Article:
Sharma, Mahima orcid.org/0000-0003-3960-2212, Abayakoon, Palika, Epa, Ruwan et al. (9 more authors) (2021) The Molecular Basis of Sulfosugar Selectivity in Sulfoglycolysis. ACS Central Science. pp. 476-487. ISSN: 2374-7943

https://doi.org/10.1021/acscentsci.0c01285

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,[6] Janice W.-Y. Mui,[2] Akira Ishihama,[7] Ethan D. Goddard-Borger,*[,3,4] Gideon J. Davies,*[,1] Spencer J. Williams*[,2]

[1]York Structural Biology Laboratory, Department of Chemistry, University of York YO10 5DD, U.K.

[2]School of Chemistry and Bio21 Molecular Science and Biotechnology Institute and University of Melbourne, Parkville, Victoria 3010, Australia

[3]ACRF Chemical Biology Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria 3010, Australia

[4]Department of Medical Biology, University of Melbourne, Parkville, Victoria 3010, Australia

[5]Meiji University, School of Agriculture, Kawasaki, Kanagawa, Japan

[6]Institute for Frontier Life and Medical Sciences, Kyoto University, Sakyo-ku, Kyoto, Japan

[7]Micro-Nano Technology Research Center, Hosei University, Koganei, Tokyo, Japan

Keywords: metabolism; sulfur cycle; enzyme mechanism; glycobiology; structural biology

Abstract:

The sulfosugar sulfoquinovose (SQ) is produced by essentially all photosynthetic organisms on earth and is metabolized by bacteria through the process of sulfoglycolysis. The sulfoglycolytic Embden-Meyerhof-Parnas pathway metabolises SQ to produce dihydroxyacetone phosphate and sulfolactaldehyde and is analogous to the classical Embden-Meyerhof-Parnas glycolysis pathway for the metabolism of glucose-6-phosphate, though the former only provides one C3 fragment to central metabolism, with excretion of the other C3 fragment as dihydroxypropanesulfonate. Here, we report a comprehensive structural and biochemical analysis of the three core steps of sulfoglycolysis catalyzed by SQ isomerase, sulfofructose (SF) kinase and sulfofructose-1-phosphate (SFP) aldolase. Our data shows that despite the superficial similarity of this pathway to glycolysis, the sulfoglycolytic enzymes are specific for SQ metabolites and are not catalytically active on related metabolites from glycolytic pathways. This observation is rationalized by 3D structures of each enzyme, which reveal the presence of conserved sulfonate-binding pockets. We show that SQ isomerase acts preferentially on the β-anomer of SQ and reversibly produces both SF and sulforhamnose (SR), a previously unknown sugar that acts as a transcriptional regulator for the transcriptional repressor CsqR that regulates SQ-utilization. We also demonstrate that SF kinase is a key regulatory enzyme for the pathway that experiences complex modulation by the metabolites AMP, ADP, ATP, F6P, FBP, PEP, and citrate, and we show that SFP aldolase reversibly synthesizes SFP. This body of work provides fresh insights into the mechanism, specificity and regulation of sulfoglycolysis and has important implications for understanding how this biochemistry interfaces with central metabolism in prokaryotes to process this major repository of biogeochemical sulfur.

Introduction

Global sulfur cycling involves the interconversion of inorganic, gaseous and organic forms of the element.[1] While there are innumerable biosulfur species on earth, just a handful comprise the majority of sulfur in the biosphere, namely the amino acids cysteine and methionine, the osmolytes dimethylsulfoniopropionate (DMSP)[2] and dimethylsulfoxonium propionate,[3] and the sulfosugar sulfoquinovose (SQ).[4] Global production of SQ is estimated to be 10^{10} tonnes per annum, comprising up to half of all biosulfur,[5] and is produced by essentially all photosynthetic organisms. SQ exists as the carbohydrate headgroup of the plant and cyanobacterial sulfolipid sulfoquinovosyl diacylglycerol (SQDG) within the photosynthetic membranes of photoautotrophs.[4] SQDG is believed to primarily enter the environment through decomposition of photosynthetic tissues and herbivory, whereupon it becomes available to environmental and intestinal bacteria.[6-10] The catabolism of SQ, a process termed sulfoglycolysis,[6-8] is therefore of great relevance to understanding the biogeochemical sulfur cycle.

The first sulfoglycolytic pathway discovered comprises a variant of the classical Embden-Meyerhof-Parnas glycolysis pathway, and is termed the sulfoglycolytic Embden-Meyerhof-Parnas (sulfo-EMP) pathway **(Figure 1a)**.[7] It was initially described in *E. coli* and involves the importation of SQ and its glycosides (which undergo intracellular hydrolysis to liberate SQ), catabolic enzymes to produce dihydroxyacetone phosphate (DHAP) and sulfolactaldehyde (SLA), a reductase to convert SLA to dihydroxypropanesulfonate (DHPS), and a permease to export DHPS from the cell.[7] The pathway is encoded within a 10-gene cluster (*ompL* and *yihO-W*; where *yihW* was renamed *csqR*) that encodes a transcriptional regulator (CsqR, formerly YihW);[11] a sulfolipid porin (OmpL) and two transmembrane permeases (YihO, YihP); a sulfoquinovosidase for cleavage of SQ glycosides (YihQ);[12,13] an SQ mutarotase (YihR);[14] three core sulfoglycolytic enzymes: SQ-sulfofructose (SF) isomerase (YihS), SF kinase (YihV), and SFP aldolase (YihT); and finally SLA reductase (YihU),[15] which forms DHPS for export **(Figure 1b)**. The three core sulfoglycolytic enzymes YihS, YihV and YihT perform roles analogous to three enzymes involved in upper glycolysis: glucose-6-phosphate (G6P) isomerase, phosphofructose kinase (PFK), and fructose bisphosphate (FBP) aldolase, respectively. However, nothing is known about the specificity of these sulfoglycolytic enzymes for intermediates from sulfoglycolysis versus glycolysis. Metabolic flux through the glycolytic pathway is tightly regulated with the most important control exerted upon PFK, with allosteric activation and inhibition by assorted cellular

metabolites such as adenosine monophosphate (AMP), adenosine diphosphate (ADP), adenosine triphosphate (ATP), fructose 6-phosphate (F6P), FBP, phosphoenolpyruvate (PEP), and citrate.[16] It is unknown whether similar metabolites control flux through the sulfoglycolytic pathway.

In prokaryotes, the gluconeogenesis pathway uses two enzymes of upper glycolysis (G6P-F6P isomerase, FBP aldolase) and the pathway specific enzyme fructose 1,6-bisphosphatase to drive the reverse of upper glycolysis and support the biosynthesis of G6P, which enters the pentose phosphate pathway (PPP) and cell wall biosynthesis.[17] In contrast, the sulfo-EMP pathway is catabolic and is presumed to operate exclusively in the forward direction and does not provide key intermediates for gluconeogenesis such as F6P and G6P. Sulfoglycolytic cells must therefore utilize the traditional gluconeogenesis pathway to satisfy demands for cell wall biosynthesis and the PPP.[10] Since these pathways must operate in tandem, it is presumably important that the sulfoglycolytic enzymes exhibit high selectivity for sulfosugar intermediates and limited activity on the analogous glycolytic intermediates to ensure the chemical unidirectionality of each pathway and avoid futile cycles.

In this work we illuminate how the three core enzymes of the sulfo-EMP pathway recognize the unique sulfonate groups of their substrates and highlight key features of their catalytic mechanisms. We show that YihS catalyses the interconversion of SQ, SF and the previously unknown metabolite sulforhamnose (SR), and provide evidence that SR, as well with SQ, can act as a transcriptional derepressor through CsqR. We demonstrate the activation and inhibition of YihV by several common cellular metabolites, revealing that this enzyme acts as an important control-point for regulating flux through the sulfoglycolytic pathway. Importantly, we also show that this triad of enzymes exhibit high selectivity for sulfoglycolytic intermediates, ensuring sulfoglycolysis and gluconeogenesis can operate in tandem.

Results

SQ-SF isomerase also produces sulforhamnose

YihS was identified as the SQ-SF isomerase in the *E. coli* sulfo-EMP pathway by Schleheck and co-workers.[7] It is classified as a member of Pfam database family PF07221,[18] which includes enzymes that catalyze the epimerization of the alpha carbon in acyclic aldoses (the Lobry de Bruyn-Alberta van Ekenstein reaction). We chose to revisit the activity of *E. coli* YihS by incubating this protein with SQ in phosphate buffered saline, heat-inactivating the enzyme, then exchanging the sample into deuterated solvent for ^{1}H NMR analysis. This analysis revealed that YihS produced two new products **(Figure 2a)**, which were not produced in the absence of enzyme. One product was identified as SF through comparison with authentic material.[19] Resonances for the second new compound did not match either anomer of SQ or SF, particularly singlets in the ^{1}H NMR spectrum at δ 4.86 and 5.10 ppm, and in the ^{13}C NMR spectrum at δ 74.6 and 71.7 ppm. Further insight into the structure of the unknown product was obtained by conducting an additional experiment on the reaction of SQ with *Ec*YihS, this time in buffered D_2O. Under these conditions, the anomeric signals for SQ collapsed from a doublet to a singlet, consistent with H/D exchange at C2, and the signals for the unknown in the ^{13}C NMR spectrum changed to triplets shifted slightly upfield, consistent with incorporation of deuterium (I = 1). Prior to the discovery that the sulfo-EMP operon encodes for sulfoglycolysis, Itoh *et al.* reported that recombinantly produced *Ec*YihS can interconvert mannose, fructose and glucose.[20] We therefore speculated that the unknown compound could be the C2-epimer of SQ, namely sulforhamnose (SR). Comparison of chemically synthesized SR with the mixture of products obtained from the reaction of *Ec*YihS and SQ provided a good match for the unknown compound.

To confirm that SQ, SF and SR are in equilibrium we individually treated each compound with *Ec*YihS until equilibrium was established. ^{1}H NMR spectra of each reaction revealed identical mixtures, providing evidence that all three compounds are substrates of *Ec*YihS (Figure S1). At equilibrium under these experimental conditions, the ratios of the three sulfosugars were SF:SR:SQ = 30:21:49. We established an HPLC-MS assay for the YihS-catalyzed isomerization reaction using a ZIC-HILIC column to follow the time-course for equilibration of SQ to the mixture of sulfosugars **(Figure 2b)**. The normalized plot for SQ depletion shows that formation of SF and SR is coincident with consumption of SQ. This data is consistent with an enzymatic mechanism that involves the formation of acyclic SQ and

enzyme-catalyzed deprotonation of C2 to form a 1,2-enediol (**Figure 2c**). Protonation at C1 then forms SF, protonation at C2 from the 'bottom' face forms SR, while protonation on the 'top' face at C2 regenerates SQ.

As the sulfo-EMP pathway is a catabolic process with the isomerization of SQ to SF the most physiologically-relevant, we chose to investigate the kinetics of this process in greater detail. *Ec*YihS exhibited Michaelis-Menten parameters for formation of SF from SQ, with k_{cat} = (7.90 ± 0.27) × 10^1 s^{-1}, K_M = 1.89 ± 0.28 mM and k_{cat}/K_M = (4.17 ± 0.51) × 10^4 M^{-1} s^{-1} (Figure S2). Comparison of this data with that for D-mannose reported by Itoh and co-workers[20] shows that SQ is 178-fold better as a substrate for *Ec*YihS (in terms of k_{cat}/K_M). Under similar conditions no activity towards G6P (10 mM) was detected (as previously reported by Itoh and co-workers[20]), demonstrating that *Ec*YihS can discriminate between a 6-phosphate and sulfonate, and that no activity is expected for YihS on glycolysis intermediates.

SQ exists as a mixture of α- and β-anomers, which are interconverted by SQ mutarotase.[14] It is unknown if *Ec*YihS acts on a single anomer or both. The mechanism utilized by *Ec*YihS to catalyze the isomerization of SQ to SR and SF also results in H/D exchange of the proton at C2 of SQ when the reaction is performed in D_2O. This leads to the ^{1}H NMR signal for the anomeric proton of SQ changing from a doublet to a singlet. We therefore monitored the multiplicity of H1 in the two SQ anomers as a function of time in the presence of *Ec*YihS. The time course ^{1}H NMR spectra showed a rapid collapse of the doublet corresponding to H1 of β-SQ, and a slower collapse of the doublet corresponding to H1 of α-SQ (Figure S3). This data is consistent with stereospecific H/D exchange at C2 of β-SQ, demonstrating that the enzyme preferentially acts on β-SQ. We interpret the slow H/D exchange responsible for the conversion of the doublet of H1 of α-SQ to a singlet to be a result of spontaneous mutarotation from the deuterated β-anomer (2-D-β-SQ), though we cannot discount a low rate of catalysis by *Ec*YihS. As an added control, this experiment was repeated in the presence of the SQ mutarotase *Hs*SQM,[14] which increases the rate of interconversion of α-SQ and β-SQ. This revealed simultaneous collapse of anomeric doublets for both α-SQ and β-SQ. Together these experiments confirm that the mechanism for H/D exchange of α-SQ is primarily by mutarotation of β-SQ, which is the substrate of *Ec*YihS. This study also demonstrates the functional significance of the SQ mutarotase in sulfoglycolysis: it ensures that the α-SQ product of the sulfoquinovosidase YihQ can be efficiently utilized by the SQ-SF isomerase YihS, which has a clear preference for β-SQ and little-to-no activity on α-SQ.

Salmonella enterica possesses a sulfo-EMP operon syntenic to that from *E. coli* and the crystal structures of wild-type *S. enterica* YihS (PDB ID: 2AFA) and *Ec*YihS (PDB ID: 2RGK) were previously solved prior to determination of their physiological function.[20] We revisited the catalytic activity of *Se*YihS to demonstrate that it also catalyzed the interconversion of SQ, SF and SR (Figure S4). To elucidate what interactions are made between YihS and the sulfonate moiety of its sulfosugar substrates, the inactive mutant *Se*YihS-H248A was produced for structural studies. Size exclusion chromatography-multiangle light scattering (SEC-MALS) revealed that *Se*YihS exists as a hexamer in solution, although an asymmetric elution peak and marginal decrease in molecular weight may signify some dissociation under experimental conditions (Figure S5). This protein was co-crystallized with SF and X-ray diffraction techniques used to determine the structure of the *Se*YihS-H248A•SF product complex, which existed as a crystallographic dimer-of-trimers (Table S1, Figure S5), commensurate with the observation of hexamers in solution by SEC-MALS.

As previously observed, the overall fold of YihS comprises an (α_6/α_6)-barrel scaffold displaying close similarity to *N*-acyl-D-glucosamine epimerases **(Figure 2d)**.[20] Clear density for a furanose sulfo-sugar was seen in all six chains that allowed modeling of β-SF in a highly compact active site. Overlaying the *Se*YihS-H248A•SF structure with the ligand-free wild-type *Se*YihS structure (2AFA.pdb) gave an rmsd of 0.7 Å over 496 residues, with the catalytic residues His248 and His383 positioned on opposite sides of C1 hydroxyl group of bound SF and shows that the active site architecture remains largely unchanged (Figure S6). However, upon binding SF, two loops surrounding the active site undergo reorganization. The loop comprising residues 228-242 moves towards the ligand pocket such that the side-chain of Phe239 reorients and projects into the active site. A previously unstructured, flexible loop 370-377 now adopts an ordered conformation and interacts with residues from loop 228-242, such that Trp375 makes π-stacking interactions with Arg238. This cation-π interaction between the residues in two separate loops encloses the pocket to form the active site.

SF is buried in a deep pocket surrounded by bulky residues on either side (Phe239, Trp316 and Trp51) **(Figure 2e,f)**. Sugar hydroxyls make several hydrogen-bonding interactions with polar residues. The sulfonate is accommodated by a network of hydrogen-bonding interactions, with one of the sulfonate oxygens hydrogen-bonding with Gln379 (2.8 Å), a second oxygen with Gln362 (3.1 Å) and an ordered water molecule (3.1 Å) and the third oxygen forming a salt-bridge with Arg55 (2.8 Å), which also hydrogen bonds to the endocyclic oxygen of SF (3.1 Å). The observation of SF binding to YihS as the β-anomer does not

definitely show which anomer is produced by the enzyme. In order to understand the ease with which SF can anomerize, we measured its unidirectional rate of uncatalyzed anomerization using an NMR based chemical exchange spectroscopy method at equilibrium previously used to determine the mutarotation rate of SQ.[14] At pD 7.5 and 25 °C, β-SF mutarotates to α-SF with $k = 0.026$ s^{-1}, corresponding to $t_{1/2} = 26$ s (Figure S7) (the corresponding data for mutarotation of β-SQ to α-SQ are $k = 3.87 \times 10^{-5}$ s^{-1}, $t_{1/2} = 299$ min). The rate was unaffected by addition of a SQ-specific mutarotase from *Herbaspirillium seropedicae*. This rapid rate of mutarotation suggests facile interconversion of SF anomers *in cellulo*.

Sulforhamnose is a transcription factor derepressor

The observation that SR is a product and substrate of SQ isomerase raised questions as to its possible role in sulfoglycolysis: is it simply an unproductive intermediate that is ultimately isomerised to SF and consumed, or does it play a role in gene regulation, akin to the role that allolactose plays in regulating the *lac* operon?[21] The gene *csqR* within the SQ-utilising gene cluster of *E. coli* encodes a DeoR-family transcription factor (TF) that has been demonstrated to bind inside the intergenic spacer between the *yihUTS* operon and the *yihV* gene, and in doing so represses expression of sulfoglycolytic genes (**Figure 1b**).[11] SQ and sulfoquinovosyl glycerol (SQGro) are derepressors for CsqR binding to DNA, as demonstrated using gel-shift analyses with DNA sequences that encompass the CsqR-binding sites and a reporter assay.[11] We used this same gel-shift assay to assess if SR can influence the CsqR-DNA interaction. Purified CsqR was used to probe dsDNA encompassing the *yihUV* intergenic region using a polyacrylamide gel-shift assay. Disappearance of the protein resulted from high levels of cooperative binding of CsqR along the probe DNA that prevents the polymeric CsqR-DNA complexes from entering the gel during electrophoresis **(Figure 3a)**. Titration with increasing [SQ] leads to dissociation of CsqR from its DNA complexes, releasing free DNA that could be detected within the gel, in agreement with the previous report.[11] Titration with increasing [SR] also led to dissociation of CsqR-DNA complexes, however completely free DNA was not observed; rather, higher molecular weight complexes were formed that decreased in size at higher [SR], indicating that SR is a weaker mediator of the CsqR-DNA binding interaction.

To gain further insight into the effect of SQ, SQGro and SR on the transcription regulator CsqR, two constructs were designed based on the predicted domain boundaries using InterPro Classification (Figure S8a): full-length CsqR, which contains two distinct domains – the winged helix-turn-helix HTH DNA binding domain (DBD, residues 2-66) and the DeoR

effector binding domain (EBD, residues 80-260); and a truncated EBD-CsqR construct comprised of only the EBD. Binding studies using a temperature unfolding assay with purified EBD-CsqR and different sulfosugars as effector molecules gave improved stability only with SQ (ΔT_m = 4.6°C at 10 mM), with no change of T_m noted for SQGro or SR at 10 mM (Figure S8b).

To provide further insight into the effect of SQ, SQGro and SR on CsqR-DNA binding, we imaged CsqR-DNA interactions by AFM (atomic force microscopy). Purified full-length CsqR was mixed with the *yihUV* probe in the presence of increasing concentrations of SQ or SR, and subjected to AFM imaging under the same conditions as employed in a previous study (**Figure 3b**).[11] Titration with increasing concentrations of SQ caused a reduction in the size of the aggregates of CsqR-DNA complexes, at low concentrations of 0.5 and even 0.05 mM. By contrast, titration of CsqR-DNA aggregates with SR did not result in dissipation of aggregates at 0.05 mM, limited dissipation at 0.5 mM, and with complete dissociation observed only at 5.0 mM. Collectively, these data suggest that SR is a transcription inducer with lower potency than SQ.

