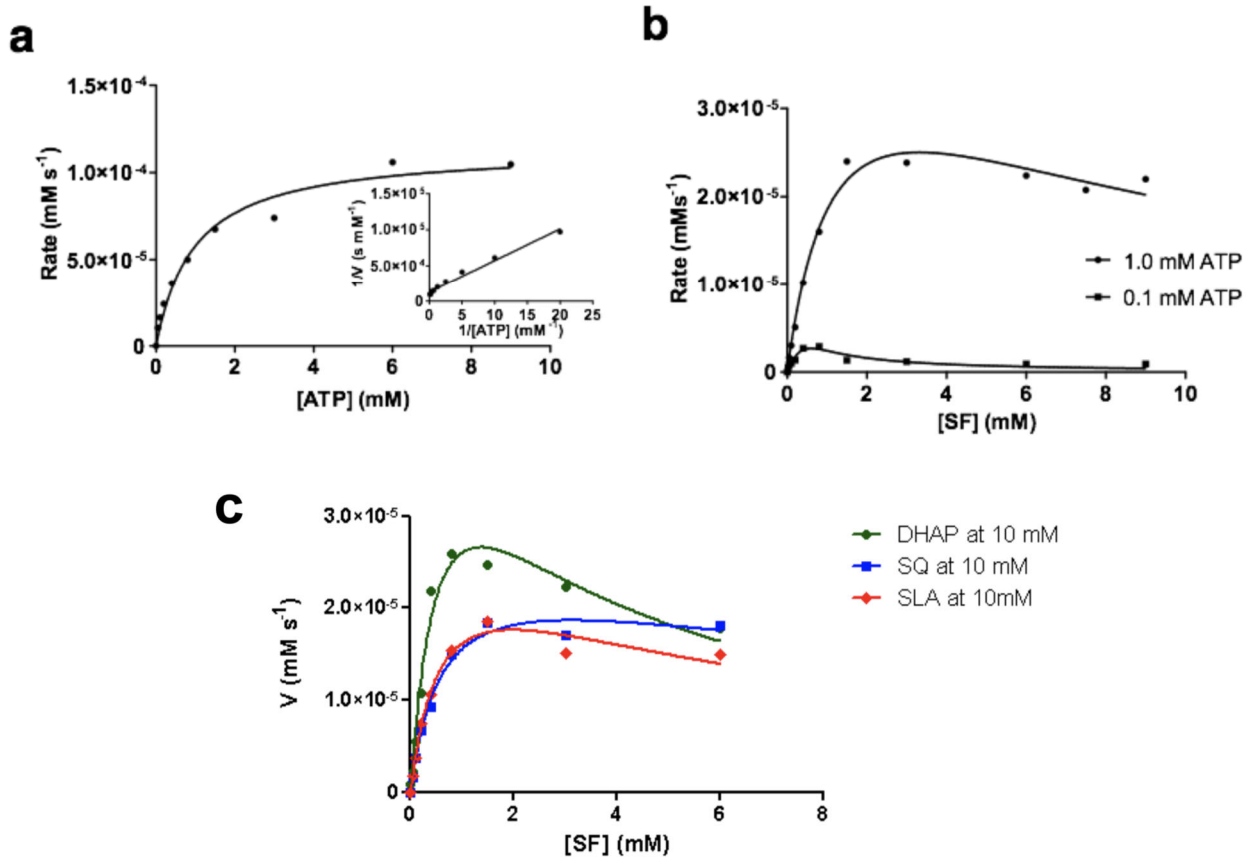


Figure S11. Characterization of *Ec*YihV. (a) Michaelis-Menten and Lineweaver-Burk plots (inset) for YihV catalyzed phosphorylation of SF to SFP at [SF] = 1.0 mM with the variation of ATP. **(b)** Substrate inhibition plots for *Ec*YihV catalyzed phosphorylation of SF to SFP at [ATP] = 1.0 mM and 0.1 mM with variation of SF. **(c)** Kinetic plots showing effect of metabolites (DHAP, SQ, SLA, each at 10 mM) on *Ec*YihV catalyzed phosphorylation of SF to SFP at [ATP] = 1.0 mM.

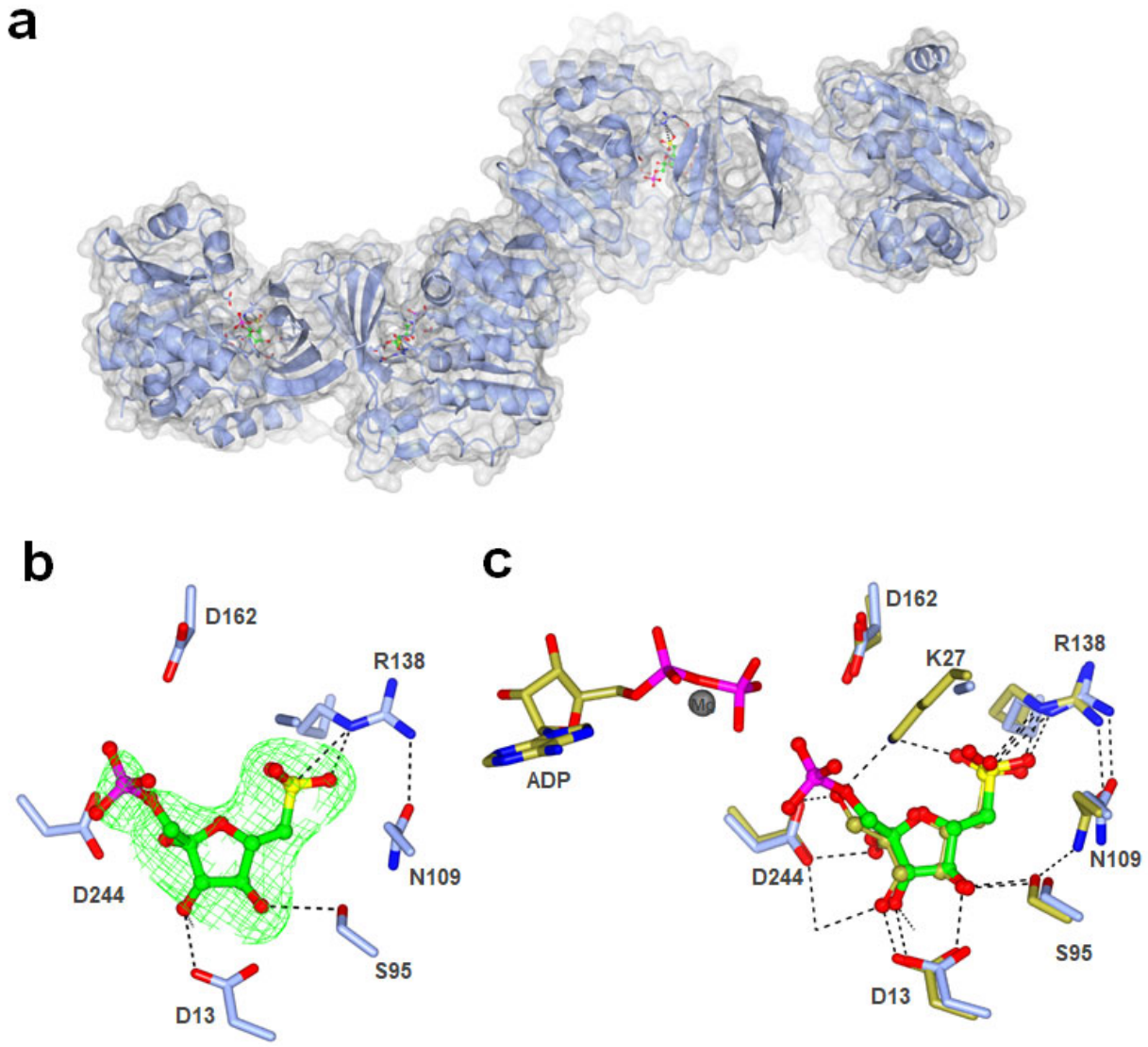


Figure S12. (a) Crystallographic dimer-of-dimer assembly of YihV-SFP complex. **(b)** Omit map of ligand SFP. Electron density represents the Fo–Fc (omit map, in green) obtained prior to modelling of the ligand, and contoured at levels of 3σ. **(c)** Overlay of SFP complex with YihV-ADP-Mg-SF complex (in gold) to show orientation of the sulfonate ligands in YihV structures with respect to active site residues.

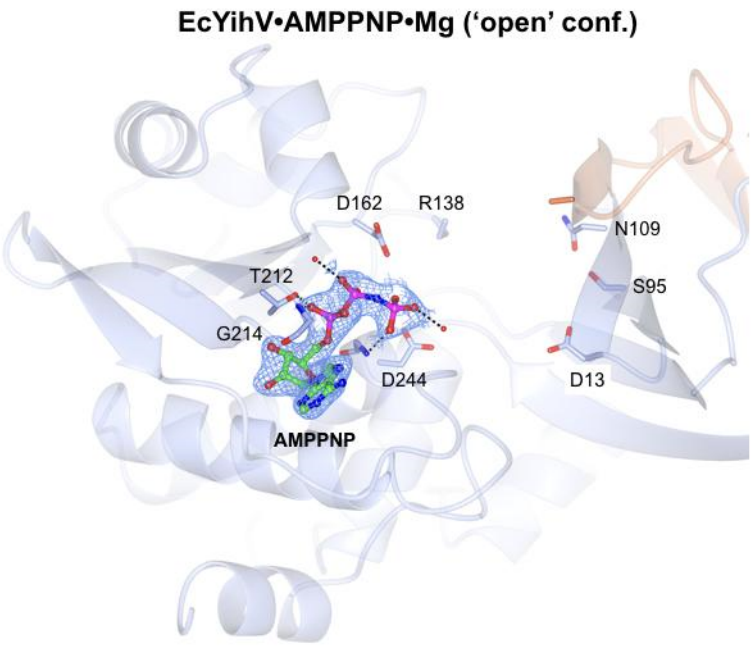
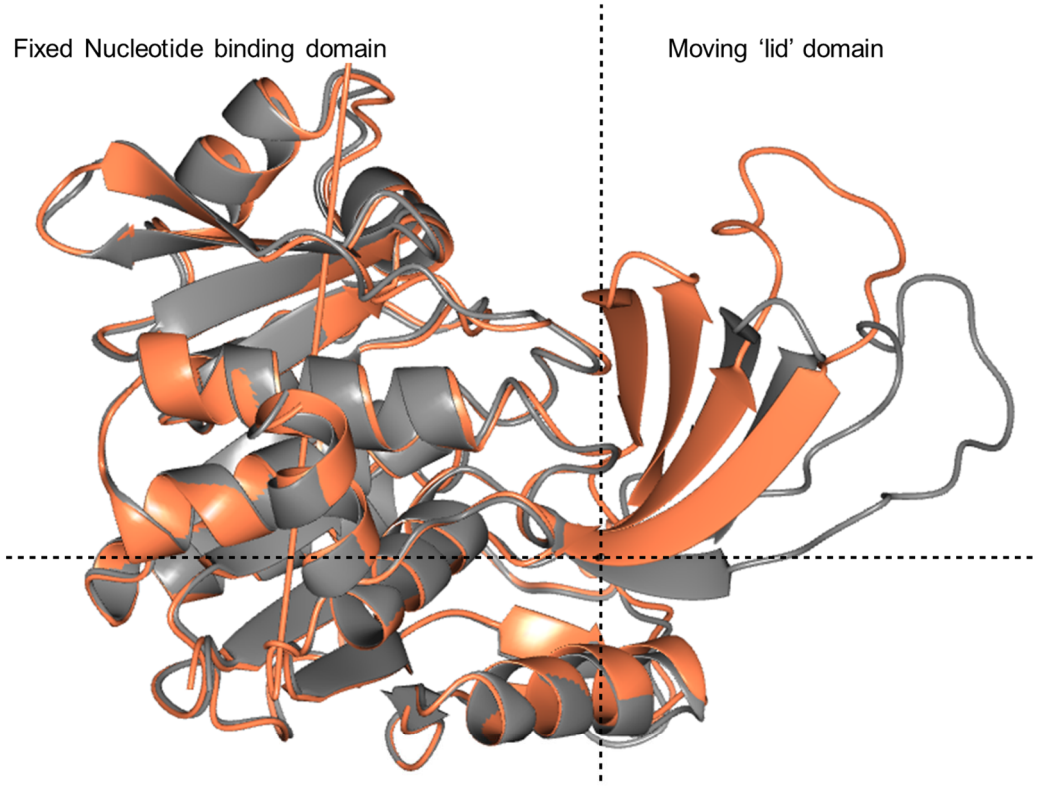


Figure S13. Crystal structures of *Ec*YihV SF kinase. *A*, Close-up view of hydrogen bonding interactions of *Ec*YihV dimer in complex with AMPPNP•Mg in open conformations with AMPPNP bound at the nucleotide binding site. Backbone and carbon atoms of subunits A and B are shown in coral and blue, respectively, and ADP, AMPPNP is shown in cylinder format. Electron density corresponds to 2Fo − Fc and in blue at 1σ contour level.



	Backbone RMSD (Å)
Domain 1	Nucleotide binding domain
Domain 2	β-clasp or 'lid' domain
Whole protein best fit RMSD	2.48
Bending Residues	11 – 12;35 - 38 86 – 94; 110 - 113
Rotation (deg)	**30.7**

Figure S14. Domain movement analysis of *Ec*YihV. DynDom analysis of the dynamic domains and hinge bending motion of the *Ec*YihV 'open' (depicted in grey) vs. ADP•Mg•SF bound conformation (in coral). The lines cross at the centre of rotation and the hinge axis is perpendicular to this crossing point.

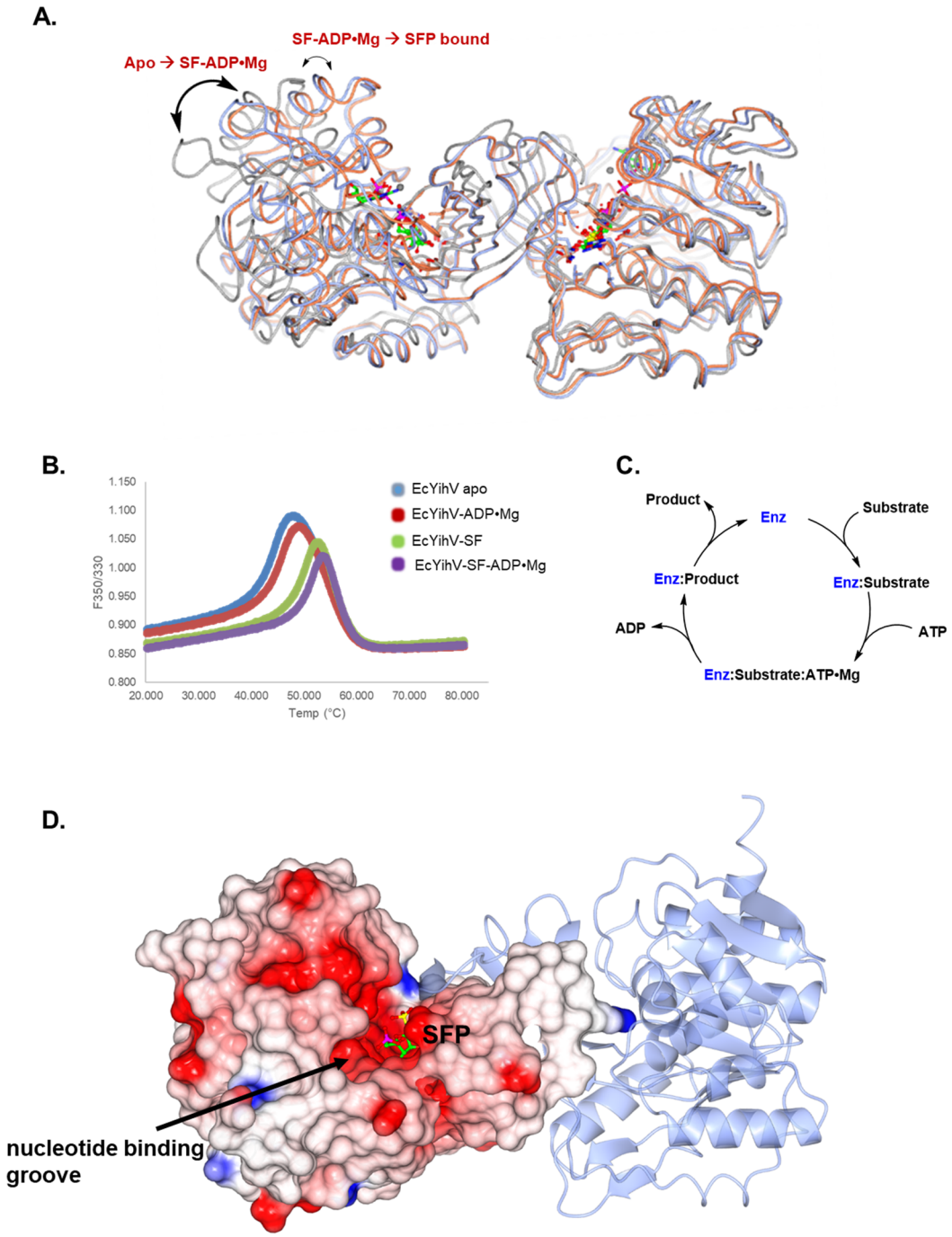


Figure S15. Comparison of *Ec*YihV•AMPPNP, *Ec*YihV•SFP and *Ec*YihV•ADP•Mg•SF complexes. *A,* Overlay of YihV dimer (in coral and blue) of three complexes to show open, closed and slightly staggered conformations. *B*, Temperature unfolding assay with *Ec*YihV reveals T_m shifts upon binding substrates or products. *C*, Ordered sequential binding order observed in ribokinases. *D*, Surface accessible nucleotide binding groove depicted in the YihV•SFP binary complex. Electrostatic potential depiction of subunit A in a dimer pair in closed conformation.

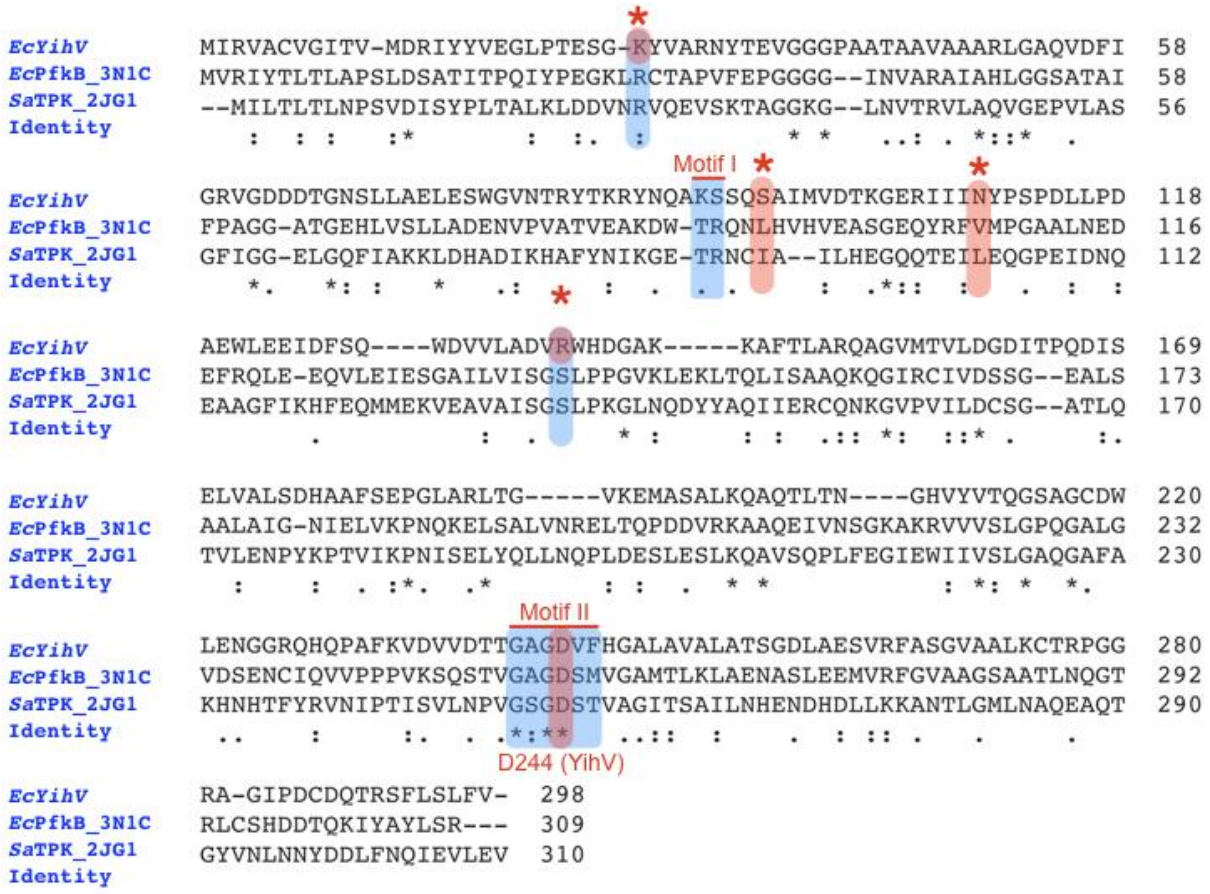


Figure S16. *Ec*YihV shares a catalytic motif with phosphosugar kinases (PfkB and TPK) and possesses a unique sulfonate substrate binding sequence. Sequence alignment of *Ec*YihV with model phosphosugar kinases shows conserved catalytic motif (*GXGDXX* sequence, Motif II) but lacks TR motif (Motif I) containing substrate binding Arg residue. Sulfonate binding KRN triad in *Ec*YihV is highlighted in red and phosphate binding residues are highlighted in blue.

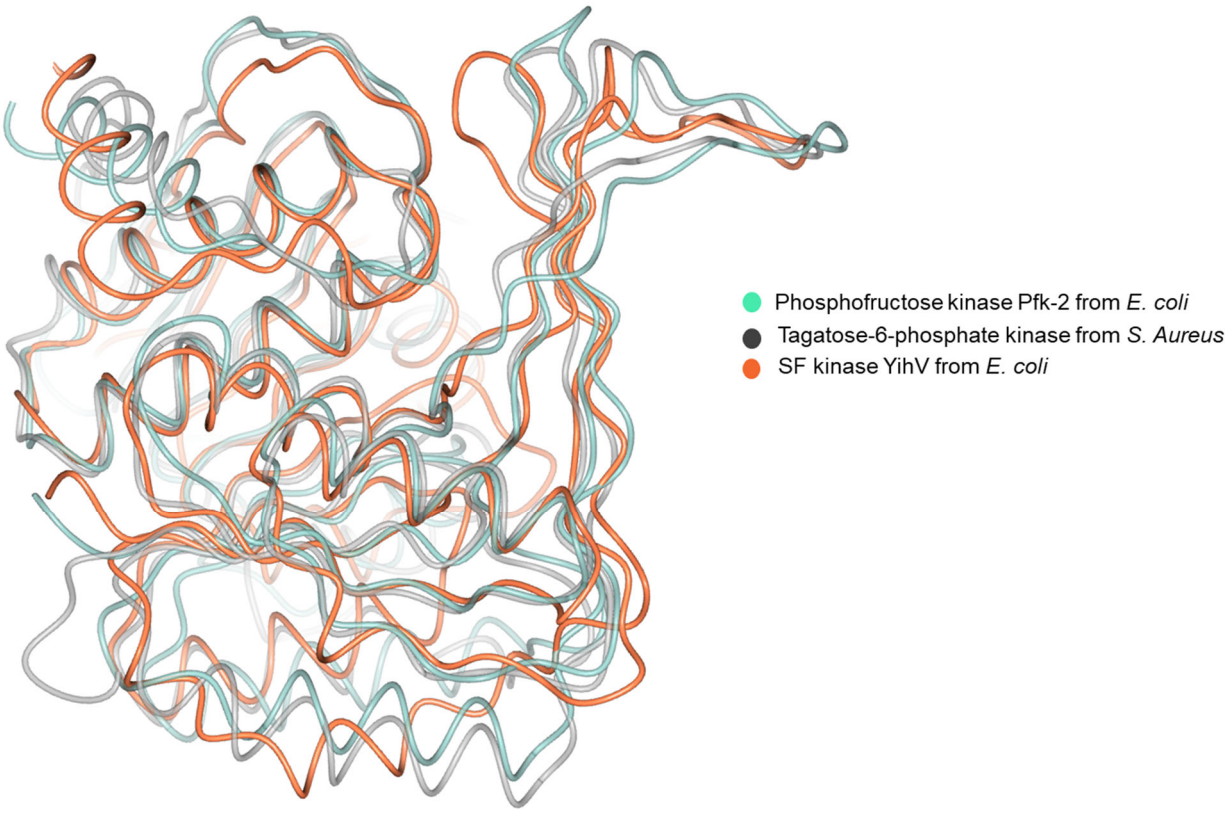


Figure S17. Closed conformations of PfkB, TPK and SF kinases show close structural relationships. *A,* Overlay of PfkB•F6P (in 3N1C.pdb) and TPK•ADP•T6P (2JG1.pdb) and *Ec*YihV•ADP•SF structures.

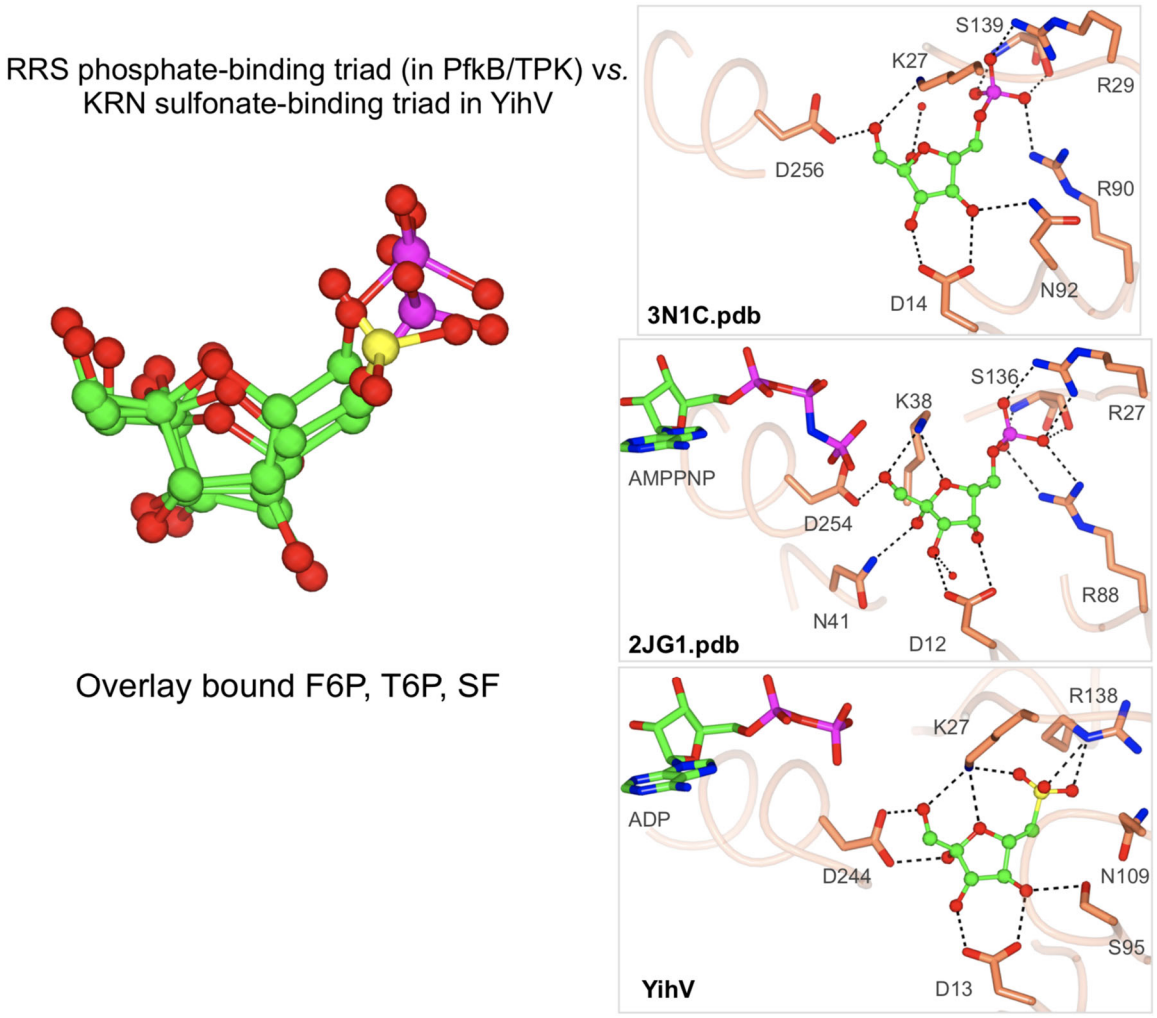
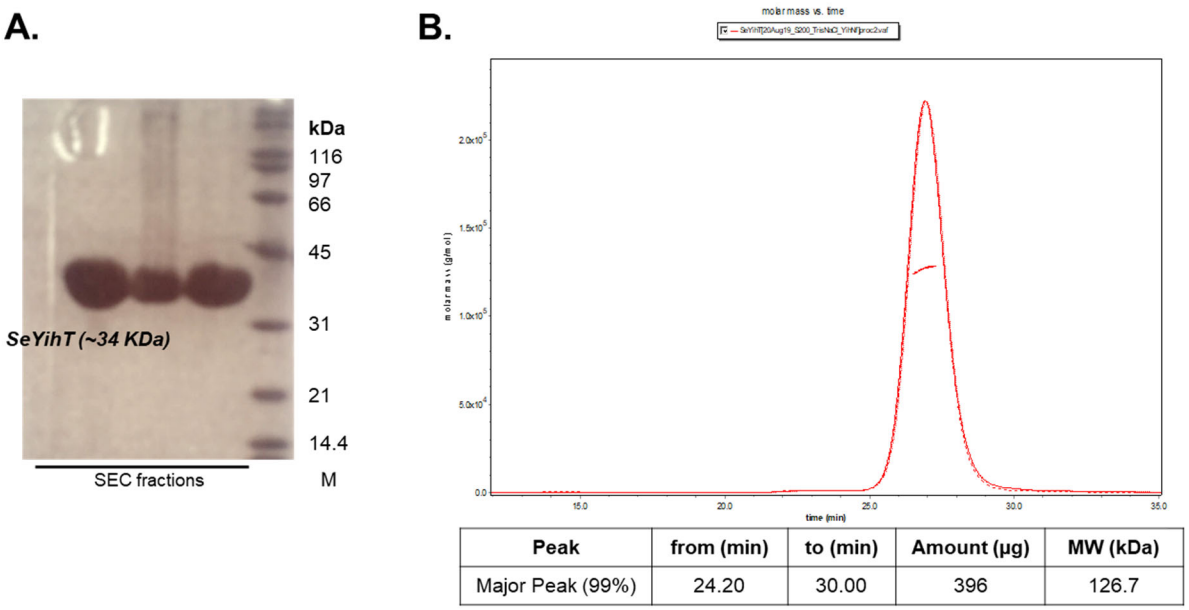


Figure S18. Structurally conserved motifs in PfkB, TPK and SF kinases provide distinct binding modes for phosphate *vs.* sulfonate substrates. (Left) Overlay of respective ligands (F6P, T6P and SF) highlighting the similarities in binding pose. (Right) Active site shown with partial **GXGDXX** and **TR** motifs in PfkB•F6P (in 3N1C.pdb) and TPK•ADP•T6P (2JG1.pdb). Comparison of phosphate-binding site in these structures and with sulfonate binding site in the *Ec*YihV•ADP•Mg•SF structure reveals KRN sulfonate recognition triad in YihV. ADP/AMPPNP molecules shown to illustrate the relative positioning of the phosphoryl donor.



Peak	from (min)	to (min)	Amount (µg)	MW (kDa)
Major Peak (99%)	24.20	30.00	396	126.7

Figure S19. Purification and properties of *Se*YihT. *A,* SDS-PAGE analysis of purified *Se*YihT after IMAC and size exclusion chromatography (expected Mw: 34000 Da). *B,* SEC-MALS plot reveals oligomeric state of *Se*YihT in solution. UV-trace and an average molecular weight trace (red), calculated from the refractive index and light scattering signal gave mass estimate of 127 kDa, which corresponds to a tetramer and comprises 99% of the eluted material confirming homogeneity of the sample.

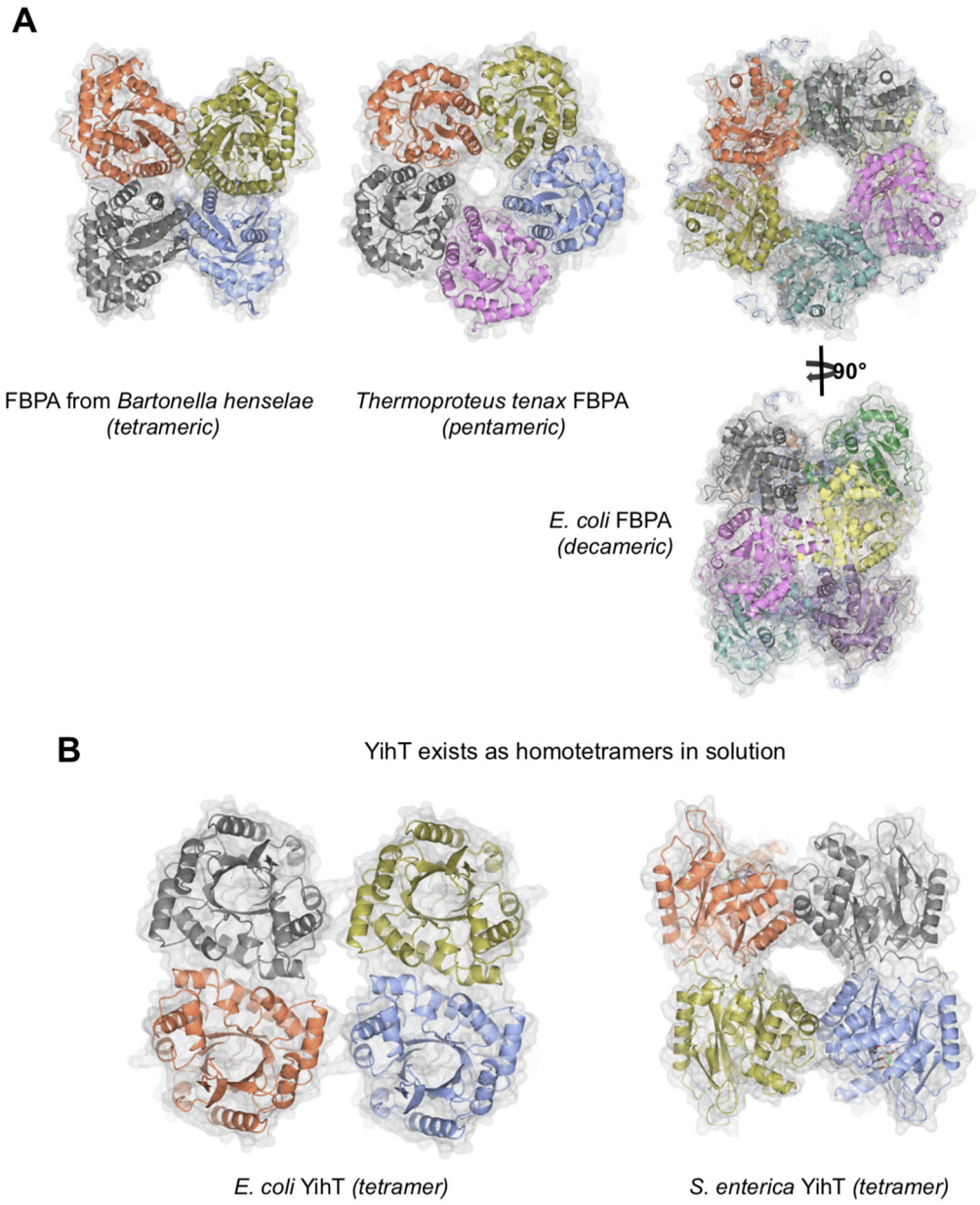


Figure S20. Oligomeric states of prokaryotic Class I fructose bisphosphate aldolases (FBPA). *A,* FBPA from *Bartonella henselae* forms homotetramers (3MMT.pdb),[1] FBPA from *Thermoproteus tenax* forms homopentamers (1W8S.pdb)[2] and Class I FBPA from *E. coli*[3] forms homodecamers in solution (model constructed using Synchrotron SAXS data of solutions of *E. coli* fbaB). *B*, Crystallographic tetramer of *E. coli* YihT and *S. enterica* YihT.