SF kinase activity is regulated by sulfoglycolytic and central metabolites

YihV is a member of the Pfam pfkB carbohydrate kinase family,[22] and is an ATP-dependent kinase that mediates phosphoryl transfer from ATP to SF to give SFP. We established an HPLC-MS/MS assay for *Ec*YihV (Figure S9) using a ZIC-HILIC column and chemically synthesized SF substrate (Figure S10).[19] Initially, we measured the kinetics for production of SFP. Under conditions of constant SF (1.0 mM) and varying ATP in the presence of $MgCl_2$, *Ec*YihV exhibited Michaelis-Menten kinetics with k_{cat} 3.1 ± 0.2 s^{-1}, K_M = 1.0 ± 0.2 mM and k_{cat}/K_M = 3.2±0.8 mM s^{-1} (**Table 1**, Figure S11). Conversely, under conditions of constant ATP (1.0 mM) and varying SF, *Ec*YihV exhibited weak substrate inhibition, with an estimated K_I value for SF of 8 mM. The K_I value was sensitive to ATP concentration, and at [ATP] = 0.1 mM, dropped to 0.3 mM. As phosphofructokinase is a critical allosteric control point for glycolysis, we investigated whether *Ec*YihV is sensitive to a range of metabolites from sulfoglycolysis (SQ and SLA), glycolysis (F6P, FBP, DHAP and PEP) or central metabolism (citrate, representing the Krebs cycle) and ADP. This revealed that *Ec*YihV is strongly inhibited by ADP, and is activated by SQ, SLA and DHAP (through effects on K_M) and by F6P, FBP, PEP and citrate (through effects on k_{cat}/K_M) (**Figure 4a, Table 1**). Taken together, these data suggest that *Ec*YihV is an important control point for flux through the sulfo-EMP pathway.

We next examined if *Ec*YihV could catalyze phosphorylation of F6P. Incubation of *Ec*YihV with 1 mM F6P, ATP and $MgCl_2$ did not lead to the formation of FBP, showing that this enzyme is specific for SF, and indicating that no cross-talk exists between sulfoglycolysis and the EMP glycolysis or gluconeogenesis pathways.

To delineate the specific interactions with substrates and products, three crystal structures of *Ec*YihV were obtained in complex with substrate SF and AMPPNP or ADP as ATP analogues, as well as product SFP (Tables S1-2, Figures 4b-4f). *Ec*YihV forms a dimeric assembly and SEC-MALS analysis of *Ec*YihV revealed a mixture of solution states with the major peak corresponding to a dimer (>60%) and a minor (13%) sub-population corresponding to a tetrameric ensemble (Figure S9). The overall structure of YihV kinase displays a two-domain architecture comprising major α/β nucleotide binding domain and β-sheet serving as a 'lid' domain covering the active site. The orthogonal packing of eight β strands within the lid domains of the two subunits form a 'β-clasp' dimerization motif, a distinctive feature previously seen in members of the ribokinase superfamily that includes tagatose-6-phosphate kinase (TPK) and the minor isozyme from the glycolytic pathway, F6P kinase (PfkB).[23-25] Close hydrophobic contacts and reciprocal interactions are observed at the dimer interface where residues from the opposite subunit in a β-clasp motif protrude into the active site, which is likely to be crucial for both catalysis and structural stability.

Co-crystallization with the non-hydrolysable ATP analogue AMPPNP resulted in the *Ec*YihV•AMPPNP•Mg complex showing a surface-accessible nucleotide binding site **(Figure 4b,c)**. Density for AMPPNP bound in an *anti*-conformation with hydrated Mg ions was present in all four subunits in the structure. YihV subunits were observed in 'open' and 'closed' conformations due to inter-domain rotation. In the open conformation, the γ-phosphate of AMPPNP is hydrogen-bonded to backbone amide of Gly243 (2.7 Å) and a water molecule (2.5 Å) and is present within 5 Å distance of conserved catalytic Asp244 (Figure S13), whereas in the closed conformation the γ-phosphate moves closer to the substrate cleft and achieves an appropriate distance for phosphoryl transfer. This closed conformation of YihV most likely denotes an inactive conformation that would prevent binding of substrate. Analysis using the DynDom program showed the β-clasp domain rotates 31 degrees about four hinge bending regions, reflective of dynamic domain movements associated with binding of substrates (Figure S14). Considering the regulatory role of this enzyme in the pathway, these dynamic movements may be promoted by binding of ligands at an allosteric site(s), thereby driving transitions between productive and unproductive conformations.

In the *Ec*YihV•SFP binary complex, the two subunits in a dimer pair are both in a closed conformation (Figure S12). SFP bound at the cleft between the nucleotide domain and β-barrel motif, showing that binding of the sulfonate ligand induces inter-domain rotation and subsequent closure. The *Ec*YihV•SFP complex revealed a sulfonate sub-pocket where one of the sulfonate oxygens of SFP is hydrogen bonded to Nε of Arg138 (2.8 Å) and Asn109 (3.2 Å), the second sulfonate oxygen forms a salt bridge with the in-trans Lys27 (2.9 Å), which projects into the active site from the β3 strand of other subunit in the dimer pair and makes reciprocal interactions with Asp162 present at 3 Å **(Figure 4d)**. Lys27 also forms hydrogen bonds with the C1 hydroxyl and ring oxygen of SFP, each present at 2.9 Å.

A compact dimer pair is seen for a dead-end complex of *Ec*YihV with SF and ADP•Mg, formed by closure of the active site **(Figure 4e,f)**. In this complex, the sugar is deeply sequestered at the edge of the inter-domain cleft and ADP molecule (with a penta-hydrated Mg ion coordinated to α-phosphate) is bound close to the substrate phosphorylation site. The C1 hydroxyl of SF is hydrogen bonded to catalytic base Asp244 (2.7 Å) and Lys27 (3.0 Å), with the substrate positioned for phosphoryl transfer. The remaining sugar hydroxyls make several hydrogen bonding interactions with other active site residues suggesting SF binding facilitates domain closure. Specifically, the C2 hydroxyl is H-bonded to Asp244 (2.6 Å), Asp13 H-bonds C3 and C4 hydroxyls each at 2.5 Å, and the C4 hydroxyl makes an additional H-bond to Ser95 (2.7 Å). The hydrogen bonding network at the sulfonate binding site is similar to the *Ec*YihV•SFP structure above, wherein two of the sulfonate oxygens interacting with Nε of Arg138 (2.8 Å) and Asn109 (3.2 Å) and the in-trans Lys27 (2.7 Å). The third sulfonate oxygen makes hydrogen bonds with a bound water molecule that in turn interacts with backbone carbonyl of Tyr28 and another water molecule.

An ordered sequential Bi-Bi mechanism is reported for ribokinases, with sugar binding interactions driving domain closure that precedes ATP binding.[23-25] Nano differential scanning fluorimetry (nanoDSF) studies using *Ec*YihV in presence of SF and ADP resulted in a T_m shift of 6 °C concurrent with binding of the sugar (and only a 1 °C shift seen upon binding ADP) indicating ligand-induced conformational stabilization of the kinase (Figure S15). This data complements the structural details revealed in the complexes reported here, showing that domain movements upon SF binding result in a closed, functional state, and ATP can access the active site through a surface groove (Figure S14-15). This may avoid premature hydrolysis of the γ-phosphate of ATP, as reported in ribokinases.[25-27]

Given the close functional relationship- of SF kinase (YihV) and PfkB and TPK, sequence- and structure-based analysis of their binding sites was undertaken. Two conserved motifs have been described in these phosphosugar kinases, the **TR** and **GXGDXX** motifs, responsible for substrate binding and catalysis, respectively.[28,29] Sequence alignment of PfkB from *E. coli* and TPK from *S. aureus* with SF kinase showed that **GXGDXX** motif comprising the catalytic aspartate D244 (*Ec*YihV numbering) is conserved in all three sub-families (Figure S16). Comparison of the closed conformation of *Ec*YihV with TPK and PfkB revealed that overall fold is highly conserved in these sugar kinases: with PfkB (3N1C.pdb), rmsd 2.6 over 248 residues; with TPK (2JG1.pdb), rmsd 2.2 over 265 residues (Figure S17). Finally, close inspection of the 6-phospho- versus 6-sulfonate-binding sites highlighted key differences consistent with the chemical structures of bound sugars (Figure S18). In PfkB and TPK, Arg88 present within the substrate-binding **TR motif** is part of the conserved RRS triad for recognition of 6-phosphate moiety. The substrate-specific **TR motif** and the consensus RRS triad are absent in SF kinases. Instead a **KRN** sulfonate recognition triad (K27-R138-N109) was identified in the SFP and SF complexes, which is conserved in annotated YihV kinases (Figures S16, S18).

SFP aldolase

SFP aldolase catalyses reversible retro-aldol cleavage of SFP to DHAP and SLA. Schleheck and co-workers detected the formation of DHAP and SLA in a coupled system containing SQ and recombinantly produced YihS, YihV and YihT, supplemented with ATP and $MgCl_2$.[7] Here, we established an LCMS-MS assay for direct analysis of the reaction catalyzed by YihT. Recombinantly expressed YihT from *E. coli* was incubated with chemo-enzymatically synthesized SFP;[19] however, we could not detect the formation of DHAP or SLA under a range of conditions. Instead, we recombinantly expressed YihT from *Salmonella enterica*, which shares 87% similarity) (Figure S19). Incubation of recombinant *Se*YihT with SFP revealed the formation of SLA and DHAP. The enzyme exhibited Michaelis-Menten kinetics with k_{cat} = 47.7 ± 2.4 s^{-1}, K_M = 3.57 ± 0.42 mM and k_{cat}/K_M = 13 ± 2 mM s^{-1} **(Figure 5a)**. Under similar conditions, no activity was noted on fructose bisphosphate (1 mM).

Benson reported that SFP can be generated from SLA and DHAP by an unspecified aldolase.[30] We revisited this observation by incubating recombinantly expressed *Se*YihT with chemically synthesized SLA and commercial DHAP. A product with identical retention time

and fragmentation pattern to authentic SFP was obtained, demonstrating the reversibility of the YihT-catalyzed reaction **(Figure 5b)**.

Fructose-bisphosphate aldolases are classified into Class I and II enzymes based on their mechanism of action. Class I aldolases utilize an active-site lysine to form Schiff base with the substrate, whereas prokaryotic and fungal Class II FBP aldolases are dependent on divalent metal ions for their activity. Class I aldolases are commonly found in algae, protozoa, plants and animals and adopt homotetrameric active forms; archaeal and bacterial enzymes utilize similar mechanisms but have low sequence similarity and are distinguished from their eukaryotic counterparts through self-association as multimers exhibiting tetrameric to decameric quaternary structures, leading to their sub-grouping into Class Ia.[31,32] In order to reveal structural features of SFP aldolase, the X-ray crystal structures of YihT from *E. coli* and *Salmonella enterica* were obtained with two and twelve molecules in an asymmetric unit (Figure S20), respectively, however SEC-MALS showed YihT exists as a homotetramer in solution (Figure S19). YihT displays an overall $(\alpha/\beta)_8$-barrel architecture similar to class I fructose bisphosphate aldolases **(Figure 5c**, Table S2**)**.[33] A DALI search using YihT against the RCSB PDB library gave annotated bacterial class I aldolases as closest structural neighbors. These included tagatose-1,6-bisphosphate aldolases from *Streptococcus mutans* (PDB ID: 3IV3 with DALI z score of 31.8, rmsd 2.4 and 21% sequence ID) and from *Streptococcus porcinus* (PDB ID: 5HJL with DALI z score of 31.7, rmsd 2.2 and 21% sequence ID), as well as fructose-bisphosphate aldolase from *Slackiahelio trinireducens* (PDB ID: 4MOZ with DALI z score of 22.9, rmsd 2.5 and 15% sequence ID), indicating high structural similarity despite low sequence ID.

In the *Se*YihT structure, density for a single sulfate ion in the proposed substrate binding pocket was observed in all twelve chains at occupancies 0.8-1 (Figure S21). The sulfate ion is located close to conserved Lys193, which is proposed to be involved in Schiff base formation. Substrate-soaking experiments at saturating concentrations of SFP resulted in trapping of the Schiff base adduct of *Se*YihT (**Figure 5c-d**, Table S3). Clear, contiguous density (at an occupancy of 0.8) was observed for a hexose covalently attached to Lys193 within two of the subunits of the tetrameric protein structure representing SFP in an open-chain conformation. In the other two subunits, there was evidence of in-crystal cleavage of the C3-C4 bond of the substrate and a new Schiff base formed with DHAP molecule could be modeled, consistent with the reaction catalyzed by this enzyme (Figure S22). The SFP fragment bound to Lys193 allowed unambiguous identification of both phosphate and sulfonate pockets within

*Se*YihT. The phosphate oxygens interact with Ser226, Ser227 and Arg253 as described in the *Se*YihT•SO_4^{2-} complex above. One sulfonate oxygen forms an electrostatic interaction with Arg253 (2.7 Å) and the other two sulfonate oxygens make hydrogen bonding interactions with two bound water molecules present at a distance of 2.7 and 2.8 Å (Figure 5e). These residues are conserved across annotated SFP aldolases and are proposed to comprise a sulfonate binding pocket that defines this subclass of enzymes (Figure S23.

Discussion

The sulfo-EMP pathway allows metabolism of the widespread sulfosugar SQ and its glycosides. The pathway shares a similar series of steps to the EMP pathway, with 6-sulfonate taking the place of phosphate. This similarity raises obvious questions about whether the sulfo-EMP pathway can act on intermediates in the EMP pathway or if the two pathways are functionally segregated. Kinetic analysis reveals that SQ-SF isomerase has no detectable activity on G6P, SF kinase has no activity on F6P, and SFP aldolase no activity on fructose bisphosphate, under conditions where robust enzymatic activity on their namesake substrates can be detected. This selectivity is likely to be of significance under conditions where SQ is the sole substrate for growth, in which case gluconeogenesis is required to supply F6P for the PPP and cell wall biosynthesis, and where reversal of flux is required. Among these steps, high selectivity for SF kinase to act solely on SF and not G6P is likely to be the most critical, as gluconeogenesis requires a switch from F6P kinase to FBP phosphatase to change directionality from catabolism to anabolism.

Using X-ray crystallography, we obtained 3D structures of all three enzymes and through complexes with SF or SFP defined sulfonate-binding pockets that show how these enzymes recognize their sulfosugar substrate. In all three enzymes the sulfonate pocket was lined with positively charged residues (for SQ isomerase: Arg55-Gln379-Gln383; for SF kinase: Lys27-Asn109-Arg138; for SFP aldolase: Arg253-Ala26(H_2O)), which in every case include an Arg residue to balance charge. This data complements X-ray data for 3D structures of sulfoquinovosidase and SLA reductase, which also exhibit selectivity for their sulfonate substrates and possessed well-defined and conserved sulfonate binding pockets. While the *Ec*SQase binding pocket is comprised of Arg301-Trp304-Tyr508(H_2O),[13] the SLA reductase binding pocket lacks a direct interaction with an Arg, rather it is comprised of Asn174-Ser178, and the backbone amide of Arg123.[15]

Transcription factors (TFs) involved in transcription regulation of the genes for metabolism are often controlled through allosteric interactions with metabolites. Previously, regulation of CsqR activity by both SQGro and SQ (sulfoquinovose) was identified.[11] Here we demonstrate that SR is also a regulator of CsqR-DNA binding, inducing the dissociation of repressor CsqR from its regulatory target DNA. The level of CsqR inactivation will therefore potentially be controlled depending on the level of three metabolites, SQGro, SQ and SR, with the induction level of sulfo-EMP operon depending on the affinity of each inducer ligand to CsqR. In *E. coli* K-12, regulation of TF activity by multiple metabolites has been found for TFs such as AllR (allantoin repressor) by allantoin and glyoxylate,[34] CysB (cysteine B) by *O*-acetyl-L-serine and thiosulfate,[35] GcvA (glycine cleavage A) by glycine and purine,[36] PurR (purine repressor) by hypoxanthine and guanine,[37] and RutR (pyrimidine utilization regulator) by uracil and thymine.[38] As in the case of CsqR, the activity and target selectivity of TyrR (tyrosine repressor) is controlled by three metabolites, phenylalanine, tryptophan and tyrosine,[39] with the level of TF activity dependent on the intracellular concentrations of all three metabolites.

The role of SF kinase in sulfoglycolysis is analogous to that of PFK in glycolysis. PFK is widely recognized as a central regulatory step in carbohydrate metabolism in most organisms and has complex allosteric properties. These effects help regulate distribution of G6P into the PPP pathway, cell wall synthesis, the production of carbohydrate storage molecules such as glycogen, maltose and trehalose, the channeling of PEP into the citric acid cycle and fatty acid synthesis, as well as ensuring ATP production is managed when in surplus. We showed modulatory effects on SF kinase by a range of central metabolites, including substrate inhibition by SF, strong inhibition by ADP, and activation by F6P, FBP, PEP and citrate. Substrate inhibition is widely recognized as a control strategy to maintain steady pathway flux even in the presence of large fluctuations in substrate concentration,[40] which may help to limit variation in flux through sulfoglycolysis as environmental concentrations of SQ and SQGro fluctuate. Because the wiring of cellular metabolism and pathway yields under conditions of sulfoglycolysis differ from glycolysis, significant differences in regulation are to be expected. Notably, while SF kinase is inhibited by ADP and not ATP, *E. coli* PFK-1 and -2 are inhibited by ATP, while only PFK-1 is activated by ADP and inhibited by ATP.[41,42] While sulfoglycolysis and glycolysis both yield PEP, unlike glycolysis, sulfoglycolysis does not produce G6P, thus there may not be the same requirement for ATP/ADP control over the pathway to control distribution of substrate into catabolic and anabolic pathways, as for G6P

in glycolysis. Instead, an important branchpoint post-sulfoglycolysis is likely to be distribution of PEP into gluconeogenesis and central metabolism (citric acid cycle and fatty acid synthesis). Activation by citrate, PEP, FBP and F6P may assist in maintaining flux into these anabolic and catabolic pathways. Additionally, it may be beneficial for glycolytic-EMP and sulfo-EMP pathways to be regulated in distinct manners, which could allow them to operate in parallel if SQ is available together with glucose, providing greater metabolic flexibility, and allowing the regulation of cellular energy metabolism as the ATP yield of the two pathways differ.

Conclusions

The present work provides the first detailed biochemical and structural analysis for the three core enzymes of the *E. coli* sulfo-EMP pathway: SQ isomerase, SF kinase, and SFP aldolase. Collectively, these data demonstrate kinetic selectivity for the core sulfoglycolytic enzymes for sulfoglycolytic intermediates over the corresponding intermediates in glycolysis. This selectively arises from conserved, sulfonate binding pockets in each enzyme. The effect of this selectivity is to allow functional segregation of this pathway from glycolysis/gluconeogenesis, preventing futile cycling, and highlights the specific structural adaptations that have led to the evolution of this pathway. The present work provides a roadmap to conduct more detailed bioinformatic analyses of this pathway in complex microbial communities and to understand how sulfoglycolysis contributes to bacterial metabolism.

Associated content

Supporting information

The supporting information is available free of charge at http//doi.org…

Additional figures (Figures S1-S23) and tables (Tables S1-3) including X-ray data, structures, enzyme kinetics, nmr analysis of products, and protein purification and characterization. Methods describing protein expression, purification, biophysical characterization and crystallography; enzymatic reactions, gel shift and atomic force microscopy, and methods for chemical synthesis of sulforhamnose (and associated NMR data).

Accession codes

The coordinate files and structure factors have been deposited in the Protein DataBank (PDB) with accession numbers **7AG4**, **7AGH**, **7AG6**, **7AGK**, **7AG1**, **7AG7** and **7NE2**.

Author Information

Corresponding authors

E-mail: sjwill@unimelb.edu.au

gideon.davies@york.ac.uk

goddard-borger.e@wehi.edu.au

Author contributions

G.J.D, A.I., and S.J.W conceived the project. G.J.D, E.D.G.-B., A.I. and S.J.W designed the experiments. M.S., Y.J. and J.L. performed molecular biology and protein chemistry. P.A., R.E. and S.J.W. conducted enzyme kinetics and analyzed results. P.A., J.M. and R.E. synthesized substrates/reagents. M.S. and J.Y. and G.J.D. conducted X-ray crystallography and analyzed results. T.S., M.S. and A.I. conducted CsqR experiments and analyzed results. M.S., S.J.W., G.J.D. wrote the manuscript with input from other authors.

Notes

The authors declare no competing financial interest.

Acknowledgements

This work was supported by the Australian Research Council (DP180101957), the Leverhulme Trust (RPG-2017-190), MEXT Cooperative Research Program of Network Joint Research Center for Materials and Devices (AI and TS), the National Health and Medical Research Council of Australia (GNT1139546 and GNT1139549), and the Royal Society for the Ken Murray Research Professorship to GJD. We acknowledge support from: The Walter and Eliza Hall Institute of Medical Research; the Australian Cancer Research Fund; and a Victorian State Government Operational Infrastructure support grant. We acknowledge the staff of the Diamond Light Source (UK) for provision of I03, IO4 and IO4-1 beamline facilities (proposal

numbers mx-13587, mx-18598 and mx-24948). We thank Noriyuki Kodera (Kanazawa Univ) for fabricating the electron beam deposited tips of AFM, Christopher Bengt, Joe Joiner, and Prof David Vocadlo for constructive advice.

References

1. Canfield, D., and Farquhar, J. (2012) The global sulfur cycle. in *Fundamentals of Geobiology* (Knoll, A. H., Canfield, D. E., and Konhauser, K. O. eds.), Blackwell Publishing Ltd. pp 49-64
2. Dickschat, J. S., Rabe, P., and Citron, C. A. (2015) The chemical biology of dimethylsulfoniopropionate. *Org. Biomol. Chem.* **13**, 1954-1968
3. Thume, K., Gebser, B., Chen, L., Meyer, N., Kieber, D. J., and Pohnert, G. (2018) The metabolite dimethylsulfoxonium propionate extends the marine organosulfur cycle. *Nature* **563**, 412-415
4. Goddard-Borger, E. D., and Williams, S. J. (2017) Sulfoquinovose in the biosphere: occurrence, metabolism and functions. *Biochem. J.* **474**, 827–849
5. Harwood, J. L., and Nicholls, R. G. (1979) The plant sulpholipid - a major component of the sulphur cycle. *Biochem. Soc. Trans.* **7**, 440-447
6. Frommeyer, B., Fiedler, A. W., Oehler, S. R., Hanson, B. T., Loy, A., Franchini, P., Spiteller, D., and Schleheck, D. (2020) Environmental and Intestinal Phylum Firmicutes Bacteria Metabolize the Plant Sugar Sulfoquinovose via a 6-Deoxy-6-sulfofructose Transaldolase Pathway. *iScience* **23**, 101510
7. Denger, K., Weiss, M., Felux, A. K., Schneider, A., Mayer, C., Spiteller, D., Huhn, T., Cook, A. M., and Schleheck, D. (2014) Sulphoglycolysis in *Escherichia coli* K-12 closes a gap in the biogeochemical sulphur cycle. *Nature* **507**, 114-117
8. Felux, A. K., Spiteller, D., Klebensberger, J., and Schleheck, D. (2015) Entner-Doudoroff pathway for sulfoquinovose degradation in *Pseudomonas putida* SQ1. *Proc. Natl. Acad. Sci. USA* **112**, E4298-4305
9. Buck, H., Kerim, K., Jessica, L., Anna, B., Alexander, F., Karin, D., Benjamin, F., Craig, H., Thomas, R., Nicolai, K., Nicola, S., David, S., and Alexander, L. (2021) Sulfoquinovose is a select nutrient of prominent bacteria and a source of hydrogen sulfide in the human gut. *Research Square*, 10.21203/rs.21203.rs-49676/v21201
10. Li, J., Epa, R., Scott, N. E., Skoneczny, D., Sharma, M., Snow, A. J. D., Lingford, J. P., Goddard-Borger, E. D., Davies, G. J., McConville, M. J., and Williams, S. J. (2020) A Sulfoglycolytic Entner-Doudoroff Pathway in *Rhizobium leguminosarum* bv. *trifolii* SRDI565. *Appl. Environ. Microbiol.* **86**, e00750-00720
11. Shimada, T., Yamamoto, K., Nakano, M., Watanabe, H., Schleheck, D., and Ishihama, A. (2019) Regulatory role of CsqR (YihW) in transcription of the genes for catabolism of the anionic sugar sulfoquinovose (SQ) in Escherichia coli K-12. *Microbiology* **165**, 78-89
12. Abayakoon, P., Jin, Y., Lingford, J. P., Petricevic, M., John, A., Ryan, E., Wai-Ying Mui, J., Pires, D. E. V., Ascher, D. B., Davies, G. J., Goddard-Borger, E. D., and Williams, S. J. (2018) Structural and Biochemical Insights into the Function and Evolution of Sulfoquinovosidases. *ACS Cent. Sci.* **4**, 1266-1273
13. Speciale, G., Jin, Y., Davies, G. J., Williams, S. J., and Goddard-Borger, E. D. (2016) YihQ is a sulfoquinovosidase that cleaves sulfoquinovosyl diacylglyceride sulfolipids. *Nat. Chem. Biol.* **12**, 215-217
14. Abayakoon, P., Lingford, J. P., Jin, Y., Bengt, C., Davies, G. J., Yao, S., Goddard-Borger, E. D., and Williams, S. J. (2018) Discovery and characterization of a

sulfoquinovose mutarotase using kinetic analysis at equilibrium by exchange spectroscopy. *Biochem. J.* **475**, 1371-1383
15. Sharma, M., Abayakoon, P., Lingford, J. P., Epa, R., John, A., Jin, Y., Goddard-Borger, E. D., Davies, G. J., and Williams, S. J. (2020) Dynamic Structural Changes Accompany the Production of Dihydroxypropanesulfonate by Sulfolactaldehyde Reductase. *ACS Catalysis* **10**, 2826-2836
16. Dunaway, G. A. (1983) A review of animal phosphofructokinase isozymes with an emphasis on their physiological role. *Mol. Cell. Biochem.* **52**, 75-91
17. Hers, H. G., and Hue, L. (1983) Gluconeogensis and related aspects of glycolysis. *Annu. Rev. Biochem.* **52**, 617-653
18. Finn, R. D., Mistry, J., Schuster-Böckler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S. R., Sonnhammer, E. L. L., and Bateman, A. (2006) Pfam: clans, web tools and services. *Nucleic Acids Res.* **34**, D247-D251
19. 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., and Williams, S. J. (2019) Comprehensive synthesis of substrates, intermediates and products of the sulfoglycolytic Embden-Meyerhoff-Parnas pathway. *J. Org. Chem.* **84**, 2910-2910
20. Itoh, T., Mikami, B., Hashimoto, W., and Murata, K. (2008) Crystal structure of YihS in complex with D-mannose: structural annotation of *Escherichia coli* and *Salmonella enterica* yihS-encoded proteins to an aldose-ketose isomerase. *J. Mol. Biol.* **377**, 1443-1459
21. Jobe, A., and Bourgeois, S. (1972) lac Repressor-operator interaction. VI. The natural inducer of the lac operon. *J. Mol. Biol.* **69**, 397-408
22. Bork, P., Sander, C., and Valencia, A. (1993) Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases. *Protein Sci.* **2**, 31-40
23. Sigrell, J. A., Cameron, A. D., and Mowbray, S. L. (1999) Induced fit on sugar binding activates ribokinase. *J. Mol. Biol.* **290**, 1009-1018
24. Sigrell, J. A., Cameron, A. D., Jones, T. A., and Mowbray, S. L. (1998) Structure of *Escherichia coli* ribokinase in complex with ribose and dinucleotide determined to 1.8 A resolution: insights into a new family of kinase structures. *Structure* **6**, 183-193
25. Park, J., and Gupta, R. S. (2008) Adenosine kinase and ribokinase-the RK family of proteins. *Cell. Mol. Life Sci.* **65**, 2875-2896
26. Murillo-López, J., Zinovjev, K., Pereira, H., Caniuguir, A., Garratt, R., Babul, J., Recabarren, R., Alzate-Morales, J., Caballero, J., Tuñón, I., and Cabrera, R. (2019) Studying the phosphoryl transfer mechanism of the *E. coli* phosphofructokinase-2: from X-ray structure to quantum mechanics/molecular mechanics simulations. *Chem. Sci.* **10**, 2882-2892
27. Roy, S. V. V., M.; Harmer, N.J. (2019) Carbohydrate Kinases: A Conserved Mechanism Across Differing Folds. *Catalysts* **9**, 29
28. Cabrera, R., Babul, J., and Guixé, V. (2010) Ribokinase family evolution and the role of conserved residues at the active site of the PfkB subfamily representative, Pfk-2 from *Escherichia coli*. *Arch. Biochem. Biophys.* **502**, 23-30
29. Miallau, L., Hunter, W. N., McSweeney, S. M., and Leonard, G. A. (2007) Structures of *Staphylococcus aureus* D-tagatose-6-phosphate kinase implicate domain motions in specificity and mechanism. *J. Biol. Chem.* **282**, 19948-19957
30. Benson, A. A., and Shibuya, I. (1961) Sulfocarbohydrate metabolism. *Fed. Proc.* **20**, 79

31. Lorentzen, E., Siebers, B., Hensel, R., and Pohl, E. (2005) Mechanism of the Schiff Base Forming Fructose-1,6-bisphosphate Aldolase: Structural Analysis of Reaction Intermediates. *Biochemistry* **44**, 4222-4229
32. Siebers, B., Brinkmann, H., Dörr, C., Tjaden, B., Lilie, H., van der Oost, J., and Verhees, C. H. (2001) Archaeal fructose-1,6-bisphosphate aldolases constitute a new family of archaeal type class I aldolase. *J. Biol. Chem.* **276**, 28710-28718
33. Marsh, J. J., and Lebherz, H. G. (1992) Fructose-bisphosphate aldolases: an evolutionary history. *Trends Biochem. Sci.* **17**, 110-113
34. Hasegawa, A., Ogasawara, H., Kori, A., Teramoto, J., and Ishihama, A. (2008) The transcription regulator AllR senses both allantoin and glyoxylate and controls a set of genes for degradation and reutilization of purines. *Microbiology* **154**, 3366-3378
35. Kredich, N. M. (1992) The molecular basis for positive regulation of cys promoters in *Salmonella typhimurium* and *Escherichia coli*. *Mol. Microbiol.* **6**, 2747-2753
36. Jourdan, A. D., and Stauffer, G. V. (1998) Mutational Analysis of the Transcriptional Regulator GcvA: Amino Acids Important for Activation, Repression, and DNA Binding. *J. Bacteriol.* **180**, 4865
37. Meng, L. M., and Nygaard, P. (1990) Identification of hypoxanthine and guanine as the co-repressors for the purine regulon genes of *Escherichia coli*. *Mol. Microbiol.* **4**, 2187-2192
38. Shimada, T., Hirao, K., Kori, A., Yamamoto, K., and Ishihama, A. (2007) RutR is the uracil/thymine-sensing master regulator of a set of genes for synthesis and degradation of pyrimidines. *Mol. Microbiol.* **66**, 744-757
39. Pittard, A. J., and Davidson, B. E. (1991) TyrR protein of *Escherichia coli* and its role as repressor and activator. *Mol. Microbiol.* **5**, 1585-1592
40. Reed, M. C., Lieb, A., and Nijhout, H. F. (2010) The biological significance of substrate inhibition: a mechanism with diverse functions. *Bioessays* **32**, 422-429
41. Guixé, V., and Babul, J. (1985) Effect of ATP on phosphofructokinase-2 from *Escherichia coli*. A mutant enzyme altered in the allosteric site for MgATP. *J. Biol. Chem.* **260**, 11001-11005
42. Zheng, R. L., and Kemp, R. G. (1992) The mechanism of ATP inhibition of wild type and mutant phosphofructo-1-kinase from *Escherichia coli*. *J. Biol. Chem.* **267**, 23640-23645

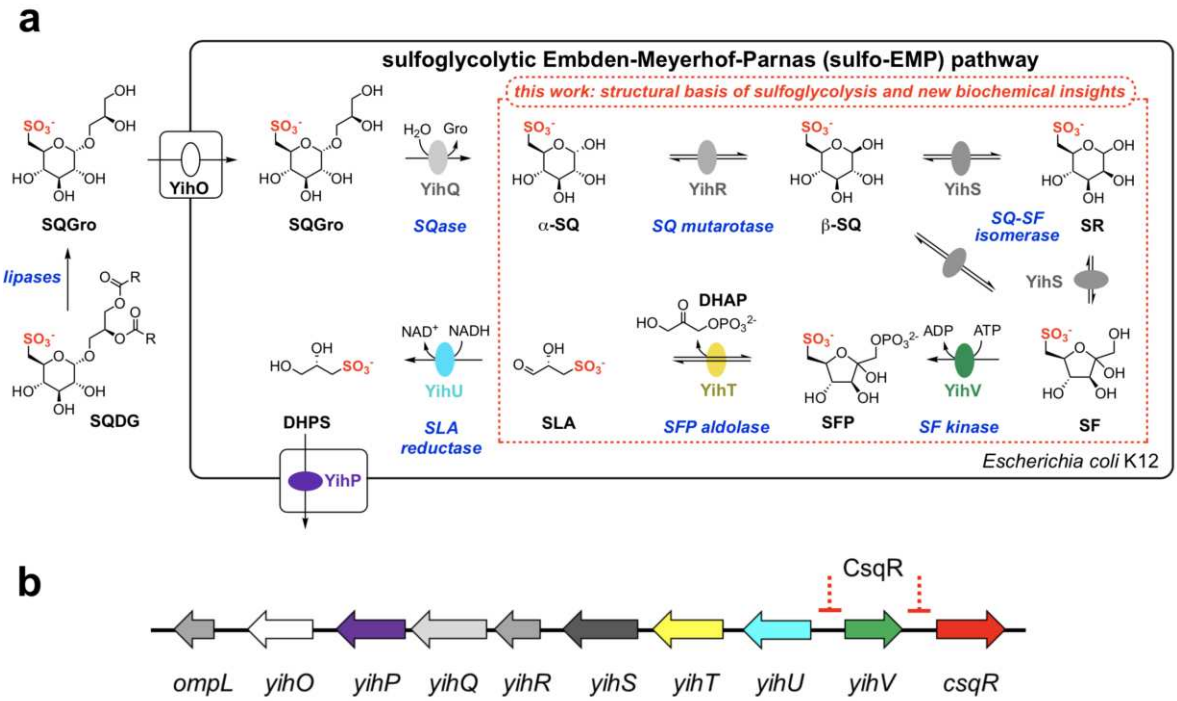


Figure 1. The Embden-Meyerhof-Parnas sulfoglycolysis pathway in *E. coli*. **(a)** Proposed biochemical pathway for metabolism of SQGro. Lipase cleavage of acyl groups in SQDG occurs in the environment. **(b)** Sulfo-EMP operon showing binding sites for the CsqR transcriptional repressor. SQDG, sulfoquinovosyl diacylglycerol; SQGro, α-sulfoquinovosyl glycerol; SQ, sulfoquinovose; SF, 6-deoxy-6-sulfofructose; SR, 6-deoxy-6-sulforhamnose; SFP, 6-deoxy-6-sulfofructose-1-phosphate; SLA, sulfolactaldehyde; DHPS, 2,3-dihydroxypropanesulfonate; DHAP, dihydroxyacetone phosphate; NADH, nicotinamide adenine dinucleotide, reduced form. Formation of SQGro is achieved by lipase action on sulfoquinovosyl diacylglycerol.

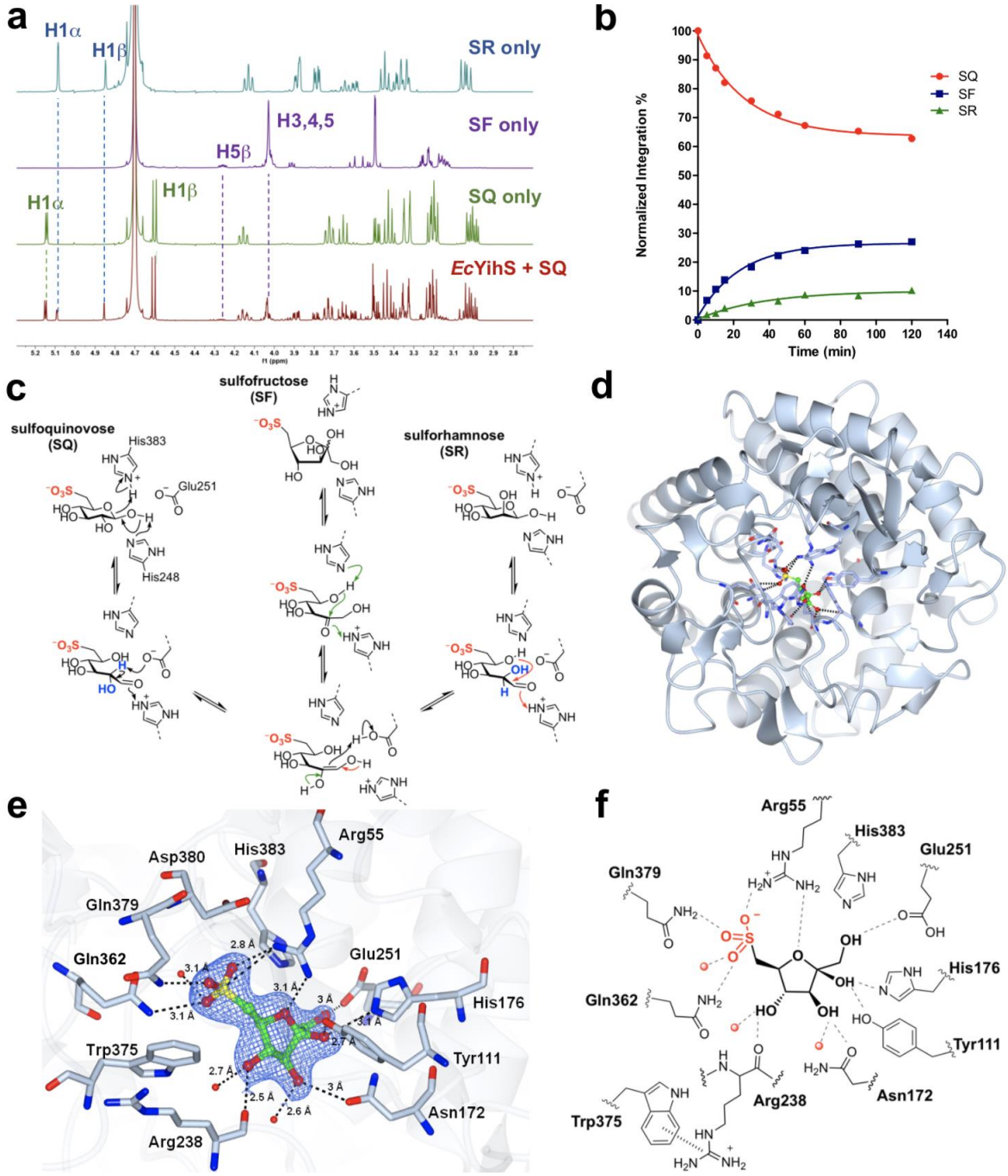


Figure 2. Characterization and crystal structure of SQ isomerase YihS. (a) ^{1}H NMR spectrum (in D_2O) showing the equilibrium mixture resulting from isomerization of SQ by *Ec*YihS, and reference spectra for pure SQ, SF and SR. **(b)** Time course plot for *Ec*YihS catalysed equilibration of SQ to SF and SR. Substrate depletion plot obtained by HPLC was normalized to final composition determined by ^{1}H NMR analysis (SF, 30.2; SR, 21.3; SQ, 48.5%). **(c)** Proposed mechanism of *Ec*YihS catalyzed isomerization of SQ to produce SF and SR. **(d)** Overall fold of *Se*YihS monomer showing location of active site. **(e)** Crystal structure

of *Se*YihS•SF showing the sulfonate pocket. Ribbon diagram of protein backbone is depicted in blue, and SF and side chains of active site residues are shown in cylinder format. Electron density corresponds to the 2Fo − Fc map (in blue) at levels of 1.5σ. **(f)** Cartoon of ligand binding pocket of *Se*YihS•SF complex depicting hydrogen bonding interactions with active site residues.