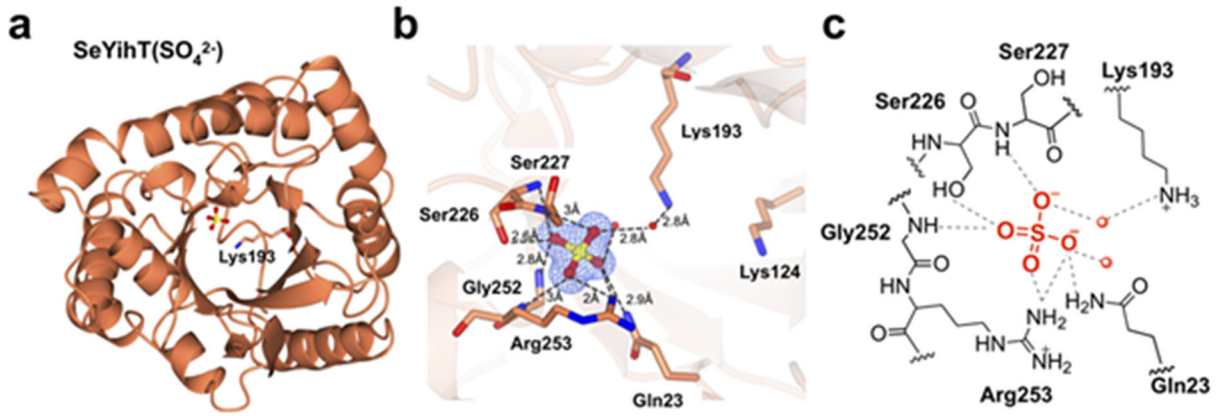


Figure S21. (a) Overview of the SeYihT•sulfate complex showing location of sulfate-binding site. **(b)** Close-up view of *Se*YihT sulfate binding site. Electron density in blue corresponds to the 2Fo − Fc at levels of 1σ. **(c)** Cartoon of sulfate binding pocket of *Se*YihT•sulfate complex depicting hydrogen bonding interactions with active site residues. Interactions in the binding site involved one sulfate oxygen hydrogen bonding to Ser226 (2.6 Å) and the backbone amide of Gly252 (2.8 Å), the second oxygen hydrogen bonding to Arg253 (2 Å), the third oxygen is hydrogen bonding to Gln23 (2.9 Å) and Arg253 (2.9 Å) and a water, and the remaining oxygen hydrogen bonding to the backbone NH of Ser226 (2.8 Å) and to catalytic Lys193 through a bridging water molecule present at a distance of 2.8 Å.

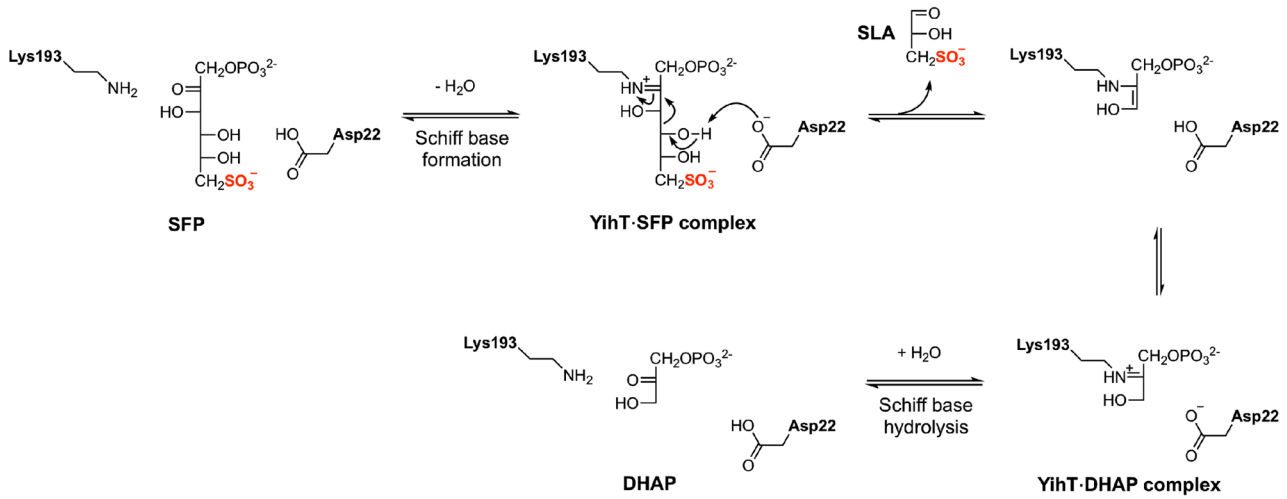


Figure S22. Proposed mechanism for SFP aldolase, with residue numbers corresponding to *Se*YihT.

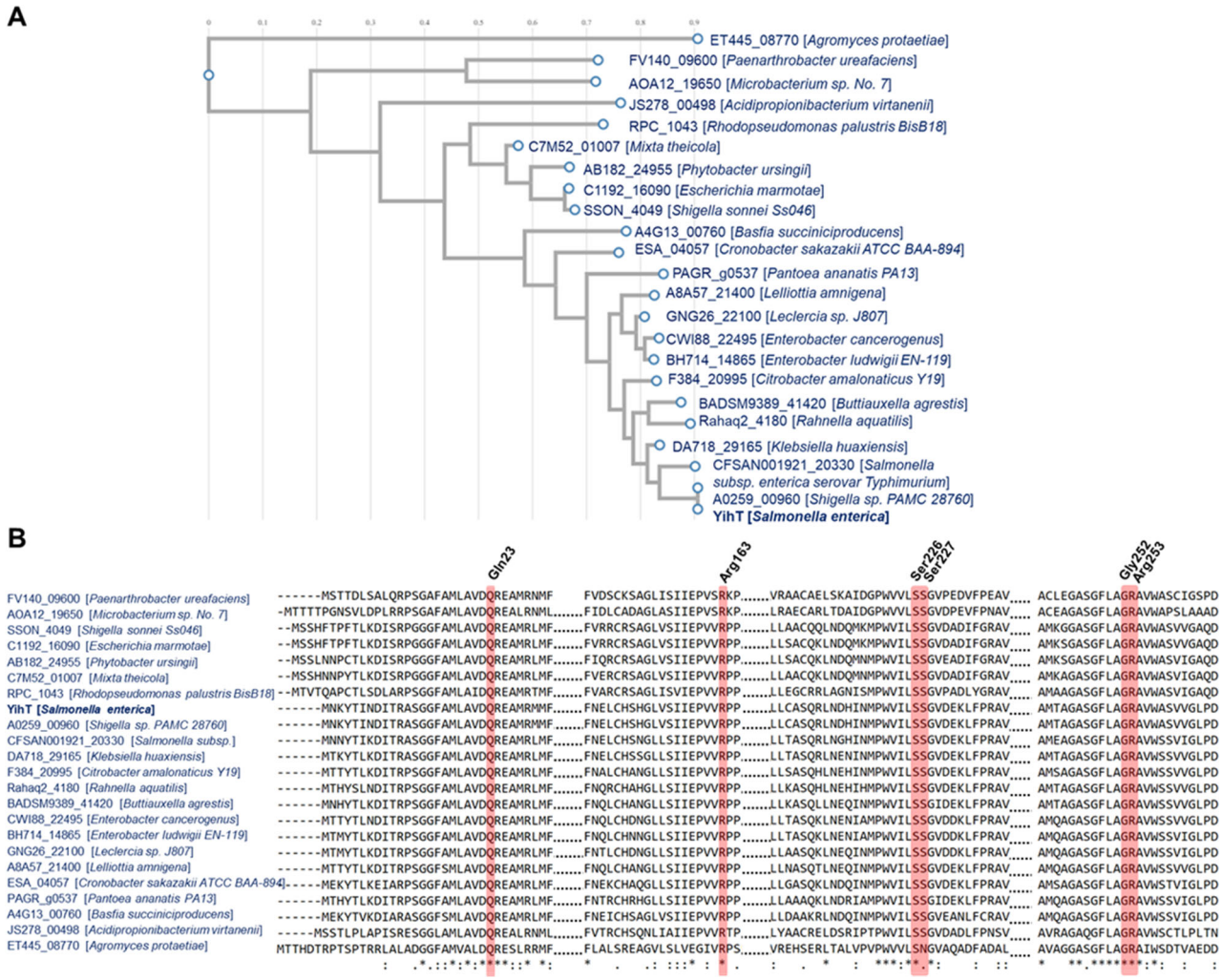


Figure S23. Phylogenetic analysis of SFP aldolases. *A,* Phylogenetic tree showing distribution of bacterial YihT enzymes. *B*, Multiple sequence alignment of YihT enzymes from different bacterial sources showing conserved anion (phosphate/sulfonate) binding residues.

Supplementary Tables

Table S1. Data collection and refinement statistics. Numbers in brackets refer to data for highest resolution shells.

	***Se*YihS-H248A•SF**	***Ec*YihV•AMPPNP•Mg**	***Ec*YihV•ADP•Mg•SF**
Data collection			
Space group	I121	P212121	P3121
Molecules in A.S.U.	6	4	2
Cell dimensions			
a, *b*, *c* (Å)	87.82, 136.83, 230.04	82.09, 82.22, 170.97	94.88, 94.88, 167.43
α, β, ɣ (°)	90, 95.35, 90	90, 90, 90	90, 90, 120
Resolution (Å)	53.13-2.13 (2.17-2.13)	59.21-2.93 (3.11-2.93)	46.17-2.08 (2.13-2.08)
R_{merge}	0.102 (0.439)	0.110 (0.340)	0.098 (1.880)
R_{pim}	0.087 (0.379)	0.070 (0.218)	0.042 (0.797)
I / σI	7.2 (2.3)	10.4 (4.7)	14.6 (1.2)
Completeness (%)	100 (99.9)	100 (100)	99.8 (99.9)
Redundancy	4.3 (4.2)	6.5 (6.5)	11.7 (12.3)
Refinement			
Resolution (Å)	53.0-2.13	59.0-2.93	46.168-2.08
No. unique reflections	150876 (7466)	25634 (4069)	53073 (3869)
R_{work} / R_{free}	0.1805/0.2009	0.2142/0.2564	0.2009/0.2374
No. atoms			
Protein	19937	8364	4375
Ligand/ion	90	124/ 3	84/ 2
Water	686	36	153
B-factors (Å^2)			
Protein	33.0	44.26	56.61
Ligand/ion	33.91	45.49/ 31.06	52.37/ 61.47
Water	30.49	23.43	57.72
R.m.s. deviations			
Bond lengths (Å)	0.0137	0.0146	0.0142
Bond angles (°)	1.73	1.92	1.91
Ramachandran Plot Residues			
In most favourable regions (%)	95.85	97.86	98.99
In allowed regions (%)	3.71	2.05	1.01
Outliers	0.44	0.09	0
PDB code	**7AG4**	**7AGH**	**7AG6**

Table S2. Data collection and refinement statistics. Numbers in brackets refer to data for highest resolution shells.

	***Ec*YihV•SFP**	***Ec*YihT-apo**	***Se*YihT•sulfate**
Data collection			
Space group	P 2 21 21	C121	P1
Molecules in A.S.U	4	2	12
Cell dimensions			
a, *b*, *c* (Å)	47.54, 91.85, 312.14	51.17, 155.83, 89.62	85.21, 107.23, 126.86
α, β, ɣ (°)	90, 90, 90	90, 102.90, 90	107.42, 95.93, 110.55
Resolution (Å)	79.16-2.97 (3.15-2.97)	47.5-2.0 (2.05-2.00)	53.71-1.80 (1.83-1.80)
R_{merge}	0.184 (1.294)	0.13 (0.861)	0.048 (0.259)
R_{pim}	0.100 (0.698)	0.11 (0.742)	0.048 (0.259)
I / σI	7.0 (1.9)	7.0 (1.8)	9.2 (2.3)
Completeness (%)	100 (100)	100 (100)	97.5 (96.1)
Redundancy	8.1 (8.2)	4.2 (4.1)	2.2 (2.1)
Refinement			
Resolution (Å)	79.0-2.97	47.00-2.0	53.01-1.8
No. unique reflections	29364 (4632)	46029 (3393)	352557 (17222)
R_{work} / R_{free}	0.2317/0.2844	0.1861/0.2143	0.1930/0.2073
No. atoms			
Protein	8088	4510	25706
Ligand/ion	57	36 (EDO)	60 (Sulfate)
Water	-	421	1870
B-factors (Å^2)			
Protein	79	27.51	29.03
Ligand/ion	61.69	42.37 (EDO)	25.24
Water	-	32.42	29.34
R.m.s. deviations			
Bond lengths (Å)	0.0151	0.0147	0.0140
Bond angles (°)	1.91	1.83	1.69
Ramachandran Plot Residues			
In most favourable regions (%)	97.13	98.46	97.96
In allowed regions (%)	2.36	1.19	1.34
Outliers (%)	0.51	0.34	0.70
PDB code	**7AGK**	**7AG1**	**7AG7**

Table S3. Data collection and refinement statistics. Numbers in brackets refer to data for highest resolution shells.

	***Se*YihT• SFP/ SeYihT•DHAP**
Data collection	
Space group	P1
Molecules in A.S.U	4
Cell dimensions	
a, *b*, *c* (Å)	48.77, 82.71, 85.20
α, β, ɣ (°)	65.59, 87.73, 78.21
Resolution (Å)	47.69 -1.50 (1.53-1.50)
R_{merge}	0.036 (0.722)
R_{pim}	0.036 (0.722)
$I / \sigma I$	12.8 (1.5)
Completeness (%)	96.6 (94.5)
Redundancy	3.6 (3.7)
Refinement	
Resolution (Å)	42.48-1.5
No. unique reflections	183408 (9048)
R_{work} / R_{free}	0.176/0.204
No. atoms	
Protein	8816
Ligand/ion	54
Water	995
B-factors (Å^2)	
Protein	25.25
Ligand/ion	24.0
Water	35.41
R.m.s. deviations	
Bond lengths (Å)	0.0112
Bond angles (°)	1.65
Ramachandran Plot Residues	
In most favourable regions (%)	98.10
In allowed regions (%)	1.04
Outliers (%)	0.86
PDB code	**7NE2**

Safety statement

No unexpected or unusually high safety hazards were encountered in this work.

Section 1. Chemical Synthesis

General

Sulfoquinovose, potassium 6-deoxy-6-sulfonato-D-fructofuranose (SF), methyl α-sulfoquinovoside, sodium salt (α-MeSQ), sulfofructose-6-phosphate, sulfolactaldehyde and dihydroxypropanesulfonate were synthesized as described.[4-5] Adenosine triphosphate disodium salt trihydrate, adenosine diphosphate sodium salt, D-fructose 6-phosphate disodium salt hydrate, D-fructose 1,6-bisphosphate tetra(cyclohexylammonium) salt, tripotassium citrate monohydrate, dihydroxyacetone phosphate hemimagnesium salt hydrate and phosphoenolpyruvic acid tri(cyclohexylammonium) salt were purchased from Sigma-Aldrich. ^{1}H NMR spectra were recorded on a 500 or 600 instrument using residual protic solvent as the internal standard. HPLC–LC–MS/MS was achieved using a Vanquish Horizon UHPLC system (Thermo Fisher Scientific) HPLC system fitted with a ZIC-HILIC column (5 μm, 150 × 2.1 mm; Merck) connected to a TSQ Altis triple quadrupole mass spectrometer (Thermofisher).

Methyl 2,3,4-tri-*O*-acetyl-6-deoxy-6-iodo-α-D-mannopyranoside

A mixture of methyl α-D-mannopyranoside (5.00 g, 0.0257 mol), triphenylphosphine (10.0 g, 0.038 mol), imidazole (5.30 g, 0.0778 mol), and iodine (9.2 g, 0.0362 mol) in toluene (90 ml) was vigorously stirred at 70 °C for 5 h. The mixture was cooled and water (5 mL) was added and the mixture stirred vigorously for 15 min. The organic phase was extracted with water (3 x 50 ml) and the combined water extract was washed with a small volume of toluene then concentrated under reduced pressure. Acetic anhydride (55 ml) and pyridine (111 ml) were added to the residue and the mixture was stirred at rt for 24 h. Water (5 mL) was added, the mixture was stirred for 5 min. The mixture was diluted with water and extracted with toluene. The toluene phase was washed with water, 1 M HCl and then sat. $NaHCO_3$ solution. The toluene layer was dried ($MgSO_4$), and concentrated. The residue was recrystallized from ethanol affording the product (5.15 g, 46%) as light white crystalline solid (m.p. 95–97 °C; lit.[6] 91–92 °C). $[\alpha]_D^{23}$ +40.6° (*c* 0.81, $CHCl_3$; lit.[6] $[\alpha]_D^{22}$ +37° in $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.31 (1 H, dd, $J_{2,3}$ 3.5, $J_{3,4}$ 10.0 Hz, H3), 5.22 (1 H, dd, $J_{1,2}$ 1.8, $J_{2,3}$ 3.4 Hz, H2), 5.11 (1 H, t, $J_{3,4} = J_{4,5}$ 9.9 Hz, H4), 4.73 (1 H, d, $J_{1,2}$ 1.7 Hz, H1), 3.80 (1 H, ddd, $J_{5,6b}$ 2.5, $J_{5,6a}$ 7.1, $J_{4,5}$ 9.6 Hz, H5), 3.47 (3 H, s, OCH_3), 3.31 (1 H, dd, $J_{5,6b}$ 2.5, $J_{6b,6a}$ 10.8 Hz, H6b), 3.18 (1 H, dd, $J_{5,6a}$ 8.9, $J_{6a,6b}$ 10.9 Hz, H6a), 2.14, 2.06, 1.99 (3 x 3 H, 3 s, Ac); ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.2, 170.01, 169.97 (3 C, C=O), 98.7 (C1), 70.3, 70.2 (2 C, C4,5), 69.8 (C2), 68.8 (C3),

55.7 (OCH_3), 21.0, 20.9, 20.8 (3 C, $COCH_3$), 4.0 (C6); HRMS $(ESI)^+$ *m/z* 431.0196 [$C_{13}H_{20}IO_8$ $(M+H)^+$ requires 431.0197].

Methyl 2,3,4-tri-*O*-acetyl-6-*S*-acetyl-6-deoxy-6-thio-α-D-mannopyranoside

A mixture of methyl 2,3,4-tri-*O*-acetyl-6-deoxy-6-iodo-α-D-mannopyranoside (0.50 g, 0.0012 mol) and KSAc (0.33 g, 0.0029 mol) in acetone (15 ml) was refluxed for 24 h at 60 °C. The solution was cooled and the solvent evaporated. The residue was partitioned between water/EtOAc, and the organic phase was separated and washed with water, then dried ($MgSO_4$), and the solvent evaporated under reduced pressure. The residue was purified by flash chromatography (EtOAc/pet. spirit., 30:70→70:30), affording the product (0.426 g, 97%) as a yellow oil. $[\alpha]_D^{23}$ +50.0° (*c* 0.78, $CHCl_3$); ^1H NMR (500 MHz, $CDCl_3$) δ 5.15 (1 H, dd, $J_{2,3}$ 3.4, $J_{3,4}$ 10.0 Hz, H3), 5.11–5.03 (2 H, m, H2,4), 4.54 (1 H, d, $J_{1,2}$ 1.6 Hz, H1), 3.76–3.69 (1 H, m, H5), 3.28 (3 H, s, OCH_3), 3.15 (1 H, dd, $J_{5,6b}$ 2.8, $J_{6a,6b}$ 14.1 Hz, H6b), 2.90 (1 H, dd, $J_{5,6a}$ 8.0, $J_{6a,6b}$ 14.2 Hz, H6a), 2.24 (3 H, s, SAc), 2.03, 1.99, 1.87 (3 x 3 H, 3 s, 3 x OAc); ^{13}C NMR (126 MHz, $CDCl_3$) δ 194.4 (SC=O), 169.9, 169.8, 169.6 (3 C, OC=O), 98.3 (C1), 69.6, 69.4, 68.9, 68.5 (4 C, C2,3,4,5), 55.0 (OCH_3), 30.3 (2 C, C6,SAc), 20.7, 20.6, 20.5 (3 C, CH_3); HRMS $(ESI)^+$ *m/z* 379.1057 [$C_{15}H_{23}O_9S$ $(M+H)^+$ requires 379.1057].

Potassium methyl 2,3,4-tri-*O*-acetyl-6-deoxy-6-sulfonato-α-D-mannopyranoside

Oxone (1.7 g, 0.0028 mol) was added to a solution of methyl 2,3,4-tri-*O*-acetyl-6-*S*-acetyl-6-deoxy-6-thio-α-D-mannopyranoside (0.41 g, 0.0011 mol) and KOAc (1.1 g, 0.011 mol) in glacial AcOH (3 ml). The mixture was stirred at rt for 24 h. The mixture was concentrated and the residue was purified by flash chromatography (EtOAc/MeOH/H_2O, 15:2:1→5:2:1), affording the product (0.453 g, 99%) as a white crystalline solid (m.p. 149–175 °C). $[\alpha]_D^{23}$ +37.8° (*c* 0.53, $CHCl_3$); ^1H NMR (500 MHz, $CDCl_3$) δ 5.26 (1 H, dd, $J_{2,3}$ 2.7, $J_{3,4}$ 9.8 Hz, H3), 5.20 (1 H, s, H2), 5.07 (1 H, t, $J_{3,4} = J_{4,5}$ 9.3 Hz, H4), 4.74 (1 H, s, H1), 4.25 (1 H, br. s, H5), 3.44 (1H, s, OCH_3), 3.25–3.00 (2 H, m, H6a,6b), 2.12, 2.04, 1.95 (3 x 3 H, 3 s, 3 x Ac); ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.8, 170.0, 169.8 (3 C, C=O), 98.1 (C1), 69.2, 69.0, 68.7 (3 C, C2,3,4), 67.2 (C5), 55.5 (OCH_3), 51.7 (C6), 20.9, 20.8, 20.7 (3 C, CH_3); HRMS $(ESI)^-$ *m/z* 383.0651 [$C_{13}H_{19}O_{11}S$ $(M–K)^-$ requires 383.0654].

Potassium methyl 6-deoxy-6-sulfonato-α-D-mannopyranoside

Potassium methyl 2,3,4-tri-*O*-acetyl-6-deoxy-6-sulfonato-α-D-mannopyranoside (0.95 g, 0.0022 mol) was dissolved in dry MeOH (14 ml) and treated with K_2CO_3 (0.22 g, 0.0016 mol) and stirred at rt for 1.5 hours. The mixture was neutralized with acetic acid, the solvent was evaporated, and the residue was purified by flash chromatography (EtOAc/MeOH/H_2O/, 10:2:1→5:2:1), affording the

product (0.537 g, 81%) as a low-melting (hygroscopic) white crystalline solid (m.p. 120–150 °C). $[\alpha]_D^{23}$ +40.4° (*c* 0.49, H_2O; lit.[6] $[\alpha]_D^{22}$ +52° in H_2O, sodium salt); ^1H NMR (600 MHz, D_2O) δ 4.70 (1 H, d, $J_{1,2}$ 1.5 Hz, H1), 3.96 (1 H, dt, $J_{5,6b}$ 1.1, $J_{4,5} = J_{5,6a}$ 9.8 Hz, H5), 3.91 (1 H, dd, $J_{1,2}$ 1.7, $J_{2,3}$ 3.5 Hz, H2), 3.74 (1 H, dd, $J_{2,3}$ 3.5, $J_{3,4}$ 9.6 Hz, H3), 3.47 (1 H, t, $J_{3,4}= J_{4,5}$ 9.8 Hz, H4), 3.42 (3 H, s, CH_3), 3.38 (1 H, dd, $J_{5,6b}$ 1.3, $J_{6b,6a}$ 14.7 Hz, H6b), 3.08 (1 H, dd, $J_{5,6a}$ 9.8, $J_{6b,6a}$ 14.7 Hz, H6a); ^{13}C NMR (126 MHz, D_2O) δ 100.6 (C1), 70.3 (C3), 69.8 (C2), 69.2 (C4), 68.6 (C5), 54.9 (CH_3), 52.0 (C6); HRMS $(ESI)^-$ *m/z* 257.0328 [$C_7H_{13}O_8S$ $(M–K)^-$ requires 257.0337].

Potassium 6-deoxy-6-sulfonato-D-mannopyranose

Potassium methyl 6-deoxy-6-sulfonato-α-D-mannopyranoside (0.53 g, 1.8 mmol) was dissolved in H_2O (4 ml) and treated with 2 M HCl (1 ml) and stirred at 100 °C for 13 h. The mixture was concentrated and the residue was purified by flash chromatography (EtOAc/MeOH/H_2O, 10:2:1→2:2:1), affording the product (0.501 g, 98%) as a low-melting (hygroscopic) white wax (1:0.75 α:β ratio of anomers). $[\alpha]_D^{23}$ +12.8° (*c* 0.59, H_2O); ^1H NMR (500 MHz, D_2O) δ 5.20 (1 H, s, H1α), 4.96 (1 H, s, H1β), 4.24 (1 H, t, $J_{4,5} = J_{5,6}$ 9.6 Hz, H5α), 4.03–3.97 (2 H, m, H2α,2β), 3.90 (1 H, dd, $J_{2,3}$ 3.2, $J_{3,4}$ 9.6 Hz, H3α), 3.76 (1 H, t, $J_{4,5} = J_{5,6}$ 9.2 Hz, H5β), 3.71 (1 H, dd, $J_{2,3}$ 3.1, $J_{3,4}$ 9.6 Hz, H3β), 3.56 (1 H, t, $J_{3,4} = J_{4,5}$ 9.7 Hz, H4α), 3.53 – 3.42 (3 H, m, H4β,6α,6β), 3.16 (2 H, dd, $J_{5,6}$ 9.6, $J_{6,6}$ 14.7 Hz, H6α,6β); ^{13}C NMR (126 MHz, D_2O) δ 94.0 (C1α), 93.7 (C1β), 72.7 (C3β), 72.3 (C5β), 71.0 (C2β), 70.4 (C2α), 70.1 (C3α), 69.4 (C4α), 69.1 (C4β), 68.6 (C5α), 52.3, 52.2 (2 C, C6α,6β); HRMS $(ESI)^-$ *m/z* 243.0159 [$C_6H_{11}O_8S$ $(M–K)^-$ requires 243.0180].

Sorbitol-6-sulfonate, sodium salt

$NaBH_4$ (15 mg, 0.41 mmol) was added to a solution of 6-deoxy-6-sulfonato-D-glucopyranoside (0.10 g, 0.41 mmol) in H_2O (2.5 ml). The mixture was stirred at rt for 24 h and concentrated residue was evaporated repeatedly with portions of EtOH. The mixture was concentrated and the residue was purified by flash chromatography (EtOAc/MeOH/H_2O, 10:2:1→2:2:1), affording the product (0.102 g, 93%) as light yellow wax. $[\alpha]_D^{23}$ +6.1° (*c* 0.54, H_2O); ^1H NMR (500 MHz, D_2O) δ 4.19 (1 H, ddd, $J_{5,6b}$ 1.9, $J_{4,5}$ 7.5, $J_{5,6a}$ 9.4 Hz, H5), 3.88 (1 H, dd, $J_{2,3}$ 2.4, $J_{3,4}$ 5.7 Hz, H3), 3.85 (1 H, td, $J_{1,2}$ 3.8, $J_{2,3}$ 6.1 Hz, H2), 3.75 (1 H, dd, $J_{1,2}$ 3.8, $J_{1a.1b}$ 11.9 Hz, H1b), 3.67–3.61 (2 H, m, H1a,4), 3.37 (1 H, dd, $J_{5,6b}$ 1.9, $J_{6a,6b}$ 14.6 Hz, H6b), 3.04 (1 H, dd, $J_{5,6a}$ 9.4, $J_{6a,6b}$ 14.6 Hz, H6a); ^{13}C NMR (126 MHz, D_2O) δ 72.9 (C4), 72.7 (C2), 69.4 (C3), 67.7 (C5), 62.4 (C1), 53.8 (C6); HRMS $(ESI)^-$ *m/z* 245.0360 [$C_6H_{13}O_8S$ $(M–Na)^-$ requires 245.0337].

Section 2. Cloning, expression and purification of target enzymes

The gene sequence coding for target enzymes were synthesised commercially with codon optimisation for expression in *E. coli* and ordered through GenScript (*Ec*YihS, *Ec*YihT, *Ec*YihV sub-cloned in pET-21b(+) vector with C-terminal His-tag and *Se*YihS, *Se*YihS-H248A, SeYihT, CsqR-EBD in pET28a(+) with N-terminal His_6 -tag).

The plasmid containing the gene for target enzyme was used to transform *E. coli* BL21(DE3) competent cells for gene expression. Pre-cultures were grown in LB-medium (5 mL) containing 100 μg mL^{-1} ampicillin (or 30 μg mL^{-1} kanamycin) for 18 h at 37 °C with shaking at 200 r.p.m. 1 L volume cultures were inoculated with the pre-culture (5 mL) and incubated at 37°C, with shaking at 200 r.p.m. until an OD_{600} of 0.6-0.8 was reached. Gene expression was induced by addition of IPTG (0.5-1 mM) and shaking was continued overnight at 18 °C with shaking at 200 r.p.m. The cells were then harvested by centrifugation at 5000 g for 20 min and resuspended in 50 mM NaPi buffer pH 7.4, containing 500 mM NaCl and 30 mM imidazole. Cells were disrupted by ultrasonication for 3 x 5 min, 30 s on, 30 s off cycles, and the suspension was centrifuged at 50,000 g for 30 min to yield a clear lysate. The C-terminal His_6-tagged protein was purified using immobilised-metal affinity chromatography (IMAC) using Ni-NTA column, followed by size exclusion chromatography (SEC) (Supplementary Figure 1). For IMAC, the lysate was loaded onto a pre-equilibrated Ni-NTA column, followed by washing with a load buffer (50 mM NaPi, 500 mM NaCl, 30 mM imidazole pH 7.4). The bound protein was eluted using a linear gradient with buffer containing 300 mM imidazole. Protein fractions were pooled, concentrated and loaded onto a HiLoad 16/600 Superdex 200 gel filtration column pre-equilibrated with 50 mM NaPi, 300 mM NaCl pH 7.4 buffer. The protein was concentrated using a Vivaspin® of 10 kDa molecular weight cut-off to a final concentration of 30-65 mg mL^{-1} for crystallization.