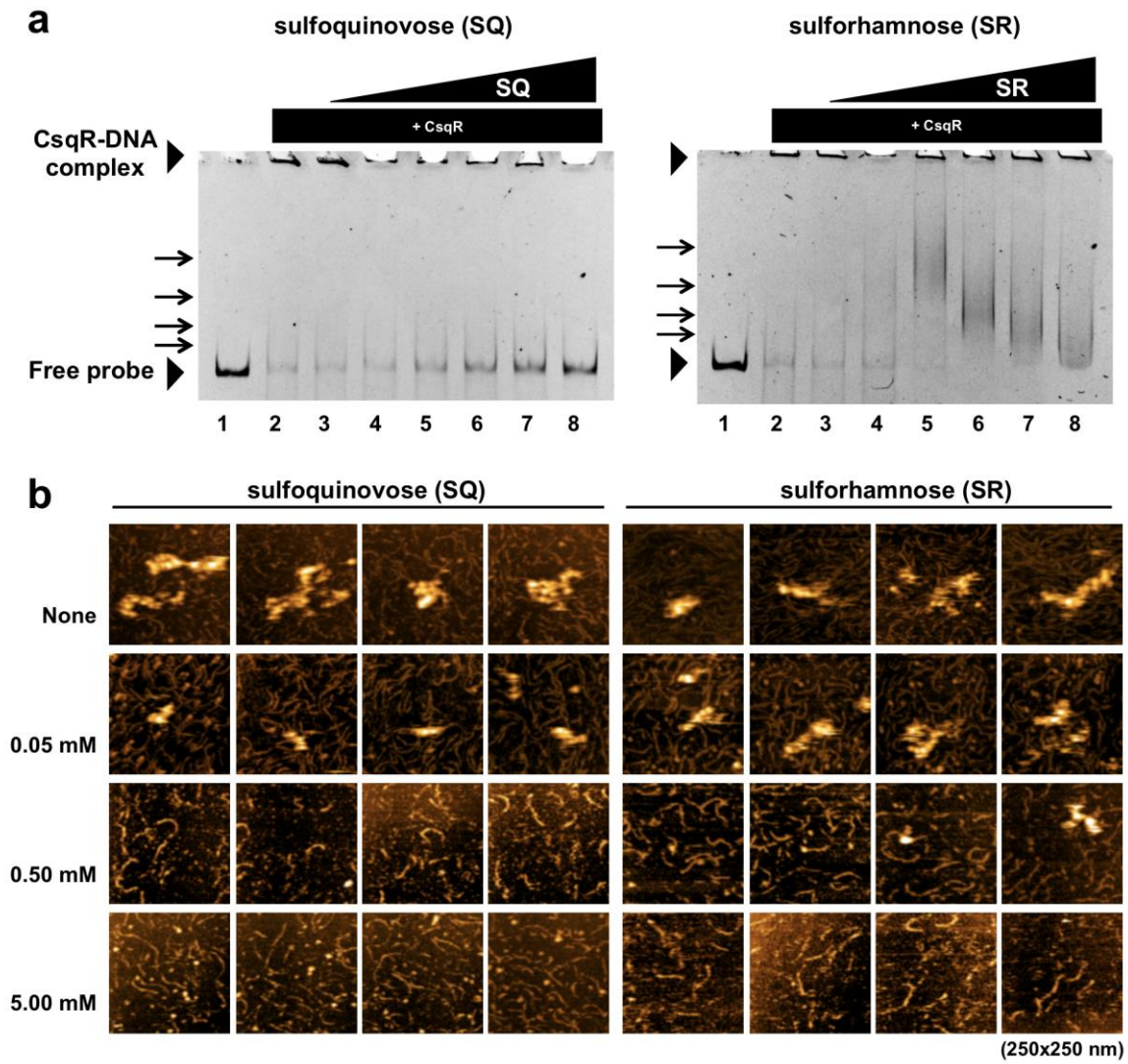


Figure 3. Sulforhamnose is a derepressor of the transcriptional repressor CsqR. (**a**) PAGE gel shift assay showing effect of SQ and SR on binding of CsqR to a DNA probe of the *yihUV* intergenic region. Effector concentrations, lanes 1-8: 0, 0, 5, 10, 25, 50, 75, 100 mM. (**b**) Representative AFM observations of the effect of SQ and SR on binding of CsqR to *yihUV* intergenic DNA probe. All images cover an area of 250×250 nm.

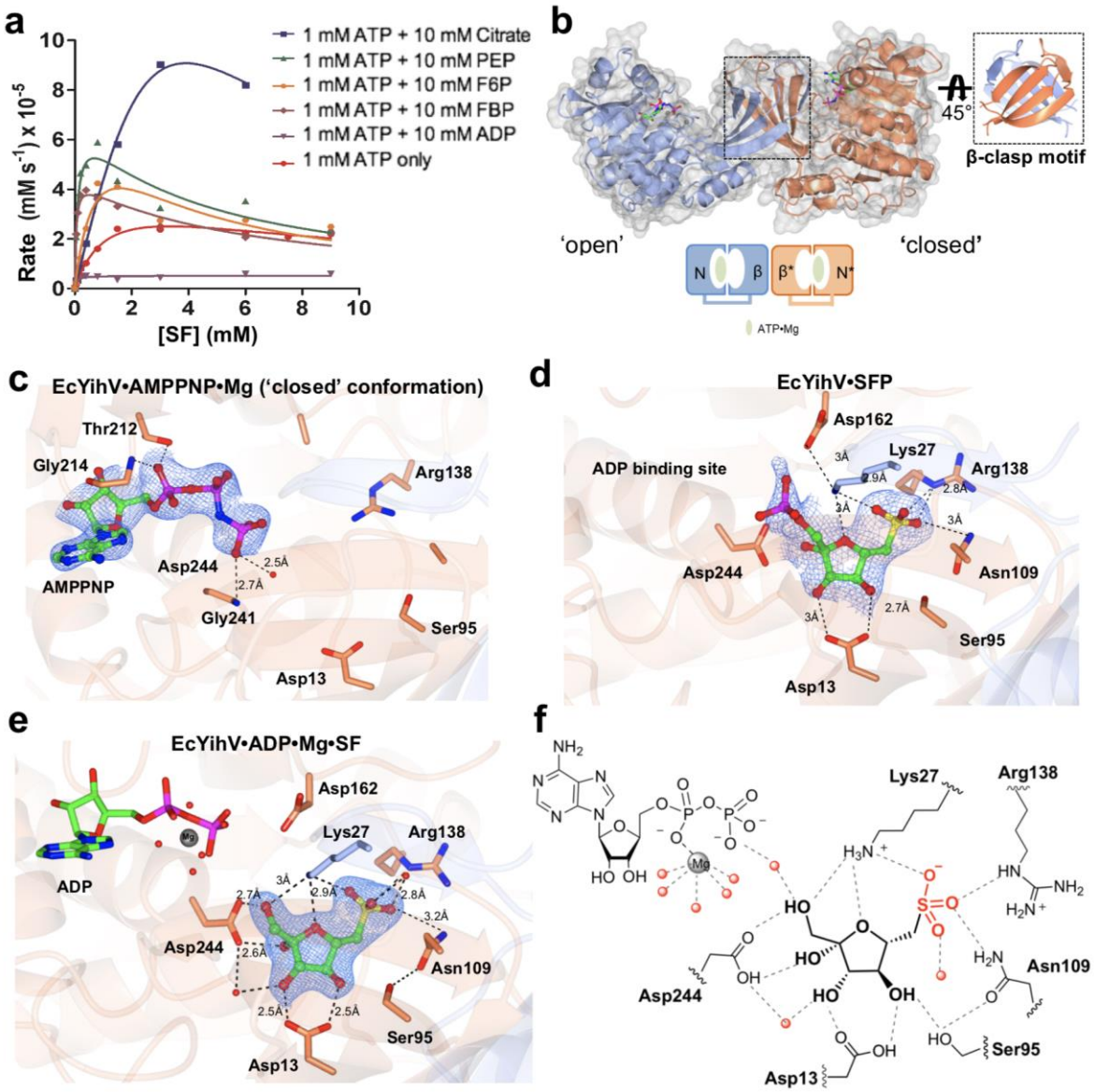


Figure 4. Crystal structures of *Ec*YihV SF kinase. **(a)** Kinetic plots showing effect of metabolites on *Ec*YihV catalyzed phosphorylation of SF to SFP at [ATP] = 1.0 mM. **(b)** (Left) *Ec*YihV dimer in complex with AMPPNP•Mg in open and closed conformations is shown in ribbon with the two subunits shown in coral and blue. Each subunit is composed of two-domain architecture with α/β nucleotide binding domain and β-sheet 'lid' domain. (Right) Lid domains of the dimer form a β-clasp dimerization motif that serves both structural and catalytic roles. **(c)** Close-up view of *Ec*YihV•AMPPNP•Mg showing nucleotide binding site in the closed conformation. **(d)** *Ec*YihV•SFP complex structure and active site interactions with bound SFP product molecule. **(e)** Close-up view of *Ec*YihV•ADP•Mg•SF active site showing hydrogen bonding interactions in quaternary complex. Backbone and carbon atoms of subunits A and B are shown in coral and blue, respectively, and ADP, AMPPNP, SF and SFP are shown in

cylinder format. Electron density corresponds to the 2Fo − Fc and in blue at levels of 1σ for 4c-e. **(f)** Cartoon of ligand binding pocket of *Ec*YihV•ADP•Mg•SF complex depicting hydrogen bonding interactions with active site residues.

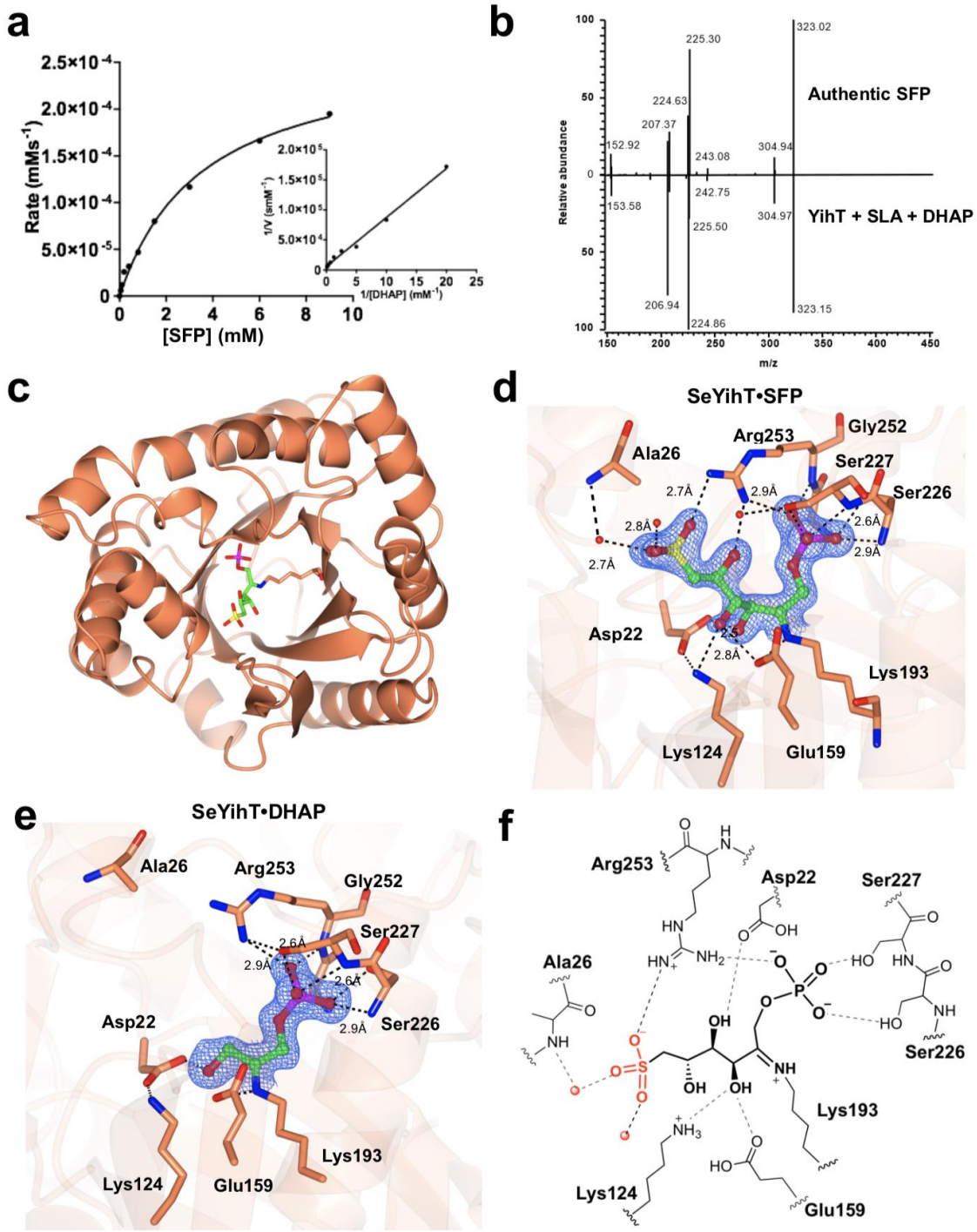


Figure 5. Kinetics, reversibility and structure of *Se*YihT SFP aldolase. (a) Michaelis-Menten and Lineweaver-Burk plots (inset) for *Se*YihT catalyzed conversion of SFP to SLA and DHAP, analysed for DHAP. **(b)** Mass spectrum mirror plot comparison of the product ion scans of the product of the reaction of *Se*YihT incubated with SLA and DHAP, and independently synthesized SFP. **(c)** Overview of the SeYihT protein showing location of substrate-binding site. **(d)** Close-up view of *Se*YihT•SFP active site with Lys193 as a Schiff base bearing the sulfonate ligand. Electron density in blue corresponds to 2Fo − Fc at levels of

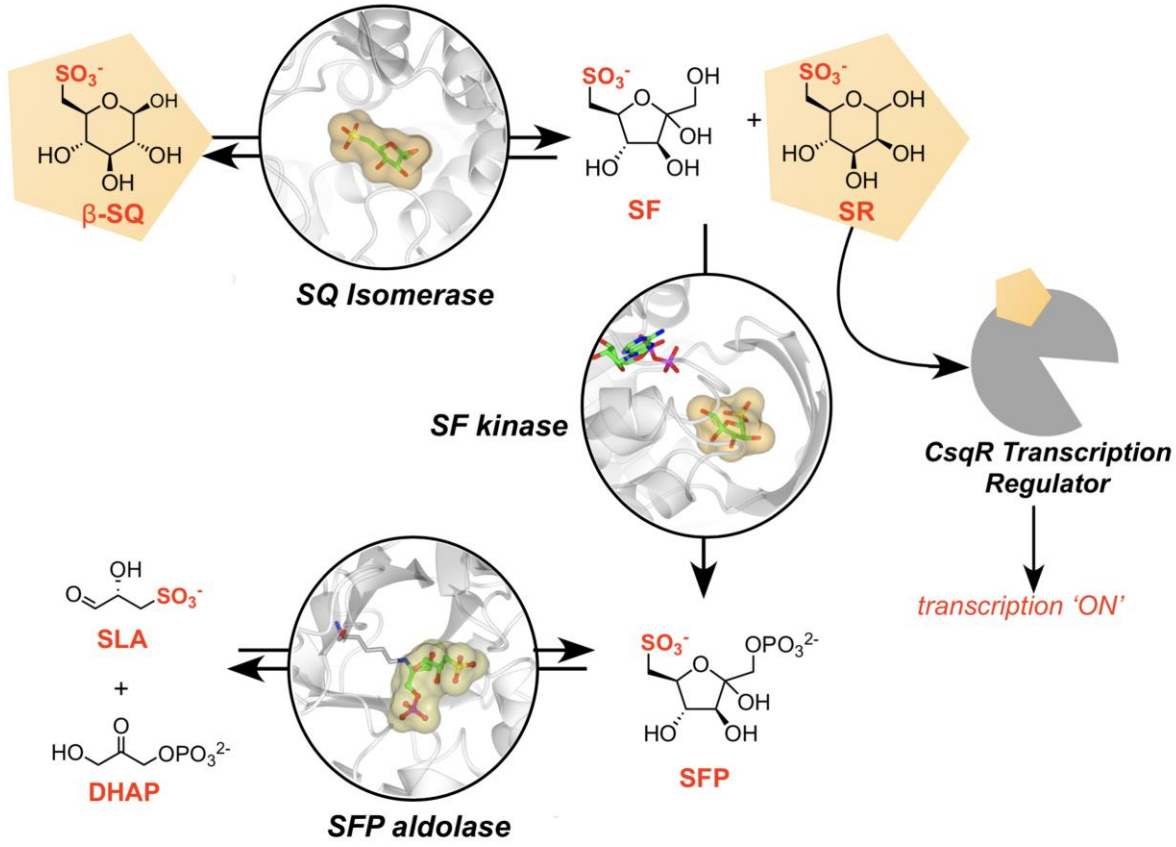
1.2σ. **(e)** Cartoon of ligand binding pocket of *Se*YihT•SFP complex depicting hydrogen bonding and electostatic interactions with active site residues.

Table 1. Kinetic parameters for *Ec*YihV in the presence and absence of cellular metabolites.

Variable[1] substrate	Additive	k_{cat} (s^{-1})	K_M (mM)	K_I (mM)	k_{cat}/K_M ($s^{-1}\,mM^{-1}$)
ATP[2,3]	___	3.1 ± 0.2	1.0 ± 0.2	___	3.2 ± 0.8
SF	___	1.2 ± 0.2	1.3 ± 0.3	8.5 ± 2.8	0.9 ± 0.4
SF[4]	___	___	___	0.30 ± 0.5	0.3 ± 1.1
SF	SQ	0.7 ± 0.1	0.6 ± 0.2	16 ± 9	1.2 ± 0.4
SF	SLA	0.77 ± 0.13	0.60 ± 0.19	6.4 ± 2.7	1.3 ± 0.6
SF	DHAP	1.4 ± 0.3	0.61 ± 0.24	3.0 ± 1.4	2.3 ± 1.4
SF[3]	citrate	3.2 ± 0.4	1.6 ± 0.5	___	2.0 ± 0.9
SF	PEP	1.9 ± 0.2	0.11 ± 0.04	4.4 ± 1.5	17.7 ± 8.6
SF	F6P	2.5 ± 1.2	0.97 ± 0.7	2.6 ± 1.9	2.5 ± 3.1
SF	FBP	1.2 ± 0.1	0.04 ± 0.01	5.8 ± 1.5	27.8 ± 10.8
SF[3]	ADP	0.14 ± 0.01	0.06 ± 0.03	___	2.3 ± 1.3

[1] Reactions were conducted in 200 µL volume in 25 mM BTP buffer (pH 7.5), 25 mM KCl, 5 mM $MgCl_2$, 0.1 mg/mL BSA, 36.6 nM *Ec*YihV at 30 °C, for 60 min. [ATP] = 1.0 mM, and kinetic data were analyzed using a substrate inhibition kinetic model, unless otherwise noted. Additives were at 10 mM.
[2] [SF] = 1.0 mM.
[3] Kinetic parameters obtained by analysis using the Michaelis-Menten equation.
[4] [ATP] = 0.1 mM.

TOC Graphic



Synopsis

Sulfoglycolysis is a pathway for metabolism of the sulfosugar sulfoquinovose. We illuminate how sulfoglycolytic enzymes bind and turnover their substrates, and identify a new transiently-produced intermediate, sulforhamnose.

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 3

SUPPLEMENTARY TABLES 23

SAFETY STATEMENT 26

SECTION 1. CHEMICAL SYNTHESIS 26

General 26

SECTION 2. CLONING, EXPRESSION AND PURIFICATION OF TARGET ENZYMES 29

SECTION 3. AMINO ACID SEQUENCES AND PHYLOGENETIC ANALYSIS 30

Sequence Alignments 30

SECTION 4. BIOPHYSICAL CHARACTERIZATION OF TARGET ENZYMES 31

SEC MALS analysis 31

Nano Differential Scanning Fluorimetry (nanoDSF) 31

SECTION 5. KINETIC ANALYSIS OF TARGET ENZYMES 31

SECTION 6. PROTEIN CRYSTALLISATION 36

Initial screening and optimised crystallization conditions 36

Data collection, processing and refinement 38

SECTION 7. ANALYSIS OF TRANSCRIPTION FACTOR CSQR 38

REFERENCES 39

NMR SPECTRA 41

Supplementary Figures

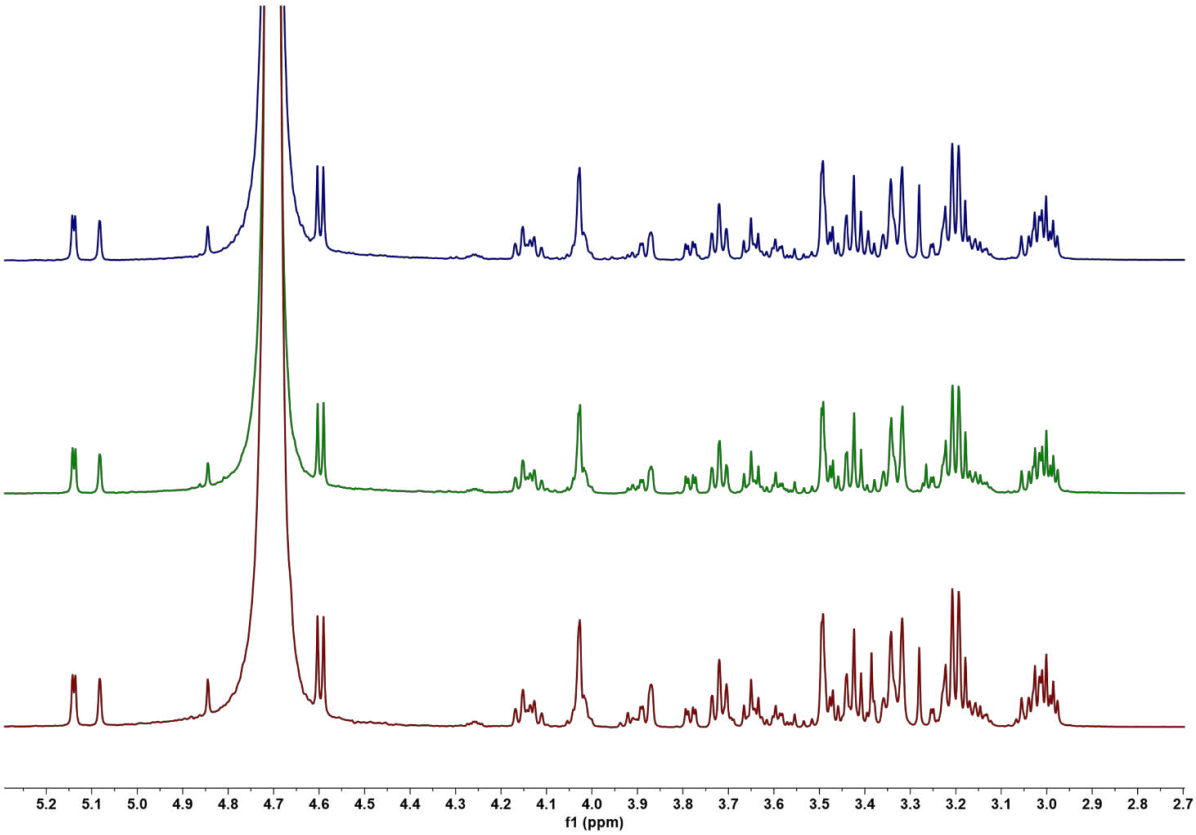


Figure S1. ^{1}H NMR spectra showing *Ec*YihS-catalyzed reactions of SQ (top), SF (middle) or SR (bottom). Reactions consisted of the sugar and *Ec*YihS in 50 mM sodium phosphate, 150 mM NaCl (pH 7.00) buffer at 37 °C for 24 h; then heat inactivated and solvent exchanged to D_2O.

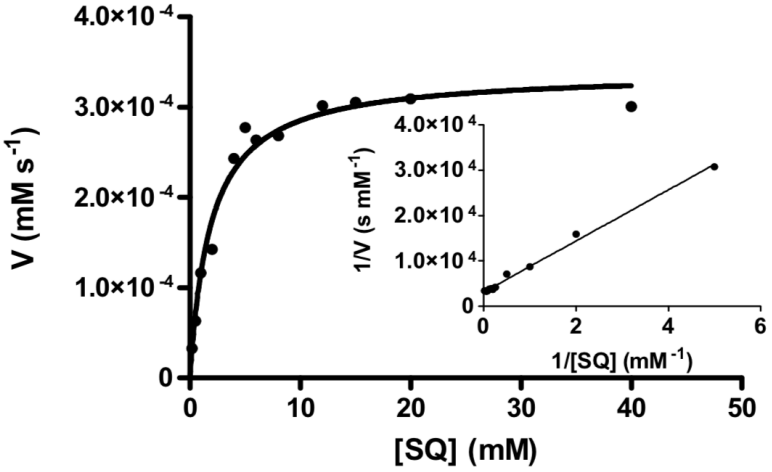


Figure S2. Kinetic characterization of SQ isomerase *Ec*YihS. Michaelis-Menten plot for *Ec*YihS catalyzed isomerization of SQ, inset, Lineweaver-Burk plot.

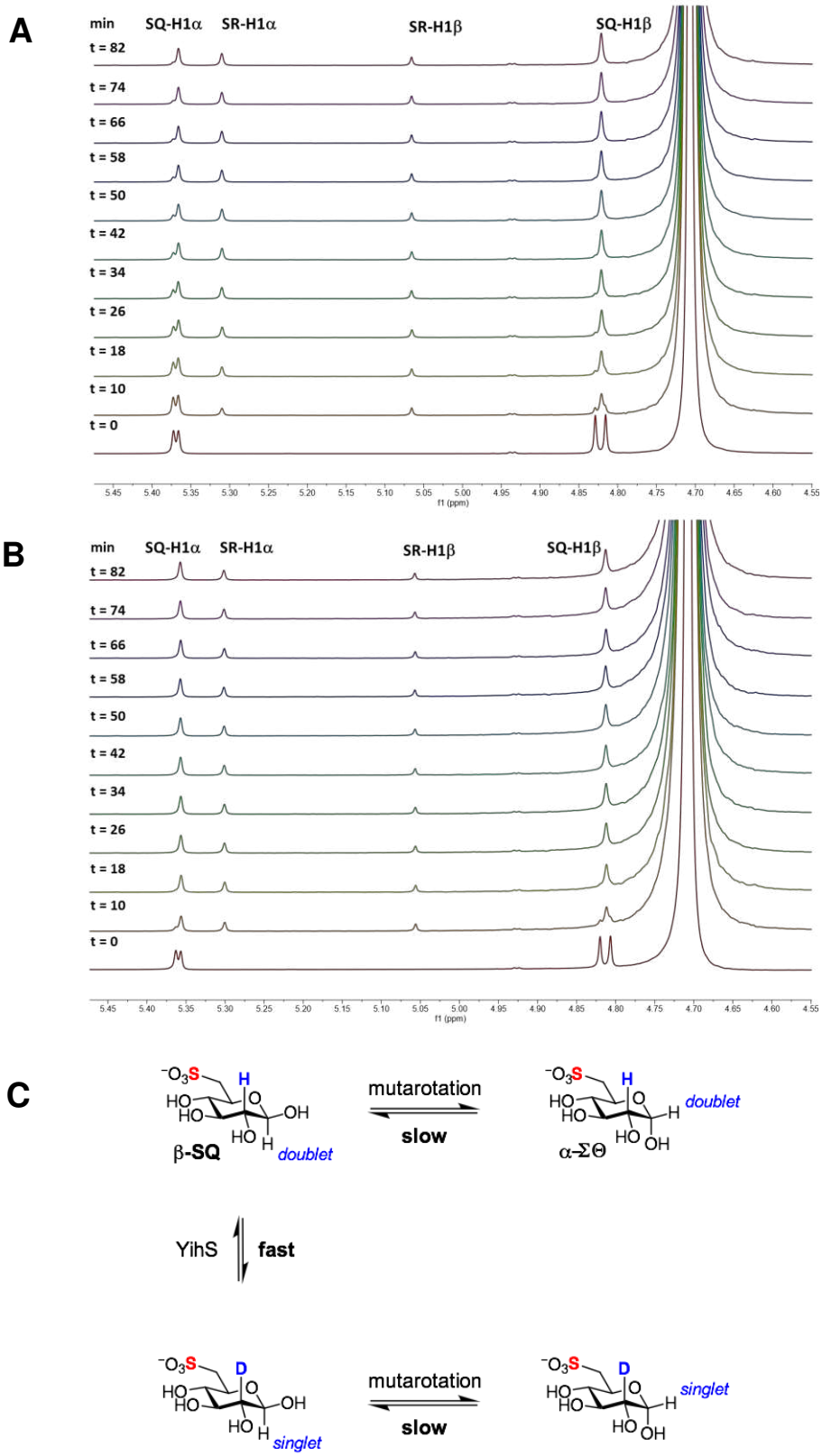


Figure S3. H/D exchange of *Ec*YihS-catalysed conversion of SQ in D_2O. *A,* Time-course ^1H NMR spectra (in D_2O) of SQ incubated with *Ec*YihS; SQ reference spectrum (*t* = 0). *B,* Time-course ^1H NMR spectra of SQ (in D_2O) incubated with *Ec*YihS and *Hs*SQM mutarotase;

SQ reference spectrum (t = 0). *C*, Scheme showing H/D exchange at C2 catalyzed by YihS isomerase.

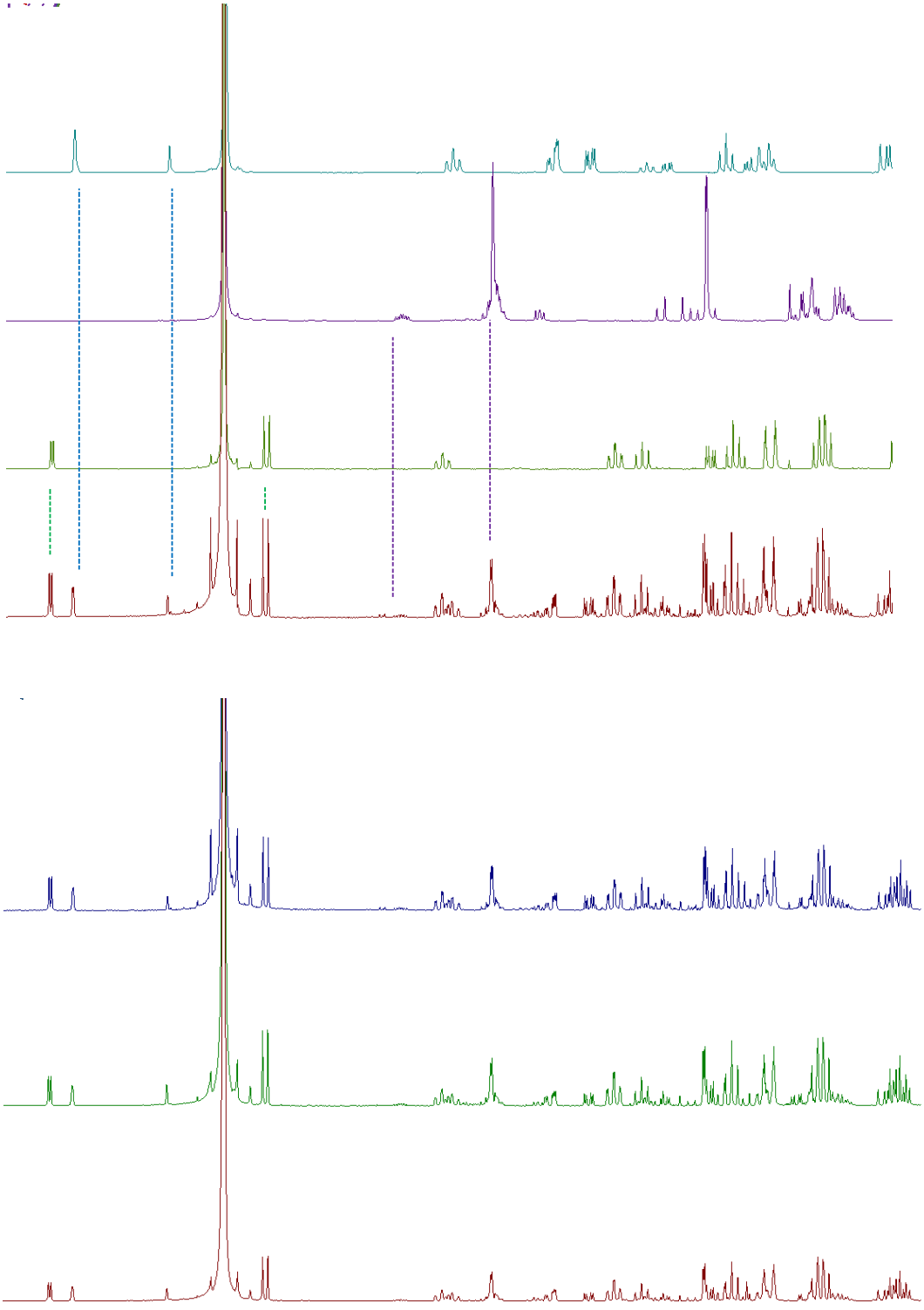
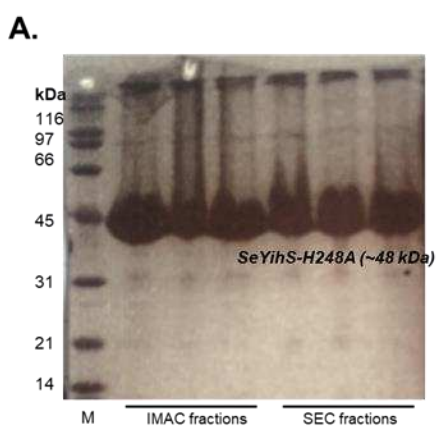
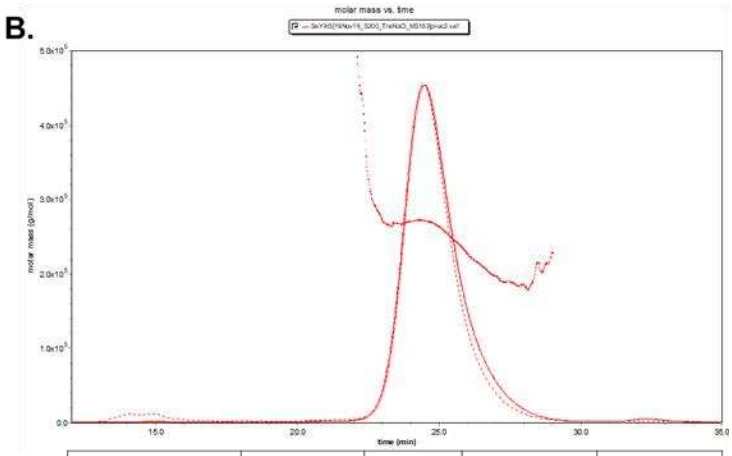


Figure S4. (Top panel) ^{1}H NMR spectrum of SQ incubated with *Se*YihS for 24 h in 50 mM sodium phosphate, 150 mM NaCl (pH 7.00) buffer at 37 °C, then heat inactivated and solvent exchanged to D_2O, along with reference spectra of SQ, SF and SR in D_2O. (Bottom panel) ^{1}H NMR spectra showing *Se*YihS-catalyzed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

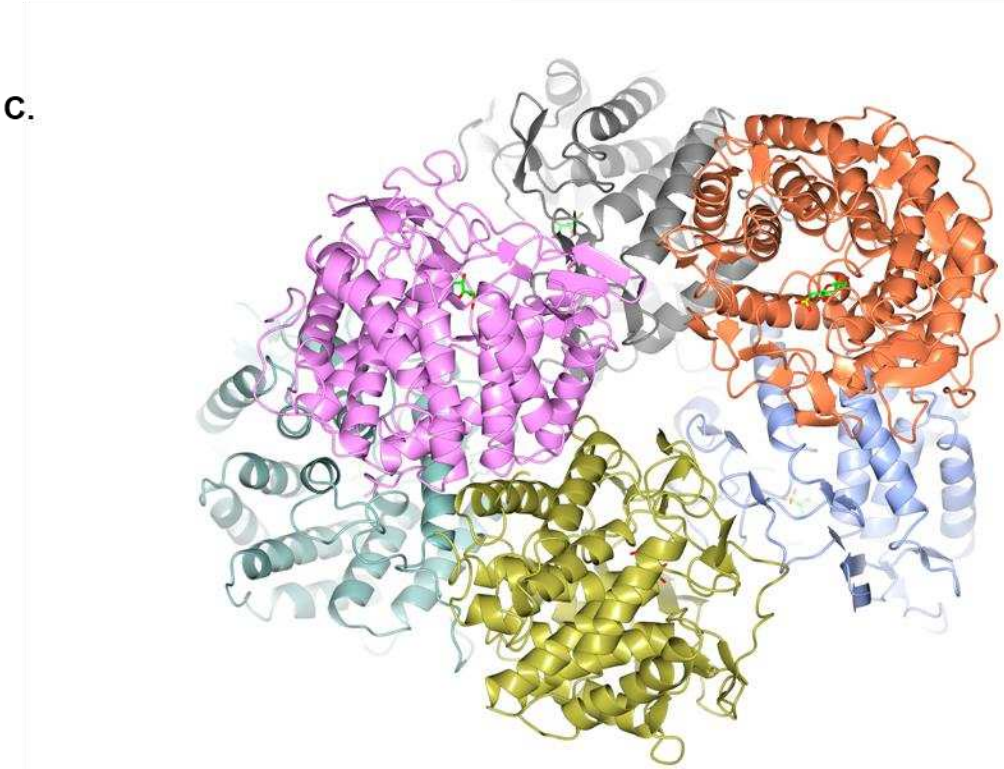


Figure S5. Purification, properties and molecular assembly of *Se*YihS-H248A. *A,* SDS-PAGE analysis of purified *Se*YihS after IMAC and size exclusion chromatography (SEC) showing expected Mw of 48,000 Da. *B,* SEC-MALS plot reveals oligomeric state of *Se*YihS-WT in solution. UV-trace and an average 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.

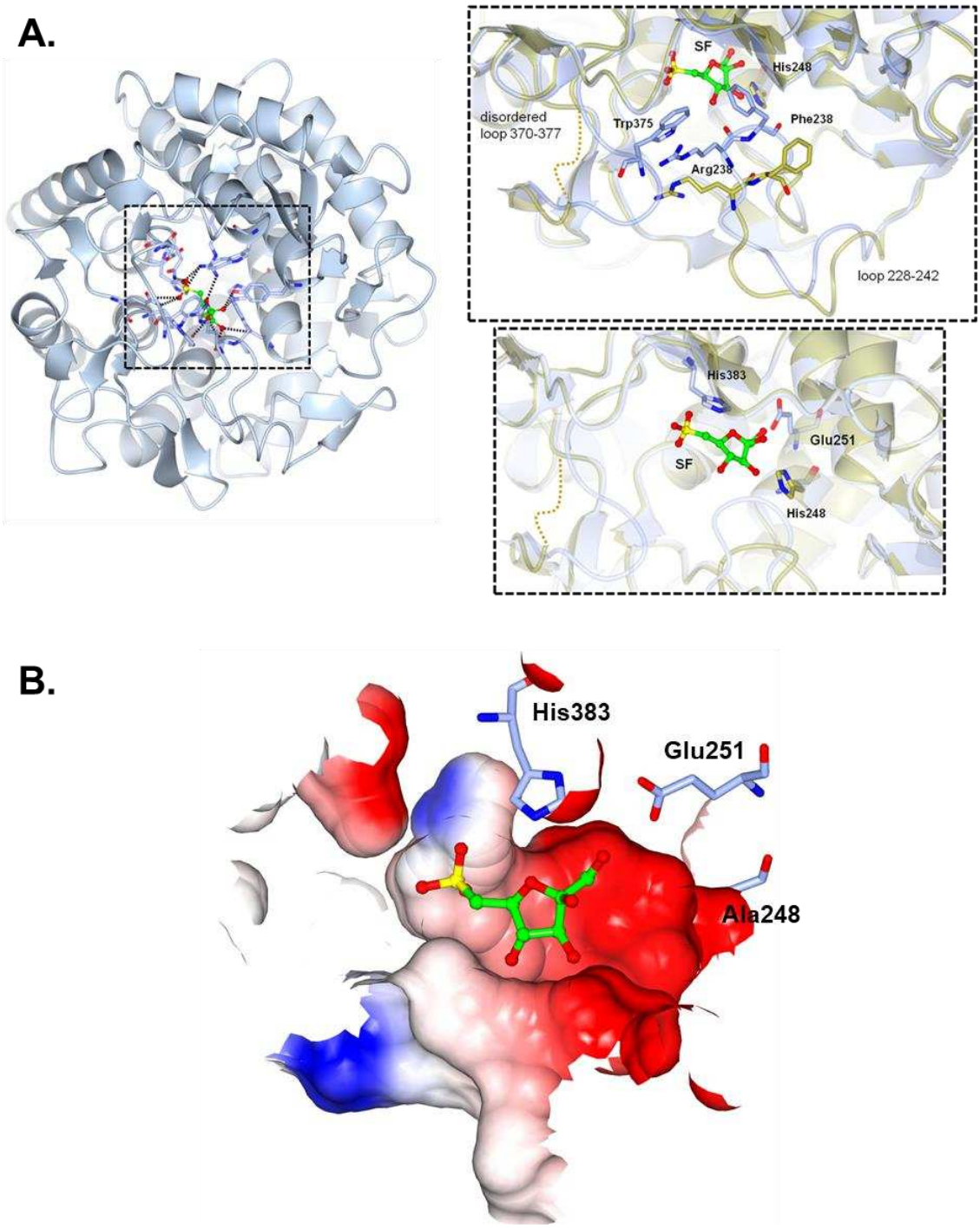


Figure S6. Overlay of wild-type *Se*YihS and active site mutant *Se*YihS-H248A•SF showing loop rearrangements upon binding sulfofructose. *A,* Active sites of wild-type (2AFA.pdb) and ligand bound structures (in gold and blue respectively) show re-ordering of loops, cation-pi stacking interaction of 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.