Section 3. Amino acid Sequences and Phylogenetic analysis

AA Sequence: wild-type SeYihS

MKWFNTLSHNRWLEQETDRIFNFGKNAVVPTGFGWLGNKGQIKEEMGTHLWITARMLHVYSVAASMGRPGAY
DLVDHGIKAMNGALRDKKYGGWYACVNDQGVVDASKQGYQHFFALLGAASAVTTGHPEARKLLDYTIEVIEK
YFWSEEEQMCLESWDEAFSQTEDYRGGNANMHAVEAFLIVYDVTHDKKWLDRALRIASVIIHDVARNGDYRV
NEHFDSQWNPIRDYNKDNPAHRFRAYGGTPGHWIEWGRLMLHLHAALEARFETPPAWLLEDAKGLFHATIRD
AWAPDGADGFVYSVDWDGKPIVRERVRWPIVEAMGTAYALYTLTDDSQYEEWYQKWWDYCIKYLMDYENGSW
WQELDADNKVTTKVWDGKQDIYHLLHCLVIPRLPLAPGLAPAVAAGLLDINAK

EcYihS wild-type

MKWFNTLSHNRWLEQETDRIFDFGKNSVVPTGFGWLGNKGQIKEEMGTHLWITARMLHVYSVAAAMGRPGAY
SLVDHGIKAMNGALRDKKYGGWYACVNDEGVVDASKQGYQHFFALLGAASAVTTGHPEARKLLDYTIEIIEK
YFWSEEEQMCLESWDEAFSKTEEYRGGNANMHAVEAFLIVYDVTHDKKWLDRAIRVASVIIHDVARNNHYRV
NEHFDTQWNPLPDYNKDNPAHRFRAFGGTPGHWIEWGRLMLHIHAALEARCEQPPAWLLEDAKGLFNATVRD
AWAPDGADGIVYTVDWEGKPVVRERVRWPIVEAMGTAYALYTVTGDRQYETWYQTWWEYCIKYLMDYENGSW
WQELDADNKVTTKVWDGKQDIYHLLHCLVIPRIPLAPGMAPAVAAGLLDINAKLEHHHHHH

EcYihV wild-type

MIRVACVGITVMDRIYYVEGLPTESGKYVARNYTEVGGGPAATAAVAAARLGAQVDFIGRVGDDDTGNSLLA
ELESWGVNTRYTKRYNQAKSSQSAIMVDTKGERIIINYPSPDLLPDAEWLEEIDFSQWDVVLADVRWHDGAK
KAFTLARQAGVMTVLDGDITPQDISELVALSDHAAFSEPGLARLTGVKEMASALKQAQTLTNGHVYVTQGSA
GCDWLENGGRQHQPAFKVDVVDTTGAGDVFHGALAVALATSGDLAESVRFASGVAALKCTRPGGRAGIPDCD
QTRSFLSLFVLEHHHHHH

EcYihT wild-type

MNKYTINDITRASGGFAMLAVDQREAMRMMFAAAGAPAPVADSVLTDFKVNAAKALSPYASAILVDQQFCYR
QVVEQNAIAKSCAMIVAADEFIPGNGIPVDSVVIDRKINPLQIKQDGGKALKLLVLWRSDEDAQQRLDMVKE
FNELCHSHGLVSIIEPVVRPPRRGDKFDREQAIIDAAKELGDSGADLYKVEMPLYGKGPQQELLCASQRLND
HINMPWVILSSGVDEKLFPRAVRVAMTAGASGFLAGRAVWASVVGLPDNELMLRDVCAPKLQQLGDIVDEMM
AKRRLE

SeYihT wild-type

MNNYTIKDITRASGGFAMLAVDQREAMRLMFAAAGAKTPVADSVLTDFKVNAAKILSPYASAVLLDQQFCYR
QAVEQNAVAKSCAMIVAADDFIPGNGIPVDNVVLDKKINAQAVKRDGAKALKLLVLWRSDEDAQQRLNMVKE
FNELCHSNGLLSIIEPVVRPPRCGDKFDREQAIIDAAKELGDSGADLYKVEMPLYGKGARSDLLTASQRLNG
HINMPWVILSSGVDEKLFPRAVRVAMEAGASGFLAGRAVWSSVIGLPDTELMLRDVSAPKLQRLGEIVDEMM
AKRR

EBD-CsqR or EcYihW (C-terminal residues 80-262 of effector binding domain)
MGSSHHHHHHSSGLVPRGSHTAEKRAIAEAVADYLPERCTVFITIGTTVEAVARALLNRRDLRIITNSLRVA
QILYKNQDIEVMVPGGTLRAHNGGIIGPGAVDFIEGFRADYLITSIGAIEHDGTLLEFDVNEALVARTMIKH
ARNTLLVADHTKFAASAAVSIGNARNVRAFFTDAPPPNSFCQLLSEENVELVVAEQEVS

Sequence Alignments

Alignment and phylogenetic reconstructions were performed using the function "build" of ETE3 v3.0.0b32[7] as implemented on the GenomeNet (https://www.genome.jp/tools/ete/). For the

generation of the phylogenetic tree, all proteins that had a Z score of ≥15 and were ≤90% identical to any other selected protein were selected.

Section 4. Biophysical characterization of target enzymes

SEC MALS analysis

Experiments were conducted on a system comprising a Wyatt HELEOS-II multi-angle light scattering detector and a Wyatt rEX refractive index detector linked to a Shimadzu HPLC system (SPD-20A UV detector, LC20-AD isocratic pump system, DGU-20A3 degasser and SIL-20A autosampler). Work was conducted at room temperature (20 ±2°C). Solvent was 0.2 µm filtered before use and a further 0.1 µm filter was present in the flow path. The column was equilibrated with at least 2 column volumes of buffer (50 mM NaPi, 300 mM NaCl pH7.4) before use and flow was continued at the working flow rate until baselines for UV, light scattering and refractive index detectors were all stable. Sample injection volume was 100 µL *Ec*YihU at 6 mg/mL in 50 mM NaPi buffer, 300 mM NaCl pH 7.4 containing 2 mM NADH; Shimadzu LC Solutions software was used to control the HPLC and Astra V software for the HELEOS-II and rEX detectors. The Astra data collection was 1 minute shorter than the LC solutions run to maintain synchronisation. Blank buffer injections were used as appropriate to check for carry-over between sample runs. Data were analysed using the Astra V software. MWs were estimated using the Zimm fit method with degree 1. A value of 0.158 was used for protein refractive index increment (dn/dc).

Nano Differential Scanning Fluorimetry (nanoDSF)

NanoDSF studies were performed on a Prometheus NT.48 (NanoTemper). Data recording and initial analysis was performed with PR.ThermControl software. All protein samples were at 5 $mg.ml^{-1}$ in 25 mM NaPi, 150 mM NaCl at pH 7.4, with a 15 µl capillary load per sample. Experiments were performed in duplicates with the temperature ramp from 15 °C to 95 °C, at 1.0 °C/min with 10% excitation power (YihV) and 100% excitation power (CsqR).

Section 5. Kinetic analysis of target enzymes

^{1}H NMR analysis of YihS catalysed conversion of SQ to SF and SR

Interconversion of SQ, SF and SR by *E. coli* or *S. enterica* YihS was monitored by ^{1}H NMR spectroscopy. Reactions were carried out in a solution of 10 mM SQ, SF or SR in 50 mM sodium phosphate, 150 mM NaCl (pH 7.00) buffer in a total volume of 700 µl. Reactions were heated at 37 °C and initiated by addition of enzyme to a final concentration of 3.92 µM (*Ec*YihS) or 0.44 µM (for *Se*YihS). The solutions were incubated for 24 h, heated at 80 °C for 3 min, concentrated using a rotary

evaporator and dissolved in D_2O and analysed by NMR spectroscopy. ^1H NMR spectra were compared with 20 mM samples of SQ, SR and SF in the same buffer.

The time-course of *Ec*YihS catalyzed conversion of SQ was monitored by ^1H NMR spectroscopy. Enzyme reactions were carried out in a solution of 20 mM SQ in 50 mM sodium phosphate, 150 mM NaCl (pD 7.50) buffer in a total volume of 700 µl using 3.92 µM YihS. ^1H NMR spectra were acquired at 47 °C at various time points for 82 min. In order to assess functional coupling of YihS and an SQ mutarotase, samples prepared as described above were co-incubated with 1.27 µM *Herbaspirillum seropedicae* SQ mutarotase (expressed and purified as described[8]), and monitored over a similar time-period.

Under similar conditions we could not detect any activity towards Glc-6-P (10 mM).

HPLC-MS/MS of YihS catalyzed isomerization of SQ

The HPLC conditions were: 75% B and 25% A for 20 min; solvent A, 22.5 mM NH_4OAc, 2.5% acetonitrile; solvent B, acetonitrile; flow rate, 0.30 ml min^{-1}; column temperature, 37 ºC. The mass spectrometer was operated in ESI negative mode. Quantification was done using the MS/MS SRM mode using Thermo Scientific XCalibur software and normalized with respect to the internal standard, α-MeSQ. Prior to analysis, for each analyte, the sensitivity for each SRM-MS/MS transition was optimized.

Retention times and ESI–MS–MS fragmentation patterns of the analytes and the internal standard were as follows:

For β-SQ: retention time, 13.1 min. ESI–MS *m/z* (precursor ion) 243; ESI–MS^2 of $[M-H]^-$ 243: 183, 153, 123.

For SF: retention time, 8.4 min. ESI–MS *m/z* (precursor ion) 243; ESI–MS^2 of $[M-H]^-$ 243: 183, 153, 123.

For α-SR: retention time, 9.9 min. ESI–MS *m/z* (precursor ion) 243; SR ESI–MS^2 of $[M-H]^-$ 243: 183, 153, 123.

For PNPGlcA: retention time, 1.9 min. ESI–MS *m/z* (precursor ion) 314; PNPGlcA ESI–MS^2 of $[M-H]^-$ 314: 138, 108.

Michaelis-Menten kinetics for YihS

Enzyme kinetics were measured for YihS catalysed conversion of SQ to SF and SR. Reactions were in 50 mM sodium phosphate, 150 mM NaCl (pH 7.00) buffer in a total volume of 250 µl. Reactions were conducted in SQ concentrations ranging from 0.2–40.0 mM, in buffer at 37 ºC using 4.29 nM YihS and the reactions run for a total duration of 1 h. At various times, 10 µl aliquots were sampled from the incubated reaction mixture, diluted with 40 µl H_2O and quenched by heating at 80 ºC for 3

min. A portion (15 µl) of the quenched solution was mixed with 10 µl of internal standard solution comprising PNPGlcA (100 µM) and 25 µl of acetonitrile. The initial rates were calculated using the build-up curves for SF peak area for each sample. The peaks were quantified using Thermo Scientific TraceFinder software and normalised with respect to PNPGlcA and fitted to a second order polynomial curve. Kinetic parameters (k_{cat}, K_M, k_{cat}/K_M) were calculated using the Prism 6 software package (GraphPad Scientific Software) using $Y=V_{max}/K_M\times((K_M\times X)/(K_M+X))$.

Time course of YihS catalyzed interconversion of SQ

Time course for YihS catalysed interconversion of SQ to SF and SR was measured in 50 mM sodium phosphate, 150 mM NaCl (pH 7.00) buffer in a total volume of 250 µl using a 0.190 mM SQ sample. Reaction was conducted in SQ in buffer at 37 ºC using 4.29 nM YihS and the reaction run for a total duration of 2 h. At various times, 10 µl aliquots were sampled from the incubated reaction mixture, diluted with 40 µl H_2O and quenched by heating at 80 ºC for 3 min. A portion (15 µl) of the quenched solution was mixed with 10 µl of internal standard solution comprising PNPGlcA (100 µM) and 25 µl of acetonitrile. Each analyte peak was quantified using Thermo Scientific TraceFinder software and normalised with respect to PNPGlcA and fitted to a one phase decay curve. Peak area for total SQ or SR was calculated by adding the measurements for α and β-anomers; α-SQ peak area was calculated using 0.27 × (β-SQ peak area) and β-SR peak area was calculated using 0.50 × (α-SR peak area).

1D ^{1}H EXSY NMR analysis of spontaneous SF mutarotation rate

1D ^{1}H EXSY measurements for SF uncatalysed mutarotation were performed at 25°C on a Bruker Avance III 600 spectrometer equipped with a TCI cryoprobe using a 1D selective NOESY pulse sequence (selnogp, Bruker). Samples were prepared in D_2O MOPS buffer (pD 7.5) using 5 mM SF. 1D ^{1}H EXSY spectra with mixing time, τ_{mix}, ranging from 20 ms to 1200 ms were acquired with 128 scans. Spectra were subsequently processed and analyzed using TOPSPIN (version 3.2 Bruker). Kinetic data for the conversion of the β-anomer of SF to the α-anomer were obtained by selectively inverting the resonance of H5β at 4.25 ppm using a Gaussian-shaped pulse of 20 ms. Rates were calculated using the Prism 6 software package (GraphPad Scientific Software). Data were fitted to a second-order polynomial function, and the tangent at t=0 provided the rate.

SF kinase (YihV) kinetics

*Ec*YihV catalyzed reactions were conducted in a 200-µL reaction mixture containing 25 mM bistrispropane buffer (pH 7.5), 25 mM KCl, 5 mM $MgCl_2$, 1 mM ATP, 0.1 mg/mL BSA and 36.6 nM *Ec*YihV. SF concentration was varied from 0–9 mM. For reactions where the effect upon rate of

ATP concentration at constant [SF] was measured, SF was at a concentration of 1 mM and [ATP] was varied from 0-9 mM. Other added metabolites specified in the text were at 10 mM concentrations. Linearity of SFP formation at the *Ec*YihV concentration used was determined under identical assay conditions with the SF concentration at 0.1 mM and ATP concentration at 1.0 mM. Enzyme reactions were conducted at 30 ºC for 60 min. The reactions were then quenched by heating at 80 ºC for 4 min. For analysis, a 20 µL aliquot was removed from the quenched mixture, diluted with 40 uL 1:1 acetonitrile/water, mixed with 20 µL of 0.025mM methyl α-sulfoquinovoside, sodium salt (α-MeSQ) (the internal standard) in 1:1 acetonitrile/water and was subjected to analysis by HPLC-MS/MS.

To assess whether *Ec*YihV could act on F6P, 200-µL reaction mixtures was incubated for 60 and 120 min that contained 1 mM F6P, 25 mM bistrispropane buffer (pH 7.5), 25 mM KCl, 5 mM $MgCl_2$, 1 mM ATP, 0.1 mg/mL BSA and 36.6 or 366 nM *Ec*YihV. No FBP was detected by HPLC-MS.

HPLC-MS/MS of YihV catalyzed phosphorylation of SF

The HPLC conditions were: from 90% B to 40% B over 15 min; then 40% B for 5 min; back to 90% B over 1 min (solvent A: 20 mM NH_4OAc in 1% acetonitrile; solvent B: acetonitrile); flow rate, 0.30 ml min^{-1}; injection volume, 2 µl. The mass spectrometer was operated in negative ionization mode. Quantification was done using the MS/MS SRM mode using Thermo Scientific XCalibur software and normalized with respect to the internal standard α-MeSQ. Prior to analysis, for each analyte, the sensitivity for each SRM-MS/MS transition was optimized.

Retention times and ESI–MS–MS fragmentation patterns of the analytes and the internal standard were as follows:

For SFP: retention time, 11.03 min. ESI–MS/MS *m*/*z* of $[M-H]^-$ 323 and product ions 225, 207, 153.

For α-MeSQ (internal standard): retention time: 6.39 min. ESI–MS/MS *m*/*z* of $[M-H]^-$ 257 and product ions 165, 81.

HPLC-MS/MS of YihT catalyzed retroaldol reaction of SFP

HPLC–LC–MS/MS analysis was performed as follows. The HPLC conditions were: from 90% B to 40% B over 15 min; then 40% B for 5 min; back to 90% B over 1 min (solvent A: 20 mM NH_4OAc in 1% acetonitrile; solvent B: acetonitrile); flow rate, 0.30 ml min^{-1}; injection volume, 2 µl. The mass spectrometer was operated in negative ionization mode. Quantification was done using the MS/MS

SRM mode using Thermo Scientific XCalibur software and normalized with respect to the internal standard, α-MeSQ. Prior to analysis, for each analyte, the sensitivity for each SRM-MS/MS transition was optimized.

Retention times and ESI–MS–MS fragmentation patterns of the analytes and the internal standard were as follows:

For SLA: Retention time: 6.21 min. ESI–MS/MS *m*/*z* of $[M-H]^-$ 153, product ions 135, 81, 71.

For α-MeSQ (internal standard): Retention time: 6.31 min ESI–MS/MS *m*/*z* of $[M-H]^-$ 257, product ions 166, 81.

For DHAP: Retention time: 7.60 min. ESI–MS/MS *m*/*z* of $[M-H]^-$ 169, product ions 151, 97, 79.

For SFP: Retention time: 10.91 min. ESI–MS/MS *m*/*z* of $[M-H]^-$ 323, product ions 225, 207, 153.

pH profile

Reactions were conducted in a 50-µL reaction mixture containing 25 mM TrisHCl buffer (pH pH 6.0 – 9.0), 25 mM NaCl, 5 mM $MgCl_2$, 0.5 mM SFP and 11.21 nM *Se*YihT. Reactions were initiated by the addition of the enzyme followed by incubating for 3 hours at 30ºC. After 3 hours, the reactions were quenched by heating at 80 ºC for 3 min. Following quenching, the reaction mixtures were analysed via MS-MS by mixing 20µL of the sample with 20 µL internal standard (0.025 mM a-MeSQ). Enzyme activity was measured by quantifying DHAP produced, normalized with respect to the internal standard (α-MeSQ).

SFP aldolase (*Se*YihT) kinetics

Linearity of reaction progress versus time was confirmed by conducting reactions in a 100-µL reaction mixture containing 25 mM TrisHCl buffer (pH 7.0), 25 mM NaCl, 5 mM $MgCl_2$, 0.5 mM SFP and 5.61 nM *Se*YihT. Reactions were initiated by the addition of the enzyme and incubating for 1 hour at 30ºC. At 20 and 40 minute intervals, 20 µL samples were removed, quenched by heating at 80 ºC for 3 min. Following quenching, the reaction mixtures were analysed via MS-MS by mixing 20µL of the sample with 20 µL internal standard (0.025 mM α-MeSQ).

Individual reactions for Michaelis-Menten kinetics for *Se*YihT were conducted in a 100-µL reaction mixture containing 25 mM TrisHCl buffer (pH 7.0), 25 mM NaCl, 5 mM $MgCl_2$, and 5.61 nM *Se*YihT. SFP concentration was varied from 0–9 mM. Reactions were initiated by the addition of the enzyme and incubating for 40 min at 30ºC. The reactions were then quenched by heating at 80 ºC for 3 min. Following quenching, the reaction mixtures were analysed via MS-MS by mixing 20 µL

of the sample with 20 µL internal standard (0.025 mM α-MeSQ). Quantification was achieved with the aid of a calibration curve for DHAP.

To examine whether *Se*YihT could catalyze the retroaldol reaction of FBP, we studied 100-µL reaction mixtures containing 25 mM TrisHCl buffer (pH 7.0), 25 mM NaCl, 5 mM $MgCl_2$, and 0.5mM FBP. *Se*YihT was studied in three concentrations: 5.61 nM, 11.2 nM and 22.4 nM. Reactions were initiated by the addition of the enzyme and incubating for 60 min at 30 ºC. The reactions were quenched by heating at 80 ºC for 3 min. Following quenching, the reaction mixtures were analysed via MS-MS by mixing 20µL of the sample with 20 µL internal standard (0.025 mM α-MeSQ). Samples were analysed for the production of DHAP and glyceraldehyde phosphate. None was detected at any of the enzyme concentrations used.

To examine whether *Se*YihT can catalyze the aldol reaction of SLA and DHAP to produce SFP, we used a 100-µL reaction mixture containing 25 mM TrisHCl buffer (pH 7.0), 25 mM NaCl, 5 mM $MgCl_2$, 5mM SLA and 5mM DHAP. Reactions were initiated by the addition of *Se*YihT at a final concentration of 11.21 nM and incubating for 120 min at 30 ºC. The reaction was quenched by heating at 80 ºC for 3 min. Following quenching, the reaction mixtures were analysed via MS-MS by mixing 20µL of the sample with 20 µL internal standard (0.025 mM α-MeSQ). Samples were analysed for the production of SFP by SRM MS-MS analysis. The observed SFP was verified by comparing the product ion scan of the SFP produced by the reaction with the product ion scan of an authentic SFP sample. A negative control experiment under identical conditions except for the enzyme, showed that no SFP was produced when the enzyme was absent.

Section 6. Protein Crystallisation

All structure figures were generated using ccp4mg. The protein interactions, surfaces, and assemblies (PISA) server (27) was used to deduce the dimerization interface and buried surface area.

Initial screening and optimised crystallization conditions

Initial screening was performed using commercially available INDEX (Hampton Research), PACT premier and CSSI/II (Molecular Dimensions) screens in 96-well sitting drop trays. Further optimization was carried out in a 48 well sitting drop or 24 well hanging-drop format to obtain optimal crystals for X-ray diffraction. For co-crystallization experiments, 0.1 M stock solution of cofactors ADP or AMPPNP and 0.5 M SF were prepared in water.

A crystal of *Se*YihS-H248A•SF was grown using a 60 mg mL^{-1} protein solution in 50 mM Tris buffer pH 7.5 containing 300 mM NaCl and 10 mM SF in a drop containing 0.15 µL protein: 0.15 µL mother liquor, the latter comprising 25% PEG (polyethylene glycol) 3350 w/v, 0.1 Tris buffer pH 8.5.

For the *Ec*YihV•AMPPNP•Mg structure, crystals were grown with *Ec*YihV at 30 mg mL^{-1} in 50 mM NaPi, 300 mM NaCl buffer pH 7.4 containing 5 mM AMPPNP and 10 mM $MgCl_2$, using a drop containing 0.6 µL protein: 0.5 µL mother liquor, the latter comprising 25% PEG (polyethylene glycol) 3350 w/v, 0.2 M $MgCl_2.6H_2O$, 0.1 M Bis-Tris pH 5.5. The *Ec*YihV•SFP crystal grew from 12 mg mL^{-1} enzyme in 50 mM Tris buffer pH 7.4 containing 10 mM $MgCl_2$ and 5 mM DTT in a drop with 0.15 µL protein: 0.15 µL mother liquor, with the reservoir solution containing 25% PEG 1500 w/v, 0.1 M PCTP (propionic acid, cacodylate, Bis-tris propane) buffer pH 8. The *Ec*YihV•ADP•Mg•SF crystal grew from 30 mg mL^{-1} enzyme in 50 mM NaPi buffer pH 7.4 containing 10 mM ADP, 10 mM $MgCl_2$, 10 mM SF and 10 mM NH_4F in a drop with 0.15 µL protein: 0.15 µL mother liquor, with the reservoir solution containing 25% PEG 1500 w/v, 0.1 M MIB (malonic acid, imidazole, boric acid) buffer pH 5.

A crystal of *Ec*YihT-apo was grown using a 65 mg mL^{-1} protein solution in 50 mM NaPi buffer pH 7.4 containing 300 mM NaCl in a drop containing 0.6 µL protein: 0.5 µL mother liquor, the latter comprising 20% PEG (polyethylene glycol) 3350 w/v, 0.25 M sodium nitrate, 0.1 M Bis-Tris propane buffer pH 6.5. For the *Se*YihT(SO_4^{2-}) structure, crystals were grown with *Se*YihT using a 30 mg mL^{-1} protein solution in 50 mM Tris buffer pH 7.5 containing 300 mM NaCl and 10 mM SF in a drop containing 0.15 µL protein: 0.15 µL mother liquor, the latter comprising 25% PEG (polyethylene glycol) 3350 w/v, 0.2 M Na_2SO_4, 0.1 M Bis-Tris propane pH 7.5. Crystal structure of *Se*YihT•SFP was was obtained using a 30 mg mL^{-1} protein solution of *Se*YihT in 50 mM Tris buffer pH 7.5 containing 300 mM NaCl in a drop containing 0.1 µL protein: 0.2 µL mother liquor, the latter comprising 30% Jeffamine M-600 v/v pH 7 and 0.1 M HEPES buffer pH 7. The crystals were soaked with purified SFP for 1 min and harvested into liquid nitrogen without any cryoprotectant.

The crystals were harvested into liquid nitrogen, using nylon CryoLoops™ (Hampton Research) using mother liquor with 25% ethylene glycol (*Ec*YihT) or with 20% PEG 550 MME *Ec*YihV•AMPPNP•Mg) as cryoprotectants or without any cryoprotectant (*Se*YihS-H248A•SF, *Se*YihT(SO_4^{2-}), *Ec*YihV•SFP and *Ec*YihV•ADP•Mg•SF). Data were collected at Diamond light source, Didcot, Oxfordshire, U.K., on beamlines I03 (*Se*YihS-H248A•SF to 2.13 Å, *Ec*YihV•ADP•Mg•SF to 2.08 Å, *Ec*YihT to 2.0 Å, *Se*YihT•SFP to 1.5 Å and *Se*YihT•SO_4^{2-} to 1.8 Å) and I04 (*Ec*YihV•AMPPNP•Mg and *Ec*YihV•SFP, to 2.93 and 2.97 Å respectively).

Data collection, processing and refinement

The data were processed and integrated using XDS[9] and scaled using SCALA[10] included in the Xia2 processing system.[11] Data collection and refinement statistics are given in Supplementary Table S1-2. The structures were solved using MOLREP,[12] using the 2ZBL.pdb (for *Se*YihS), 1TO3.pdb (*Se*YihT, *Ec*YihT), 1RKD.pdb (*Ec*YihV structures) as initial search models. The structure was built and refined using iterative cycles using Coot[13] and REFMAC,[14] the latter employing local NCS restraints. Following building and refinement of the protein and water molecules, clear residual density was observed in the omit maps for sulfonate ligands. The coordinate and refinement library files for ligands (SF, SFP) were prepared using ACEDRG.[15] SFP was modelled at occupancy of 0.7-0.8 in *Ec*YihV•SFP structure and sulfate ions were modelled in at 0.8-1 occupancy in *Se*YihT•sulfate and SFP/DHAP were covalently linked to Lys193 at occupancy of 0.7-0.8 in *Se*YihT•SFP structures; all other ligands/ions were modelled at occupancy of 1. The coordinate files and structure factors have been deposited in the Protein DataBank (PDB) with accession numbers **7AG4** (*Se*YihS-H248A•SF), **7AGH** (*Ec*YihV•AMPPNP•Mg), **7AG6** (*Ec*YihV•ADP•Mg•SF), **7AGK** (*Ec*YihV•SFP), **7AG1** (*Ec*YihT), **7NE2** (*Se*YihT •SFP) and **7AG7** (*Se*YihT •SO_4^{2-}).

Structure-Based Analyses

Crystal packing interactions were analyzed using PISA.[16] Structural comparisons and structure-based sequence alignments were conducted using a Dali search25 of the Protein Data Bank.[17]

Section 7. Analysis of transcription factor CsqR

Purification of CsqR protein

Plasmid pCsqR for expression and purification of CsqR was constructed according to the published procedure.[18] In brief, CsqR coding sequences were PCR-purified using the *E. coli* K-12 W3110 genome DNA as a template, and inserted into the pET21a vector. The expression plasmid pCsqR was transformed into *E. coli* BL21, and CsqR was expressed by the addition of IPTG in the middle of the exponential phase and purified using the published procedure.[18] The purity of CsqR was more than 95% as judged by SDS-PAGE.[19]

Gel shift assay

Gel shift assay was performed according to the published procedure.[19-20] The 167 bp-long probe of *yihU/yihV* intergenic region was generated by PCR amplification using a pair of primers (yihUV-F; CGCGGATCCAACCCTCTCCTGAATACAGT and yihUV-R; CGCGGATCCGGTCATTCCTTAAACATTTTG) and Ex Taq DNA polymerase (Takara). For gel

shift assay, a mixture of 0.5 pmol each of the probes and CsqR was incubated at 37 °C for 30 min in the gel shift binding buffer (50 mM Tris-HCl, pH 7.9, 40 mM KCl, 5 mM $MgCl_2$ and 1 mM DTT). After addition of a DNA loading solution, the mixture was directly subjected to 5% PAGE. DNA in gels was stained by GelRed (Biotium) and detected using LuminoGraph (Atto).

AFM observation

AFM observation of CsqR-DNA complexes was carried out as described previously.[19, 21] DNA probes were prepared by PCR amplification using pairs of primers. A mixture of 0.1 pmol probe DNA and 10–100 pmol CsqR was incubated in 10 μl of the binding buffer (50 mM Tris-HCl, pH 7.9, 40 mM KCl, 5 mM $MgCl_2$ and 1 mM DTT) for 30 min at 37°C. The samples were directly spotted onto a freshly cleaved mica surface and stored at room temperature for 10 min to facilitate attachment of the sample to the substrate. The mica surface was washed thoroughly with the imaging buffer (50 mM Tris-HCl, pH 9, 5 mM $MgCl_2$, 40 mM KCl and 1 mM DTT). The samples were imaged in the imaging buffer using a High-Speed AFM System (Nano Explorer Model, RIBM). Images were taken at a rate of 2 per second using cantilevers with a 0.10 Nm spring constant and a resonance frequency in water of 0.6 MHz (Olympus, BL-AC10DS). The electron beam-deposited tips were fabricated using ferrocene powder and these were used in AFM imaging to obtain high-resolution images. All AFM images were viewed and analyzed by Kodec 4.4.7.39.[22] A low-pass filter and a flattening filter were applied to individual images to remove spike noise and to make the *xy*-plane flat, respectively.

References

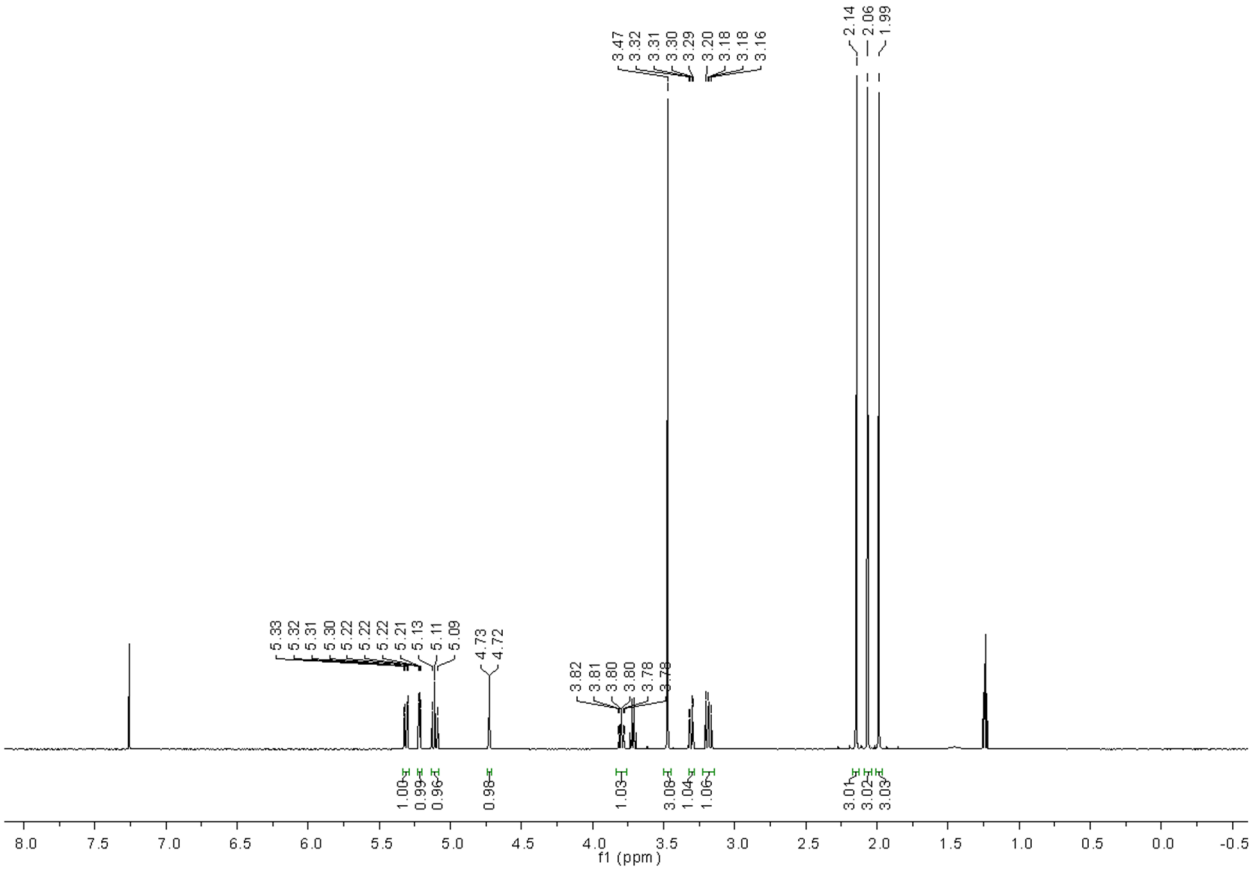
1. Gardberg, A., Abendroth, J., Bhandari, J., Sankaran, B., Staker, B. , Structure of fructose bisphosphate aldolase from *Bartonella henselae* bound to fructose 1,6-bisphosphate. *Acta Crystallogr. Sect. F* **2011,** *67*, 1051--1054.
2. Lorentzen, E.; Siebers, B.; Hensel, R.; Pohl, E., Mechanism of the Schiff Base Forming Fructose-1,6-bisphosphate Aldolase: Structural Analysis of Reaction Intermediates. *Biochemistry* **2005,** *44*, 4222-4229.
3. Dadinova, L. A.; Shtykova, E. V.; Konarev, P. V.; Rodina, E. V.; Snalina, N. E.; Vorobyeva, N. N.; Kurilova, S. A.; Nazarova, T. I.; Jeffries, C. M.; Svergun, D. I., X-Ray Solution Scattering Study of Four *Escherichia coli* Enzymes Involved in Stationary-Phase Metabolism. *PLOS One* **2016,** *11*, e0156105.
4. Abayakoon, P.; Epa, R.; Petricevic, M.; Bengt, C.; Mui, J. W. Y.; van der Peet, P. L.; Zhang, Y.; Lingford, J. P.; White, J. M.; Goddard-Borger, E. D.; Williams, S. J., Comprehensive synthesis of substrates, intermediates and products of the sulfoglycolytic Embden-Meyerhoff-Parnas pathway. *J. Org. Chem.* **2019,** *84*, 2910-2910.
5. Zhang, Y.; Mui, J. W.; Arumaperuma, T.; Lingford, J. P.; Goddard-Borger, E. D.; White, J. M.; Williams, S. J., Concise synthesis of sulfoquinovose and sulfoquinovosyl diacylglycerides, and development of a fluorogenic substrate for sulfoquinovosidases. *Org. Biomol. Chem.* **2020,** *18*, 675-686.

6. Lehmann, J.; Benson, A. A., Plant Sulfolipid .9. Sulfosugar Syntheses from Methyl Hexoseenides. *J. Am. Chem. Soc.* **1964,** *86*, 4469-4472.
7. Huerta-Cepas, J.; Serra, F.; Bork, P., ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Molecular biology and evolution* **2016,** *33*, 1635-8.
8. Abayakoon, P.; Lingford, J. P.; Jin, Y.; Bengt, C.; Davies, G. J.; Yao, S.; Goddard-Borger, E. D.; Williams, S. J., Discovery and characterization of a sulfoquinovose mutarotase using kinetic analysis at equilibrium by exchange spectroscopy. *Biochem. J.* **2018,** *475*, 1371-1383.
9. Kabsch, W., Xds. *Acta Crystallogr., Section D: Biol. Crystallogr.* **2010,** *66*, 125-132.
10. Evans, P., Scaling and assessment of data quality. *Acta Crystallogr. Sect. D* **2006,** *62*, 72-82.
11. Winter, G., xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **2010,** *43*, 186-190.
12. Vagin, A.; Teplyakov, A., MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **1997,** *30*, 1022-1025.
13. Emsley, P.; Cowtan, K., Coot: Model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004,** *60*, 2126-2132.
14. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J., Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. Sect. D* **1997,** *53*, 240-255.
15. Long, F.; Nicholls, R. A.; Emsley, P.; Grazulis, S.; Merkys, A.; Vaitkus, A.; Murshudov, G. N., AceDRG: a stereochemical description generator for ligands. *Acta Crystallogr. Sect. D* **2017,** *73*, 112-122.
16. Krissinel, E., Stock-based detection of protein oligomeric states in jsPISA. *Nucleic Acids Res.* **2015,** *43*, W314-W319.
17. Holm, L.; Rosenström, P., Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **2010,** *38*, W545-W549.
18. Yamamoto, K.; Hirao, K.; Oshima, T.; Aiba, H.; Utsumi, R.; Ishihama, A., Functional Characterization in Vitro of All Two-component Signal Transduction Systems from *Escherichia coli*. *J. Biol. Chem.* **2005,** *280*, 1448-1456.
19. Shimada, T.; Yamamoto, K.; Nakano, M.; Watanabe, H.; Schleheck, D.; Ishihama, A., Regulatory role of CsqR (YihW) in transcription of the genes for catabolism of the anionic sugar sulfoquinovose (SQ) in Escherichia coli K-12. *Microbiology (Reading, England)* **2019,** *165*, 78-89.
20. Shimada, T.; Bridier, A.; Briandet, R.; Ishihama, A., Novel roles of LeuO in transcription regulation of *E. coli* genome: antagonistic interplay with the universal silencer H-NS. *Mol. Microbiol.* **2011,** *82*, 378-397.
21. Sugino, H.; Usui, T.; Shimada, T.; Nakano, M.; Ogasawara, H.; Ishihama, A.; Hirata, A., A structural sketch of RcdA, a transcription factor controlling the master regulator of biofilm formation. *FEBS Lett.* **2017,** *591*, 2019-2031.
22. Ngo, K. X.; Kodera, N.; Katayama, E.; Ando, T.; Uyeda, T. Q. P., Cofilin-induced unidirectional cooperative conformational changes in actin filaments revealed by high-speed atomic force microscopy. *eLife* **2015,** *4*, e04806.