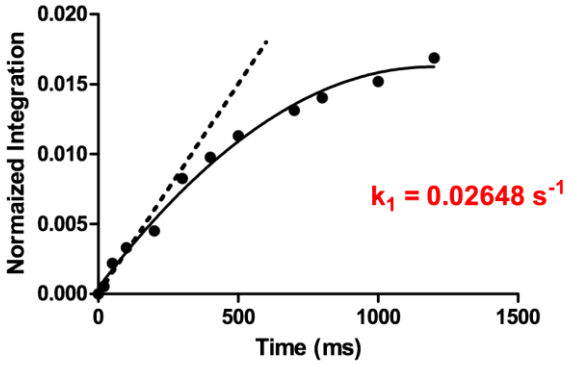
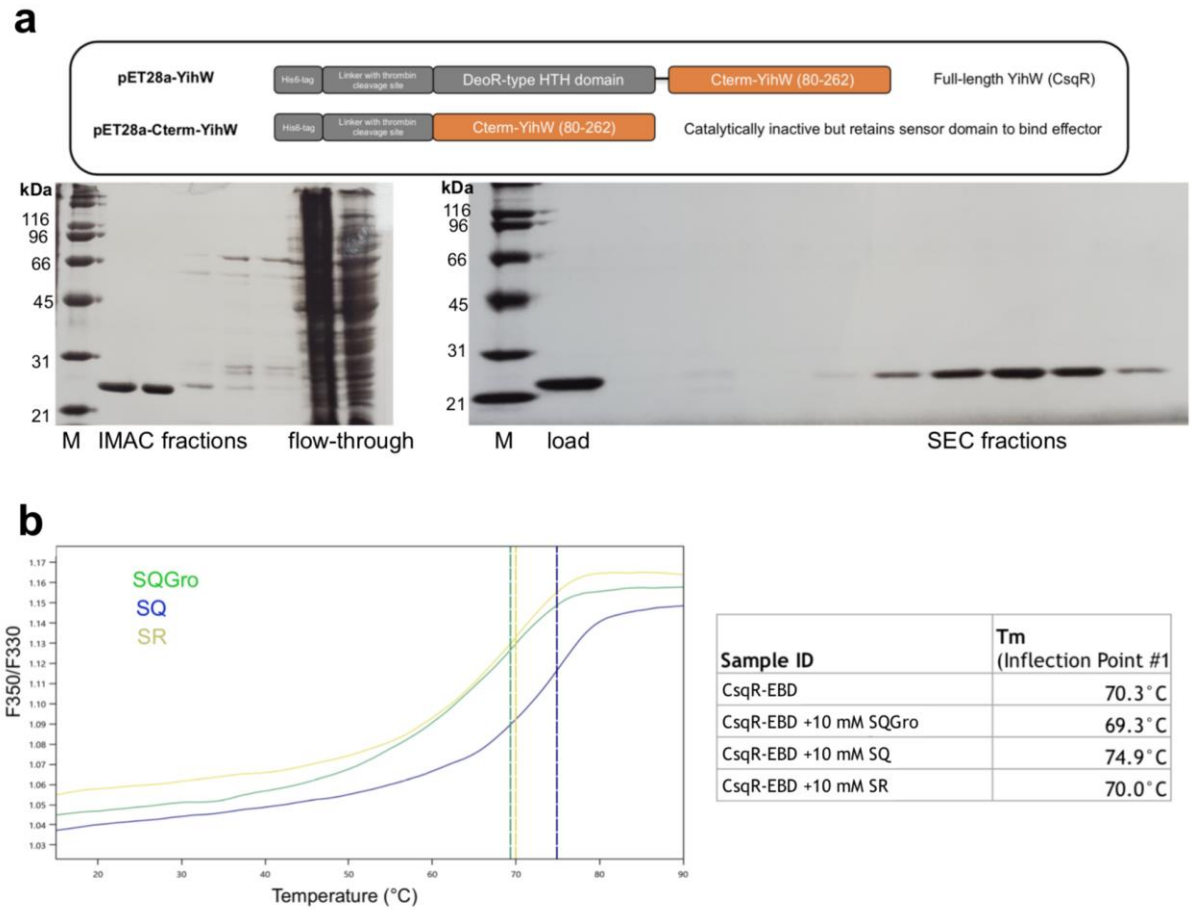
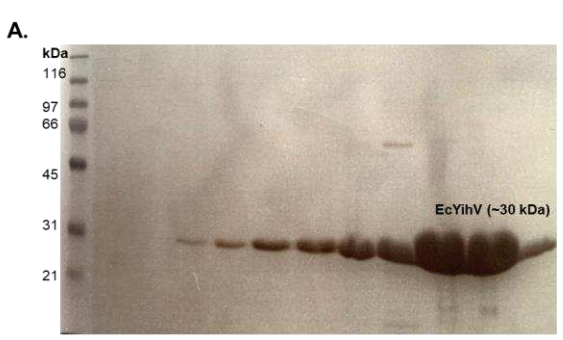


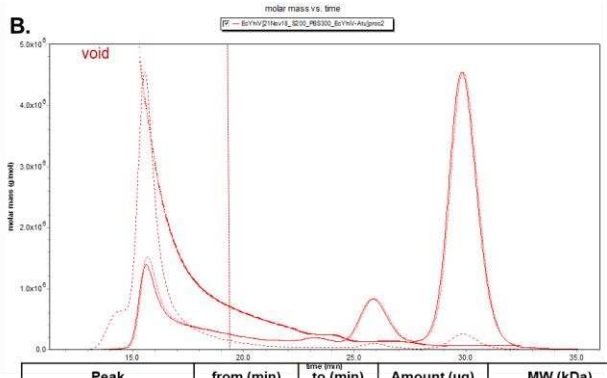
Figure S7. Kinetic analysis of SF mutarotation by inversion recovery 1D ^{1}H exchange spectroscopy. Inversion recovery curves for 5 mM β-SF at δ 4.25 ppm using a Gaussian-shaped pulse of 20 ms. Dashed line indicates tangent to the fitted curve at $t = 0$.



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).

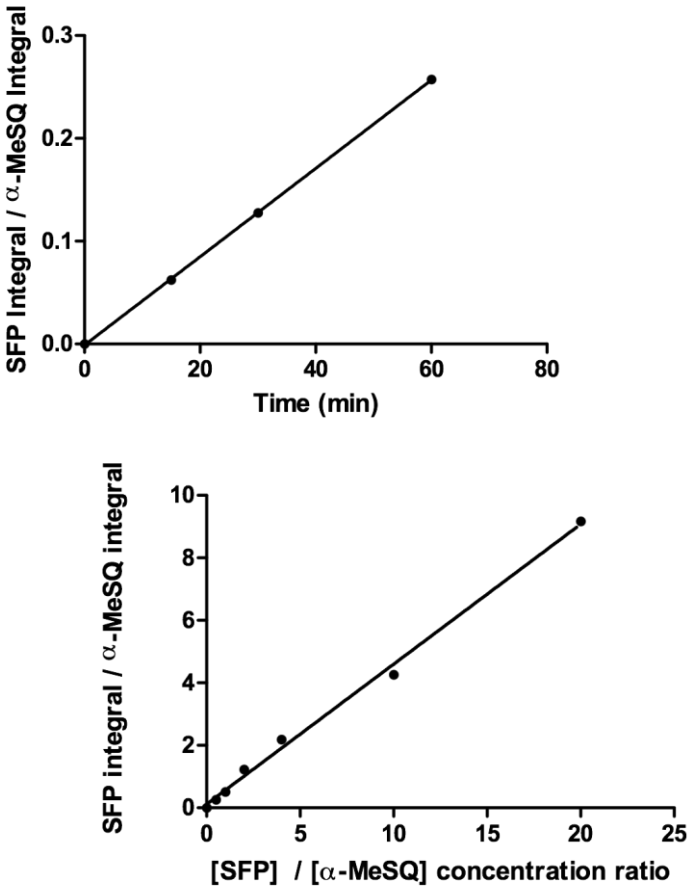


Figure S10. LC-MS/MS assay for SF kinase. (Left) Linearity of the *Ec*YihV kinase reaction with time. (Right) Calibration curve for detection of SFP using methyl α-sulfoquinovoside (α-MeSQ) as an internal standard. Reactions were carried out in 25 mM BTP (pH 7.5), 25 mM KCl, 5 mM $MgCl_2$, 0.1 mg/mL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

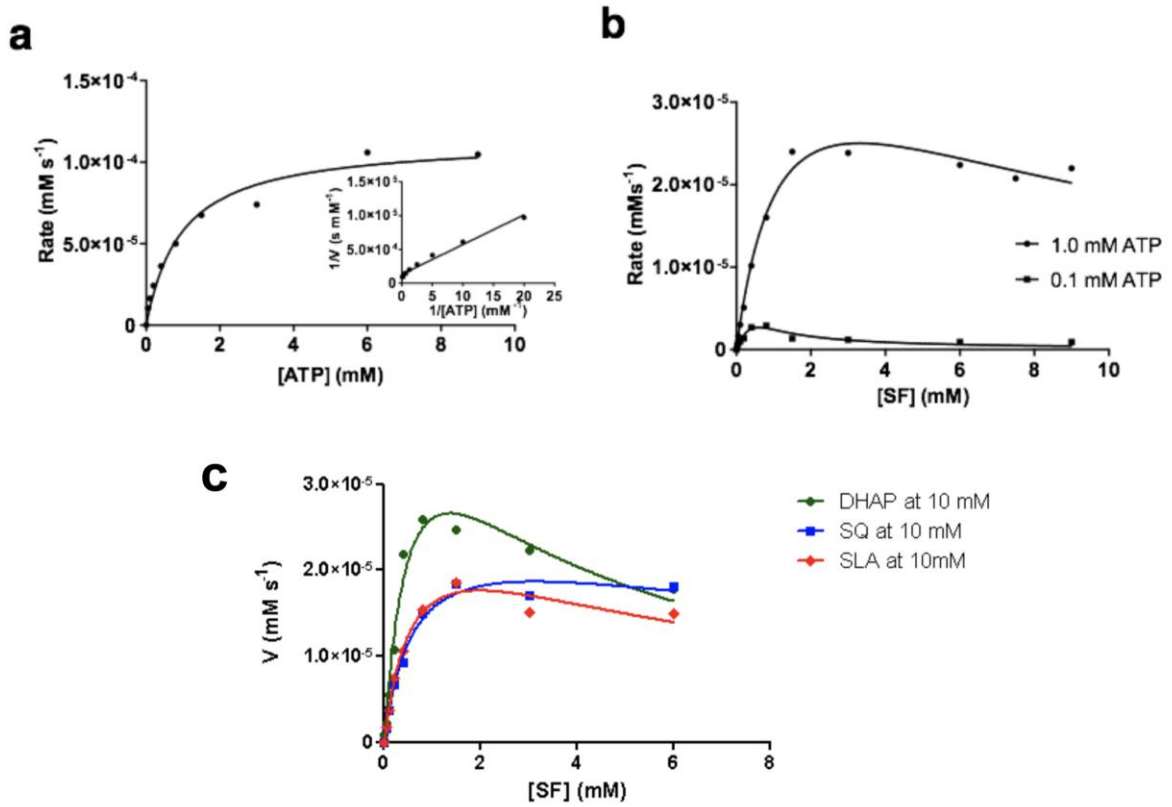


Figure S11. Characterization of *Ec*YihV. (a) Michaelis-Menten and Lineweaver-Burk plots (inset) for YihV catalyzed phosphorylation of SF to SFP at [SF] = 1.0 mM with the variation of ATP. **(b)** Substrate inhibition plots for *Ec*YihV catalyzed phosphorylation of SF to SFP at [ATP] = 1.0 mM and 0.1 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.

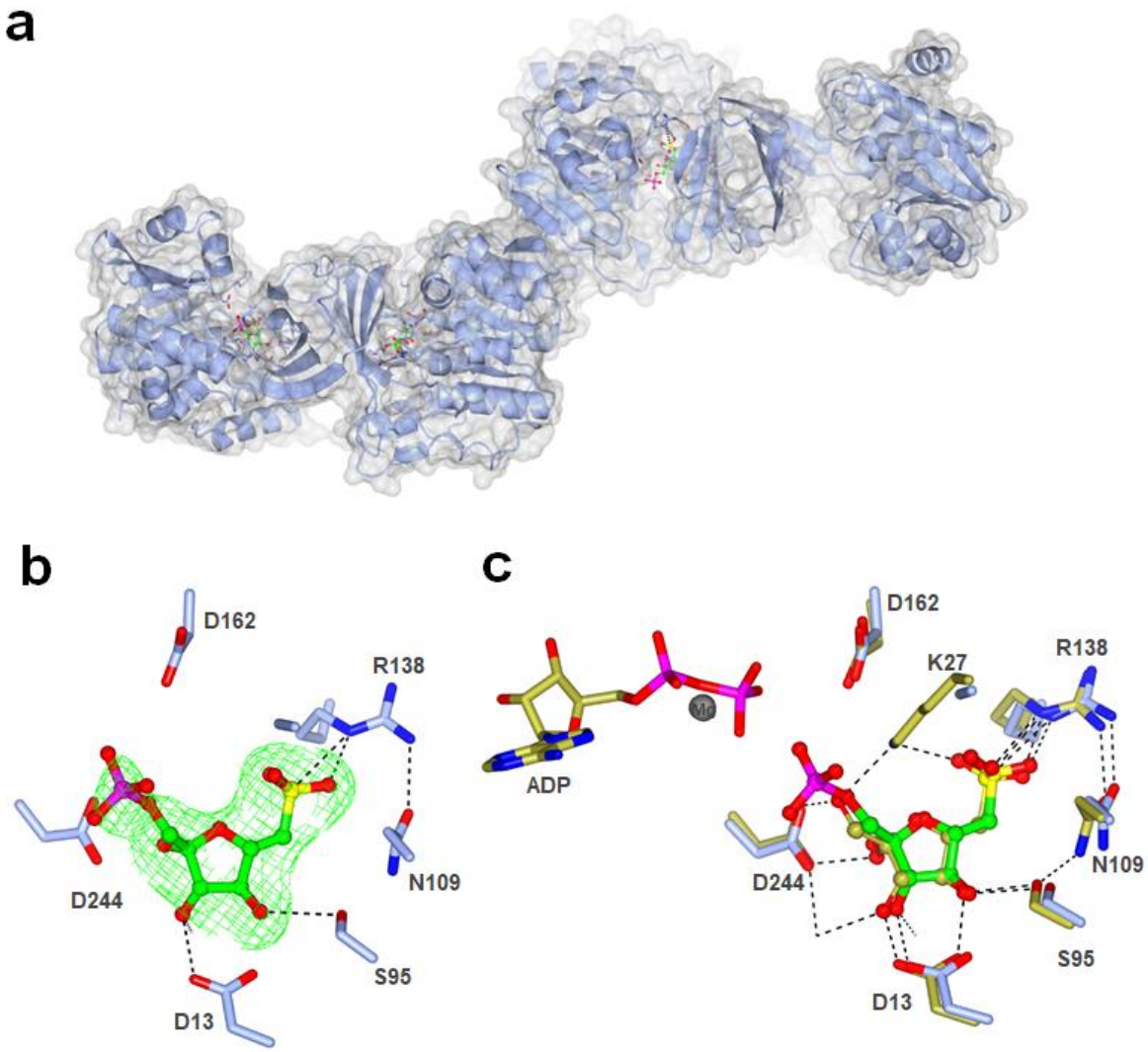


Figure S12. (a) Crystallographic dimer-of-dimer assembly of YihV-SFP complex. **(b)** Omit map of ligand SFP. Electron density represents the Fo–Fc (omit map, in green) obtained prior to modelling of the ligand, and contoured at levels of 3σ. **(c)** Overlay of SFP complex with YihV-ADP-Mg-SF complex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

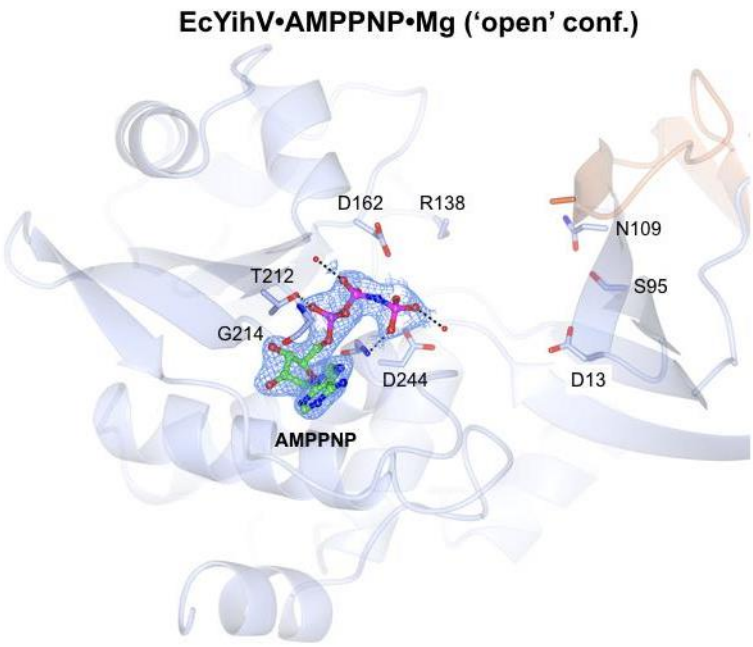
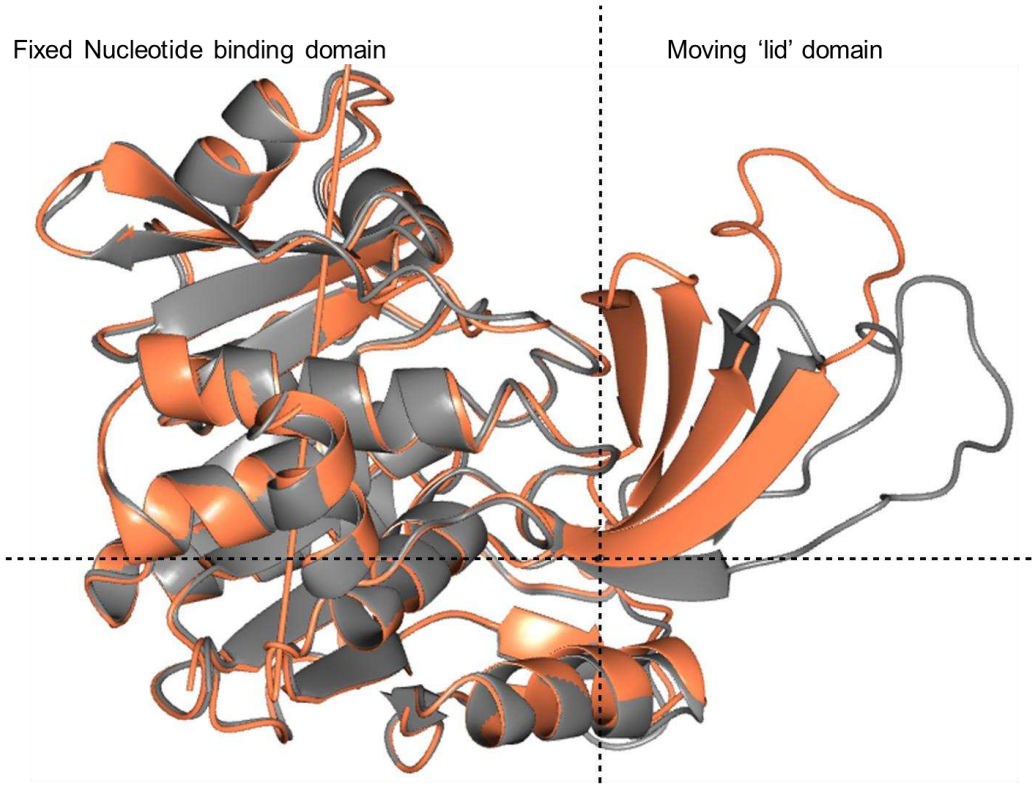


Figure S13. Crystal structures of *Ec*YihV SF kinase. *A*, Close-up view of hydrogen bonding interactions of *Ec*YihV dimer in complex with AMPPNP•Mg in open conformations with AMPPNP bound at the nucleotide binding site. Backbone and carbon atoms of subunits A and B are shown in coral and blue, respectively, 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 **Domain 2**	Nucleotide binding domain 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.