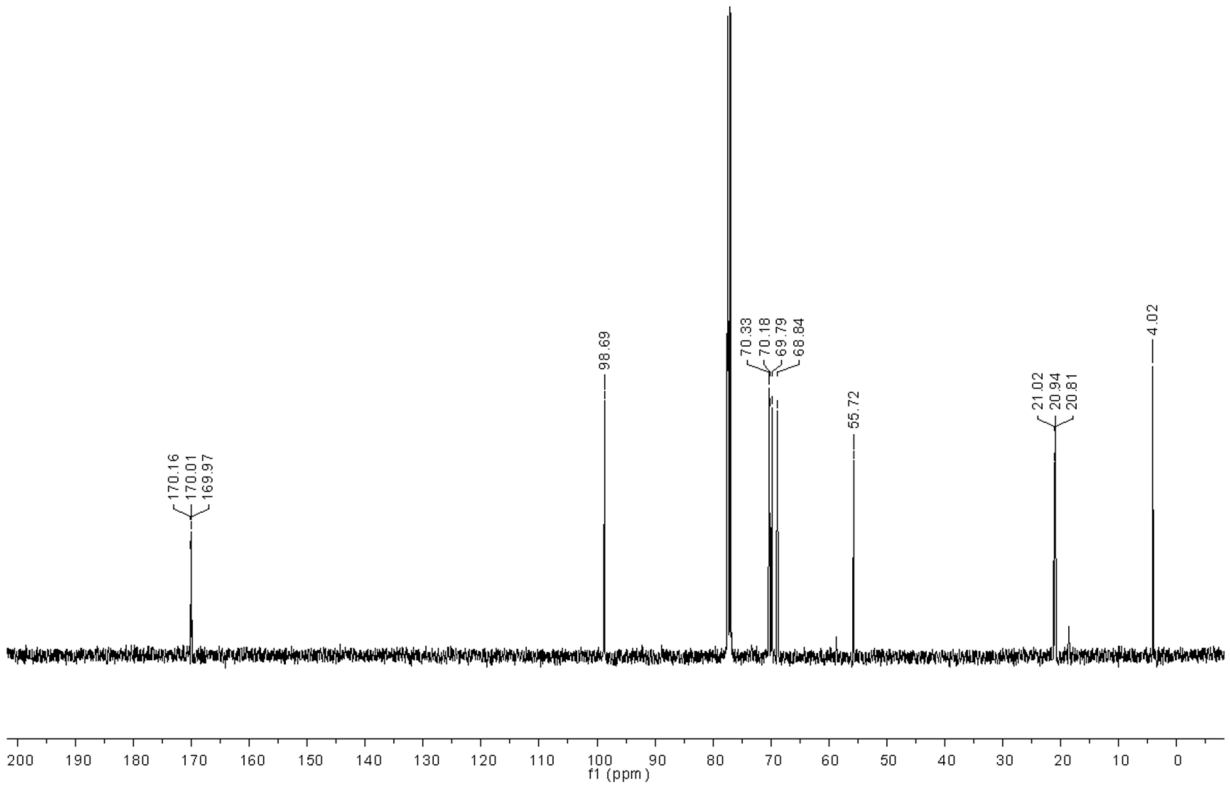
NMR spectra

Methyl 2,3,4-tri-*O*-acetyl-6-deoxy-6-iodo-α-D-mannopyranoside

^{1}H NMR

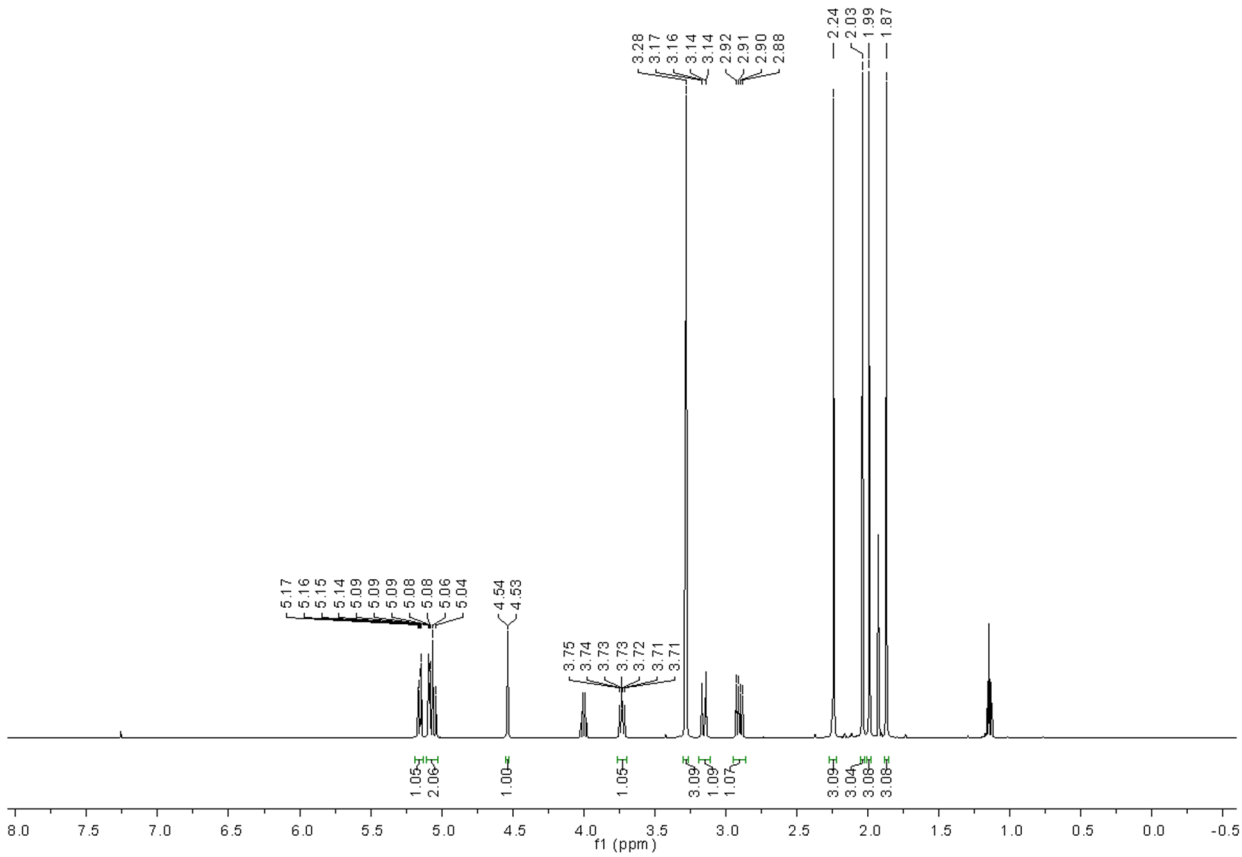


^{13}C NMR

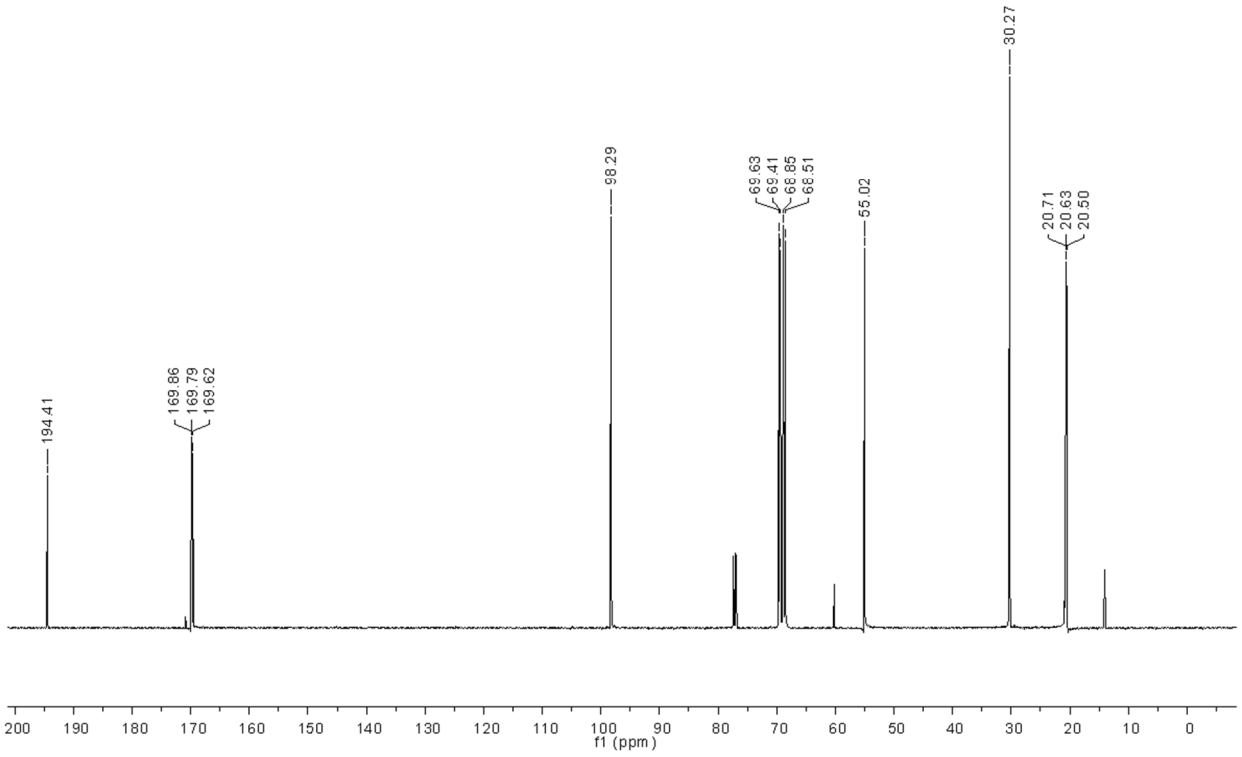


Methyl 2,3,4-tri-*O*-acetyl-6-*S*-acetyl-6-deoxy-6-thio-α-D-mannopyranoside

^{1}H NMR

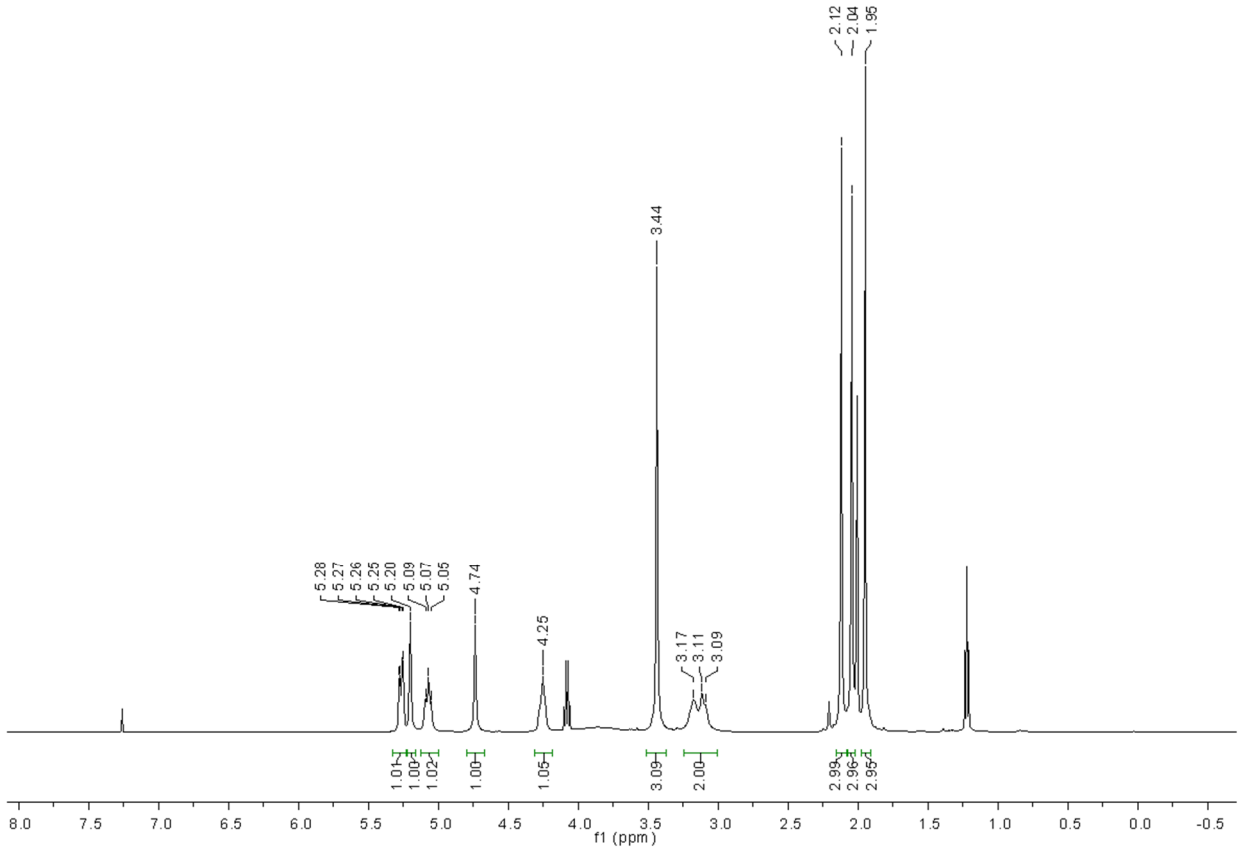


^{13}C NMR

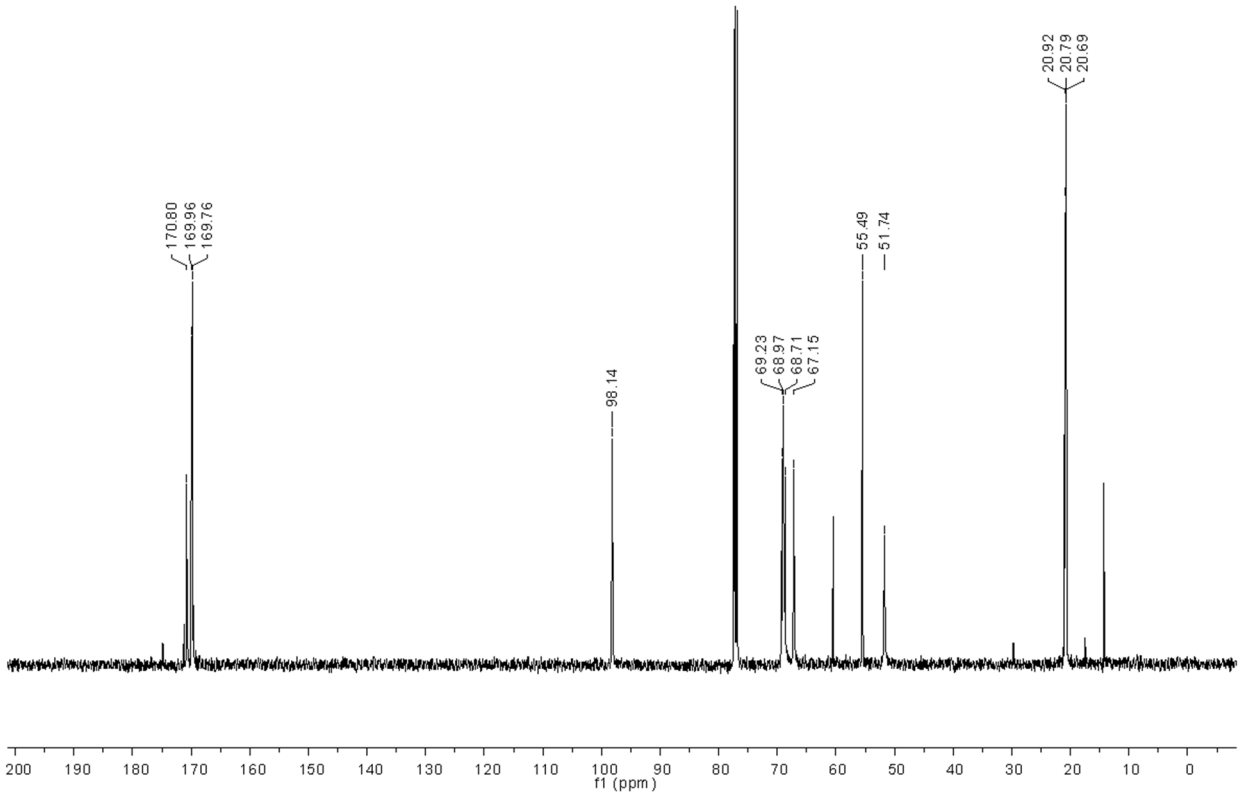


Potassium methyl 2,3,4-tri-*O*-acetyl-6-deoxy-6-sulfonato-α-D-mannopyranoside