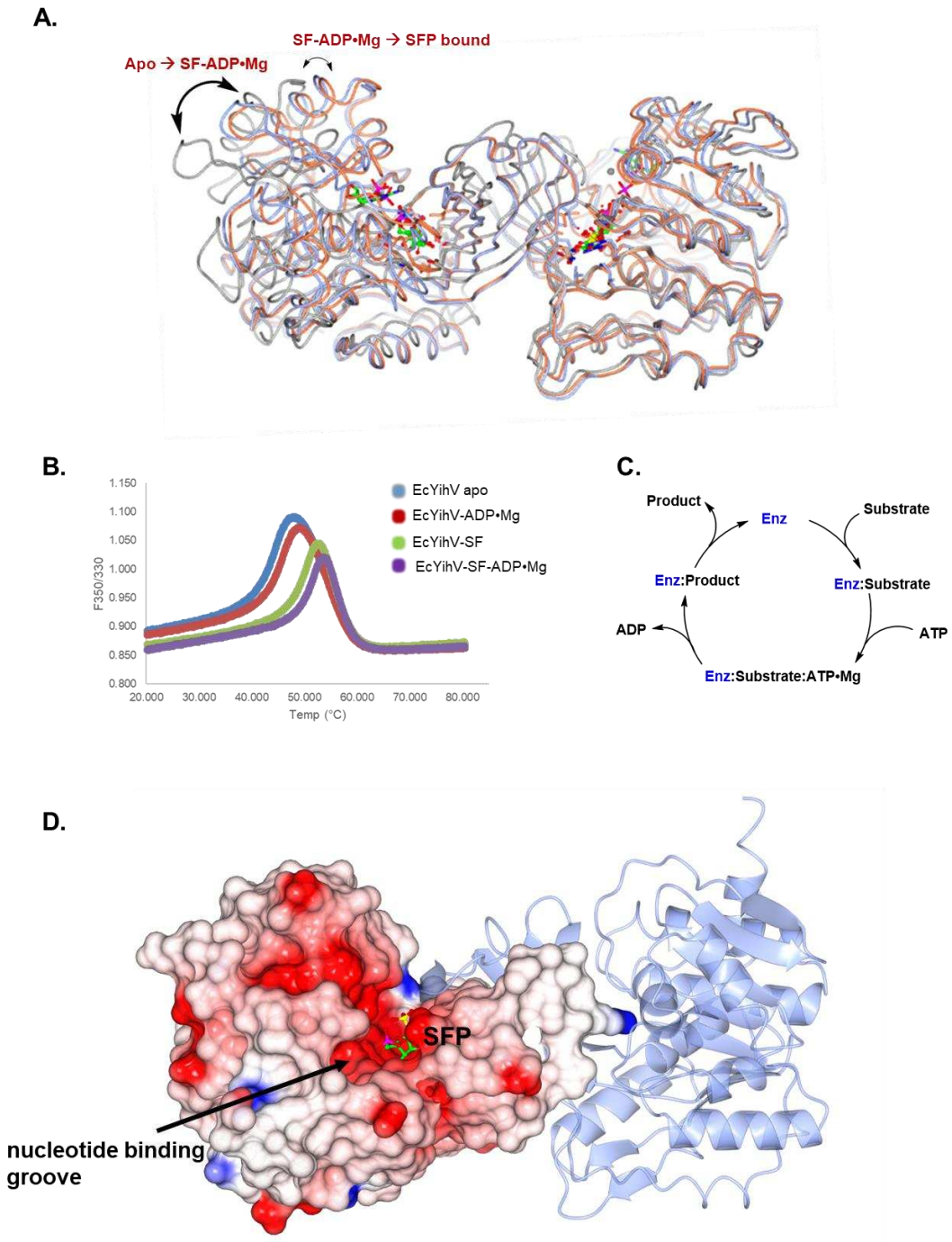


Figure S15. Comparison of *Ec*YihV•AMPPNP, *Ec*YihV•SFP and *Ec*YihV•ADP•Mg•SF complexes. *A,* Overlay of YihV dimer (in coral and blue) of three complexes to show open, closed and slightly staggered conformations. *B*, Temperature unfolding assay with *Ec*YihV reveals T_m shifts upon binding substrates 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.

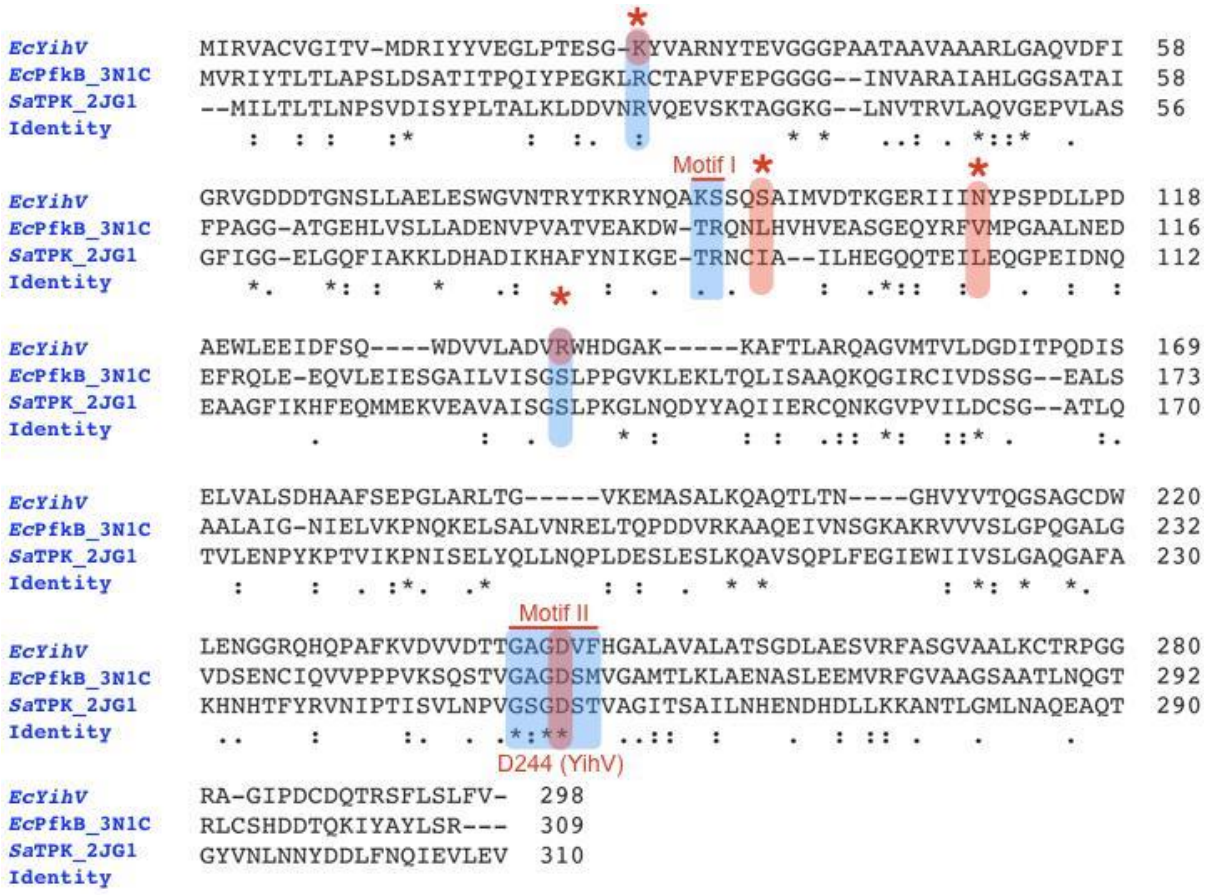


Figure S16. ***Ec*YihV shares a catalytic motif with phosphosugar kinases (PfkB and TPK) and possesses a unique sulfonate substrate binding sequence.** Sequence alignment of *Ec*YihV with model phosphosugar kinases shows conserved catalytic motif (*GXGDXX* sequence, Motif II) but lacks TR motif (Motif 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.

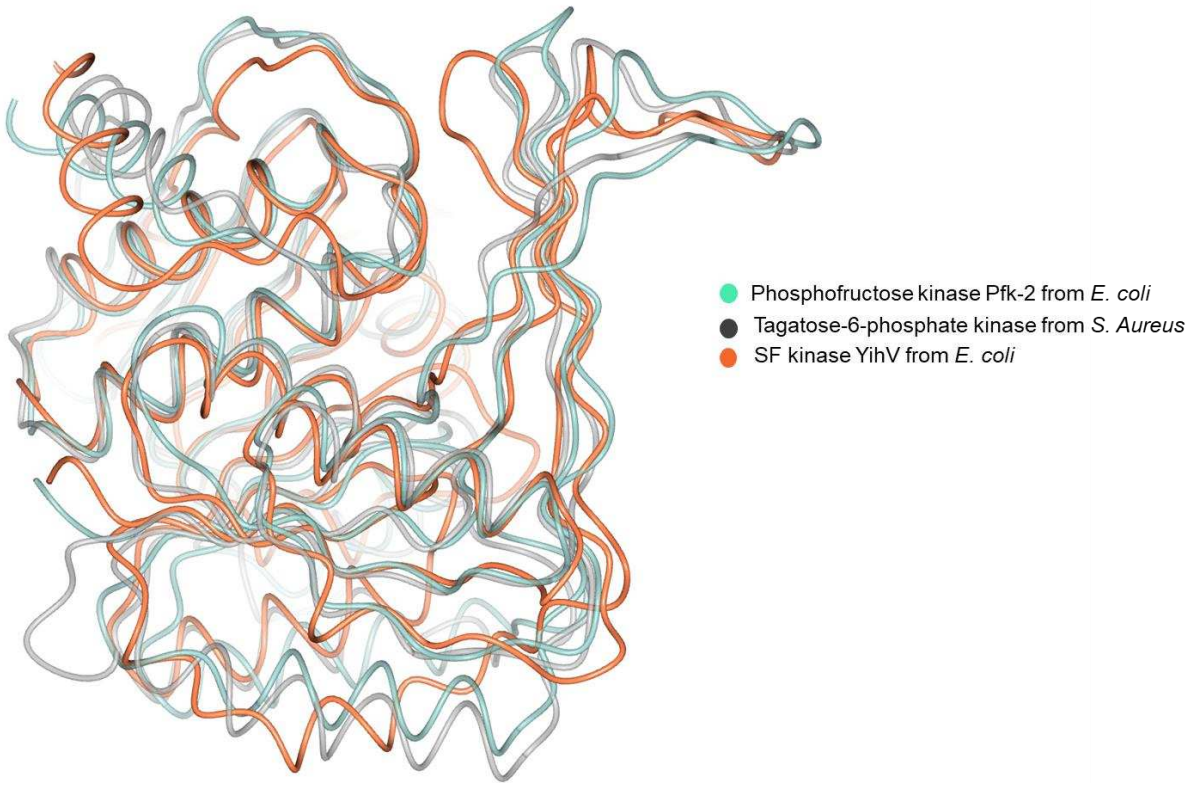


Figure S17. Closed conformations of PfkB, TPK and SF kinases show close structural relationships. *A,* Overlay of PfkB•F6P (in 3N1C.pdb) and TPK•ADP•T6P (2JG1.pdb) and *Ec*YihV•ADP•SF structures.

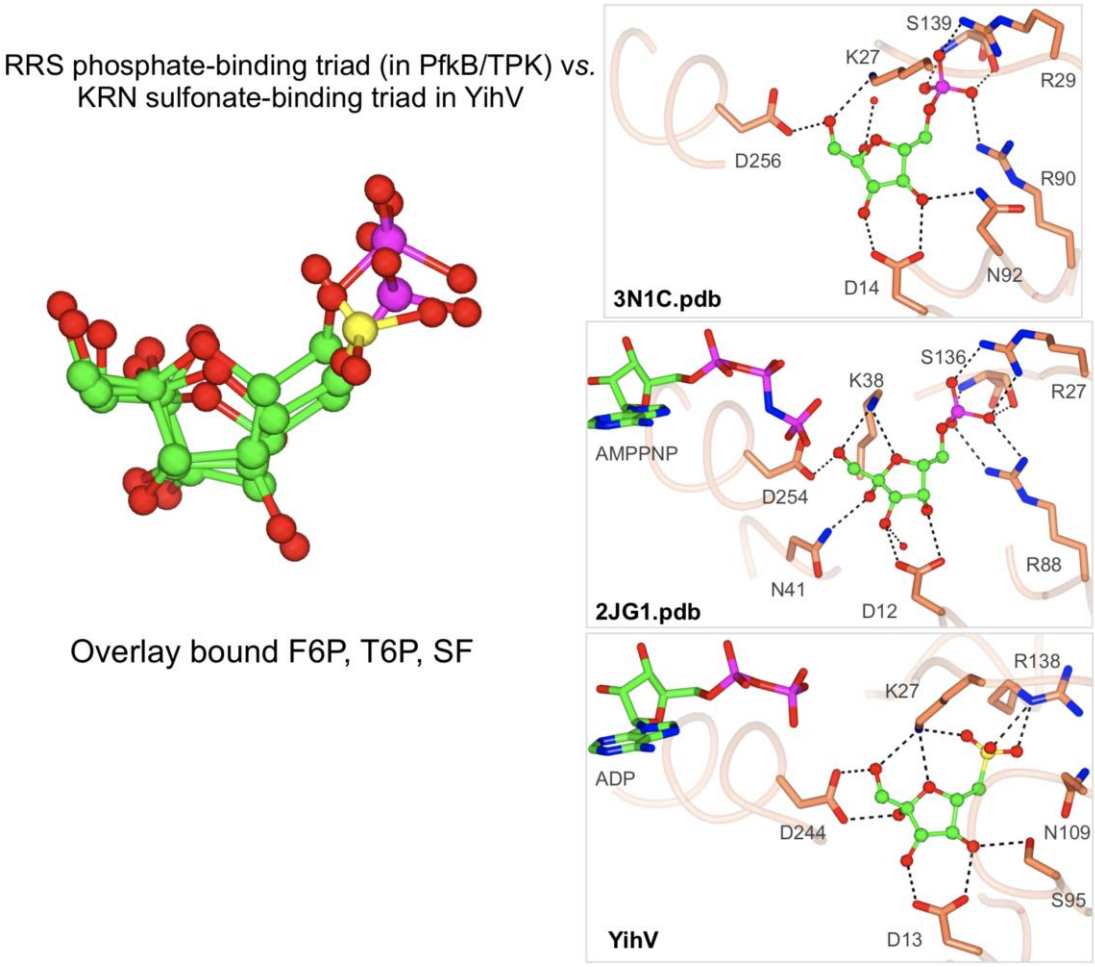
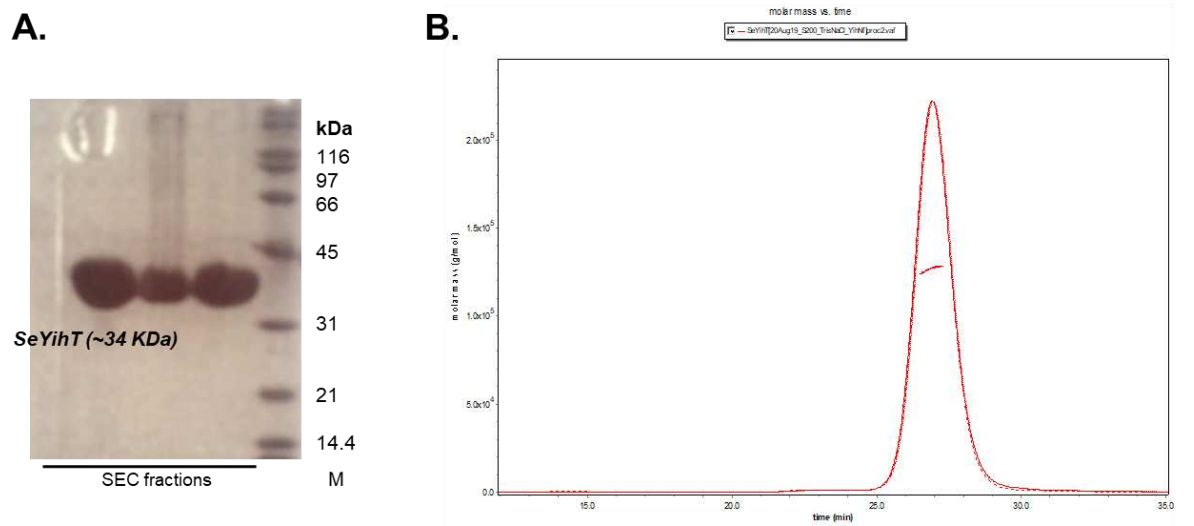


Figure S18. Structurally conserved motifs in PfkB, TPK and SF kinases provide distinct binding modes for phosphate *vs.* sulfonate substrates. (Left) Overlay of respective ligands (F6P, T6P and SF) highlighting the similarities in binding pose. (Right) Active site shown with partial **GXGDXX** and **TR** 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.

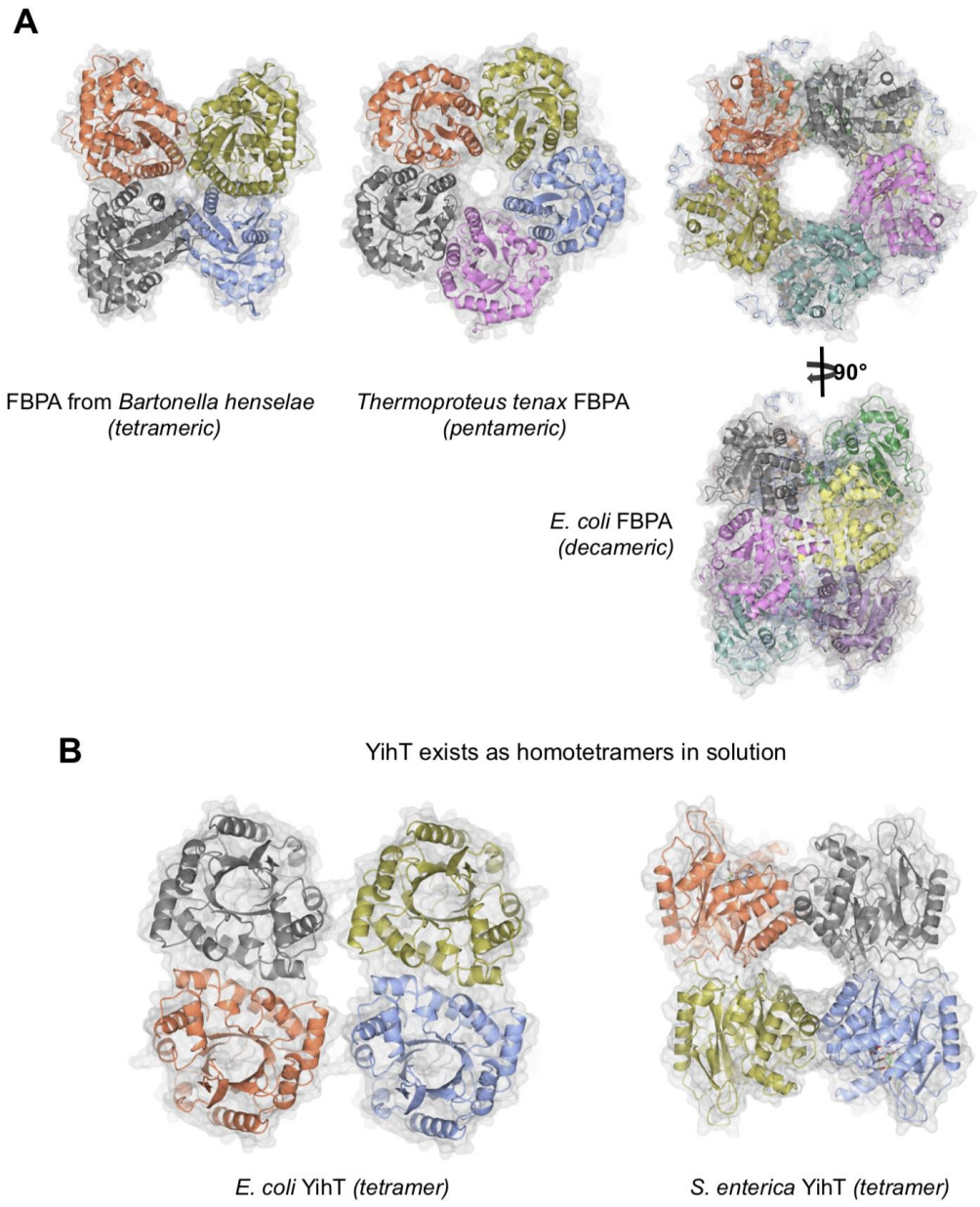


Figure S20. Oligomeric states of prokaryotic Class I fructose bisphosphate aldolases (FBPA). *A,* FBPA from *Bartonella henselae* forms homotetramers (3MMT.pdb),[1] FBPA from *Thermoproteus tenax* forms homopentamers (1W8S.pdb)[2] and Class I FBPA from *E. coli*[3] forms homodecamers in solution (model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

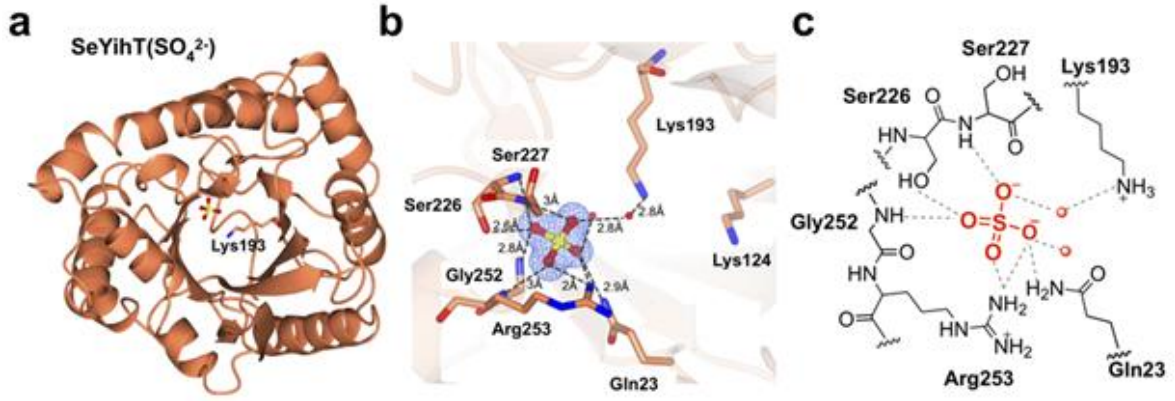


Figure S21. (a) Overview of the SeYihT•sulfate complex showing location of sulfate-binding site. **(b)** Close-up view of *Se*YihT sulfate binding site. Electron density in blue corresponds to the 2Fo − Fc at levels of 1σ. **(c)** Cartoon of sulfate binding pocket of *Se*YihT•sulfate complex depicting 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.