^{1}H NMR

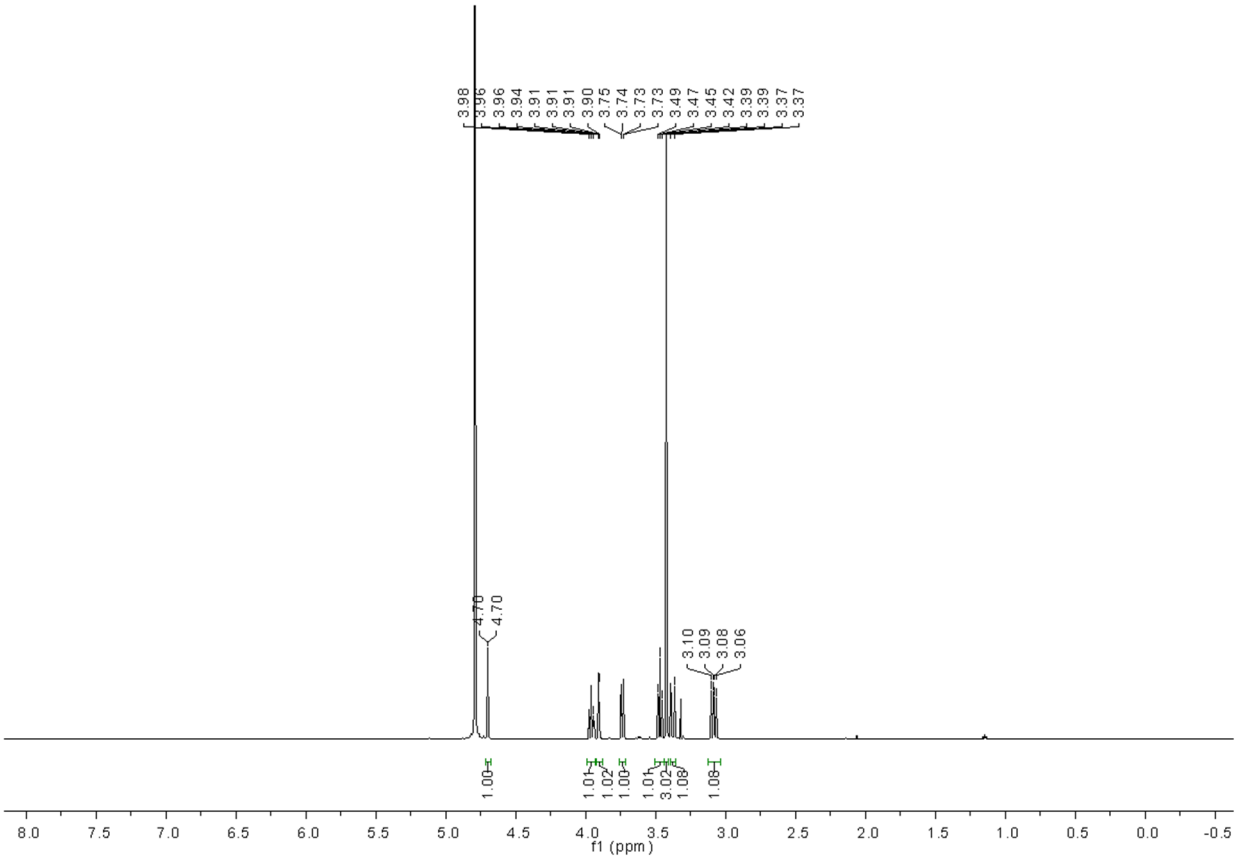


^{13}C NMR

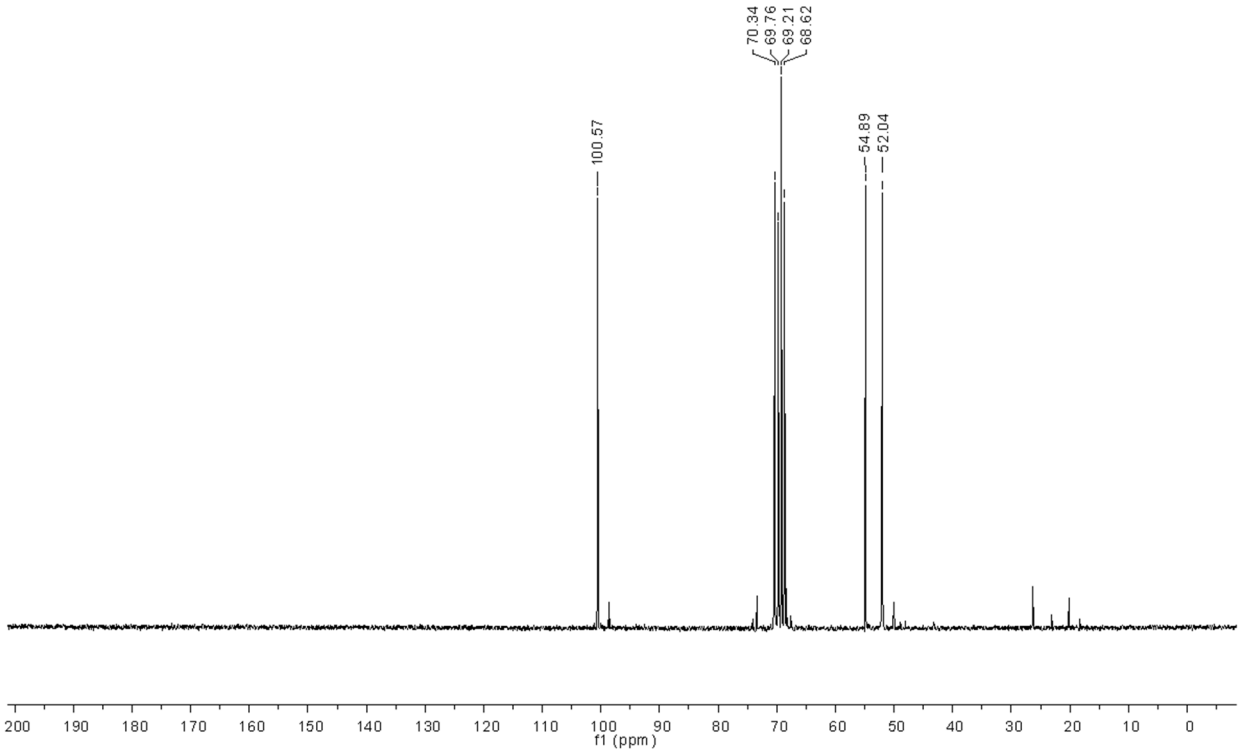


Potassium methyl 6-deoxy-6-sulfonato-α-D-mannopyranoside

^{1}H NMR

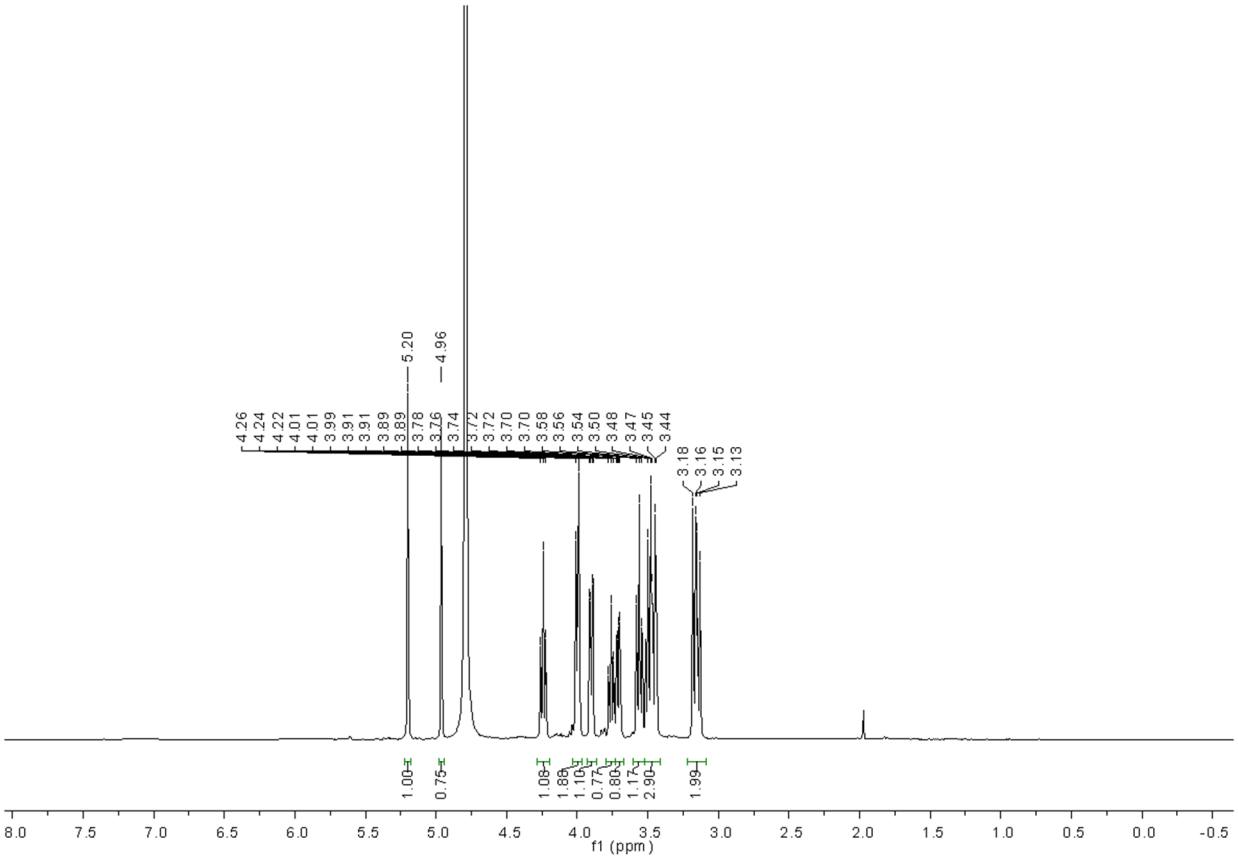


^{13}C NMR

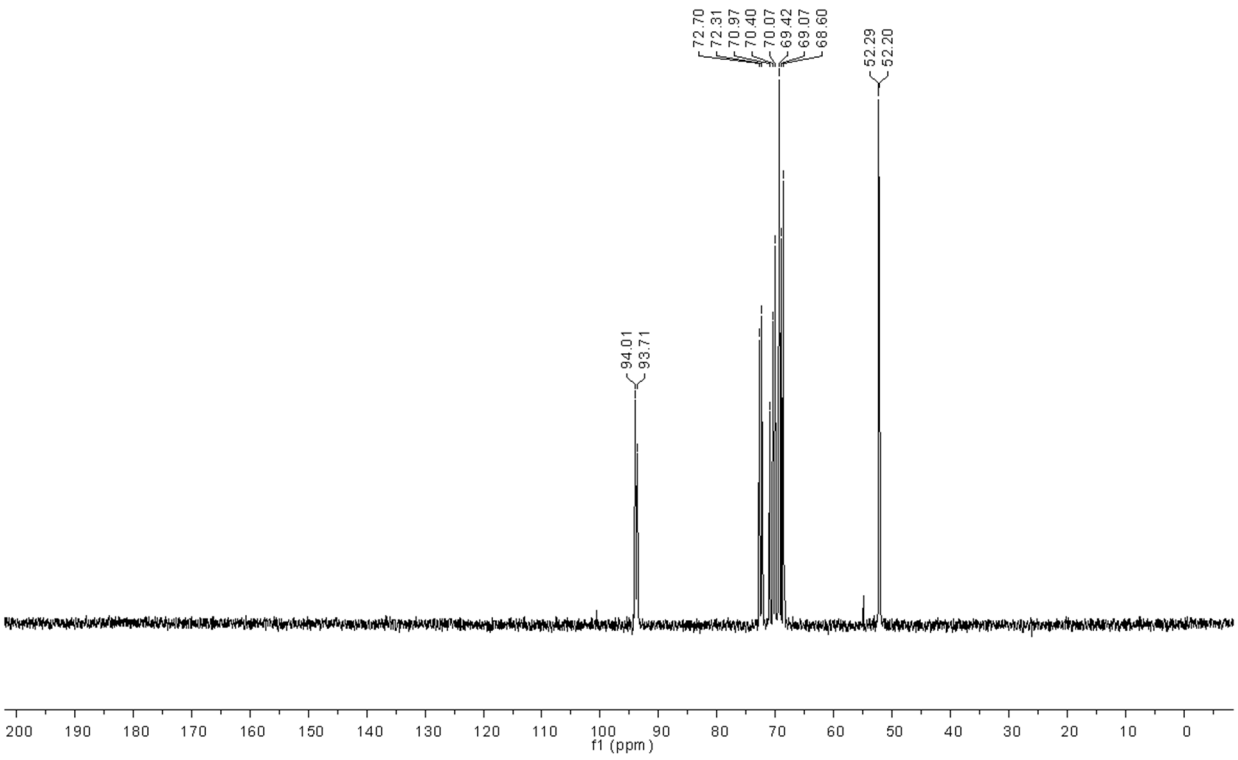


Potassium 6-deoxy-6-sulfonato-D-mannopyranose

^{1}H NMR

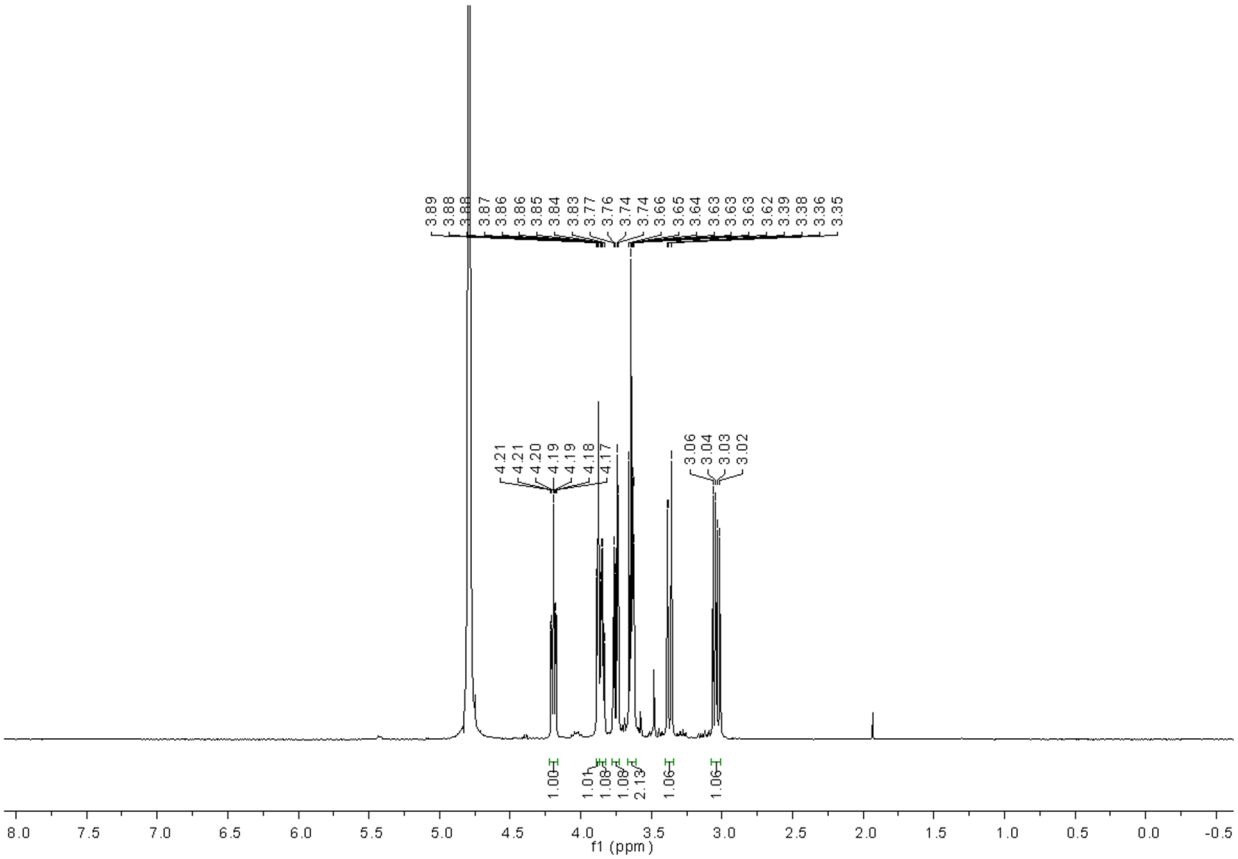


^{13}C NMR



Sorbitol-6-sulfonate, sodium salt

^{1}H NMR



^{13}C NMR