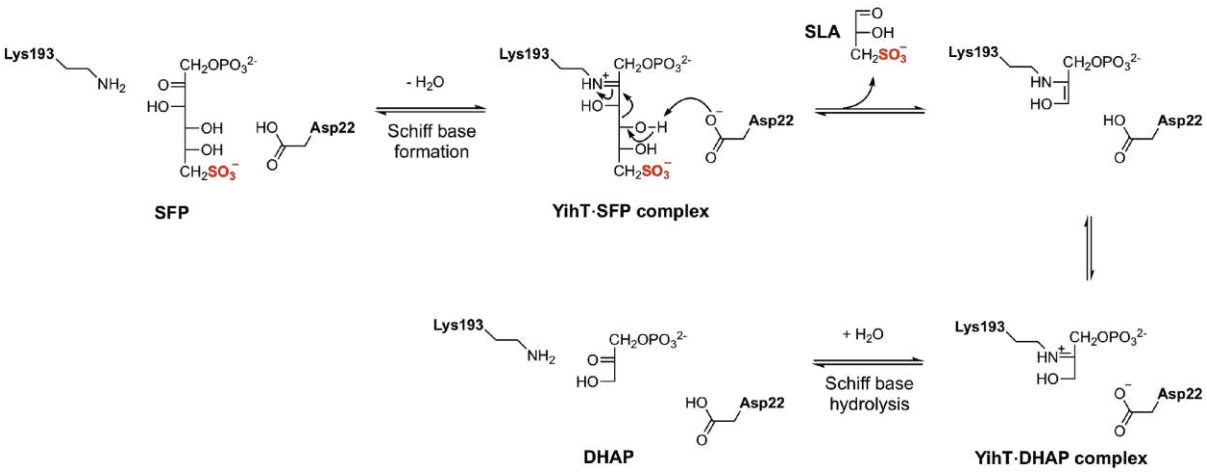


Figure S22. Proposed mechanism for SFP aldolase, with residue numbers corresponding to *Se*YihT.

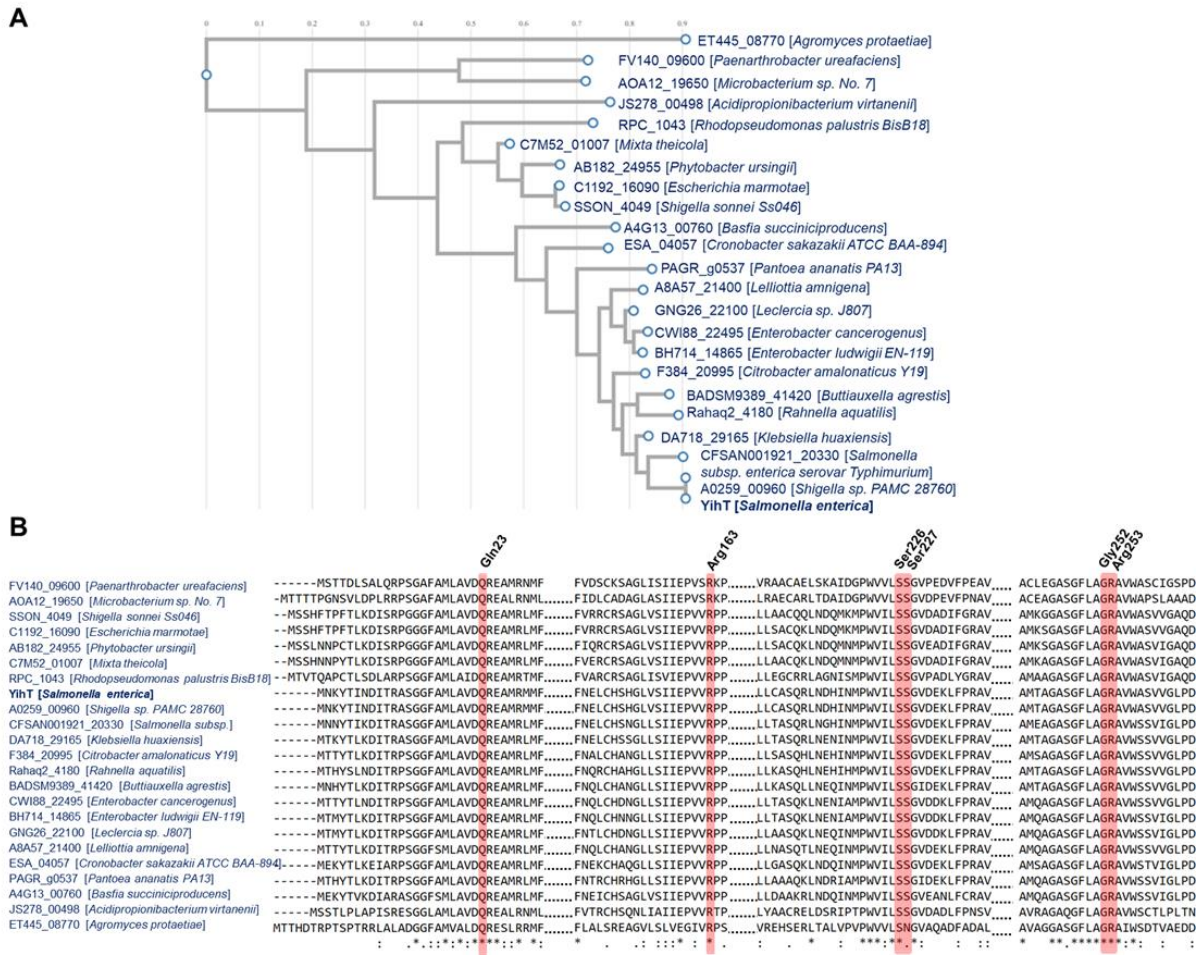


Figure S23. Phylogenetic analysis of SFP aldolases. *A,* Phylogenetic tree showing distribution of bacterial YihT enzymes. *B*, Multiple sequence alignment of YihT enzymes from different bacterial sources showing conserved anion (phosphate/sulfonate) binding residues.

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)
$I / \sigma I$	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)
$I / \sigma I$	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, CO**C**H_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, $CHCl_3$); ^{1}H 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, $CHCl_3$); ^{1}H 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); ^{1}H 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); ^{1}H 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); ^{1}H 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 ractions 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

MKWFNTLSHNRWLEQETDRIFNFGKNAVVPTGFGWLGNKGQIKEEMGTHLWITARMLHVYSVAASMGR
PGAYDLVDHGIKAMNGALRDKKYGGWYACVNDQGVVDASKQGYQHFFALLGAASAVTTGHPEARKLLD
YTIEVIEKYFWSEEEQMCLESWDEAFSQTEDYRGGNANMHAVEAFLIVYDVTHDKKWLDRALRIASVI
IHDVARNGDYRVNEHFDSQWNPIRDYNKDNPAHRFRAYGGTPGHWIEWGRLMLHLHAALEARFETPPA
WLLEDAKGLFHATIRDAWAPDGADGFVYSVDWDGKPIVRERVRWPIVEAMGTAYALYTLTDDSQYEEW
YQKWWDYCIKYLMDYENGSWWQELDADNKVTTKVWDGKQDIYHLLHCLVIPRLPLAPGLAPAVAAGLL
DINAK

EcYihS wild-type

MKWFNTLSHNRWLEQETDRIFDFGKNSVVPTGFGWLGNKGQIKEEMGTHLWITARMLHVYSVAAAMGR
PGAYSLVDHGIKAMNGALRDKKYGGWYACVNDEGVVDASKQGYQHFFALLGAASAVTTGHPEARKLLD
YTIEIIEKYFWSEEEQMCLESWDEAFSKTEEYRGGNANMHAVEAFLIVYDVTHDKKWLDRAIRVASVI
IHDVARNNHYRVNEHFDTQWNPLPDYNKDNPAHRFRAFGGTPGHWIEWGRLMLHIHAALEARCEQPPA
WLLEDAKGLFNATVRDAWAPDGADGIVYTVDWEGKPVVRERVRWPIVEAMGTAYALYTVTGDRQYETW
YQTWWEYCIKYLMDYENGSWWQELDADNKVTTKVWDGKQDIYHLLHCLVIPRIPLAPGMAPAVAAGLL
DINAKLEHHHHHH

EcYihV wild-type

MIRVACVGITVMDRIYYVEGLPTESGKYVARNYTEVGGGPAATAAVAAARLGAQVDFIGRVGDDDTGN
SLLAELESWGVNTRYTKRYNQAKSSQSAIMVDTKGERIIINYPSPDLLPDAEWLEEIDFSQWDVVLAD
VRWHDGAKKAFTLARQAGVMTVLDGDITPQDISELVALSDHAAFSEPGLARLTGVKEMASALKQAQTL
TNGHVYVTQGSAGCDWLENGGRQHQPAFKVDVVDTTGAGDVFHGALAVALATSGDLAESVRFASGVAA
LKCTRPGGRAGIPDCDQTRSFLSLFVLEHHHHHH

EcYihT wild-type

MNKYTINDITRASGGFAMLAVDQREAMRMMFAAAGAPAPVADSVLTDFKVNAAKALSPYASAILVDQQ
FCYRQVVEQNAIAKSCAMIVAADEFIPGNGIPVDSVVIDRKINPLQIKQDGGKALKLLVLWRSDEDAQ
QRLDMVKEFNELCHSHGLVSIIEPVVRPPRRGDKFDREQAIIDAAKELGDSGADLYKVEMPLYGKGPQ
QELLCASQRLNDHINMPWVILSSGVDEKLFPRAVRVAMTAGASGFLAGRAVWASVVGLPDNELMLRDV
CAPKLQQLGDIVDEMMAKRRLE

SeYihT wild-type

MNNYTIKDITRASGGFAMLAVDQREAMRLMFAAAGAKTPVADSVLTDFKVNAAKILSPYASAVLLDQQ
FCYRQAVEQNAVAKSCAMIVAADDFIPGNGIPVDNVVLDKKINAQAVKRDGAKALKLLVLWRSDEDAQ
QRLNMVKEFNELCHSNGLLSIIEPVVRPPRCGDKFDREQAIIDAAKELGDSGADLYKVEMPLYGKGAR
SDLLTASQRLNGHINMPWVILSSGVDEKLFPRAVRVAMEAGASGFLAGRAVWSSVIGLPDTELMLRDV
SAPKLQRLGEIVDEMMAKRR

EBD-CsqR or EcYihW (C-terminal residues 80-262 of effector binding domain)
MGSSHHHHHHSSGLVPRGSHTAEKRAIAEAVADYLPERCTVFITIGTTVEAVARALLNRRDLRIITNS
LRVAQILYKNQDIEVMVPGGTLRAHNGGIIGPGAVDFIEGFRADYLITSIGAIEHDGTLLEFDVNEAL
VARTMIKHARNTLLVADHTKFAASAAVSIGNARNVRAFFTDAPPPNSFCQLLSEENVELVVAEQEVS

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 rotary evaporator and dissolved in D_2O and analysed by NMR spectroscopy. ^{1}H 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 ^{1}H 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. ^{1}H 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, a 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 30°C. After 3 hours, the reactions were quenched by heating at 80 °C 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 30°C. At 20 and 40 minute intervals, 20 µL samples were removed, quenched by heating at 80 °C 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 30°C. The reactions were then quenched by heating at 80 °C 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 °C. The reactions were quenched by heating at 80 °C 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 °C. The reaction was quenched by heating at 80 °C 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

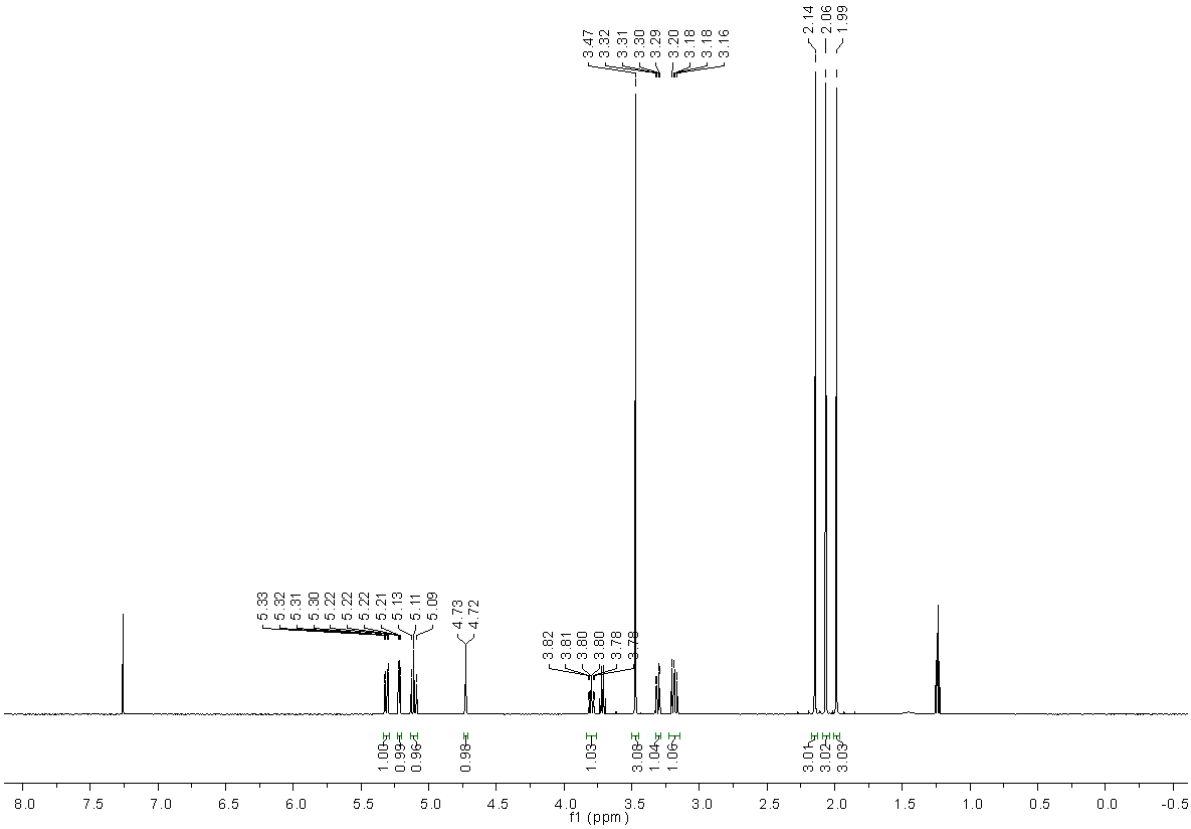
1. Gardberg, A., Abendroth, J., Bhandari, J., Sankaran, B., Staker, B. , Structure of fructose bisphosphate aldolase from *Bartonella henselae* bound to fructose 1,6-bisphosphate. *Acta Crystallogr. Sect. F* **2011,** *67*, 1051--1054.
2. Lorentzen, E.; Siebers, B.; Hensel, R.; Pohl, E., Mechanism of 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.

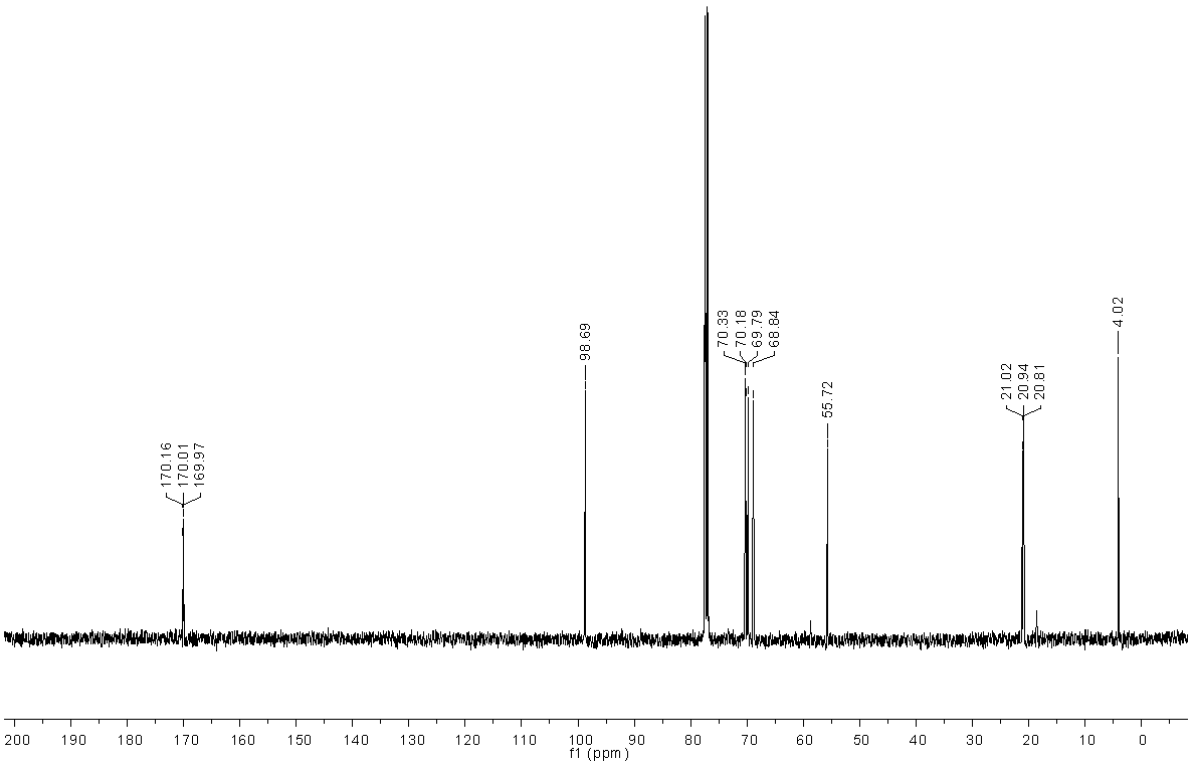
NMR spectra

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

^{1}H NMR

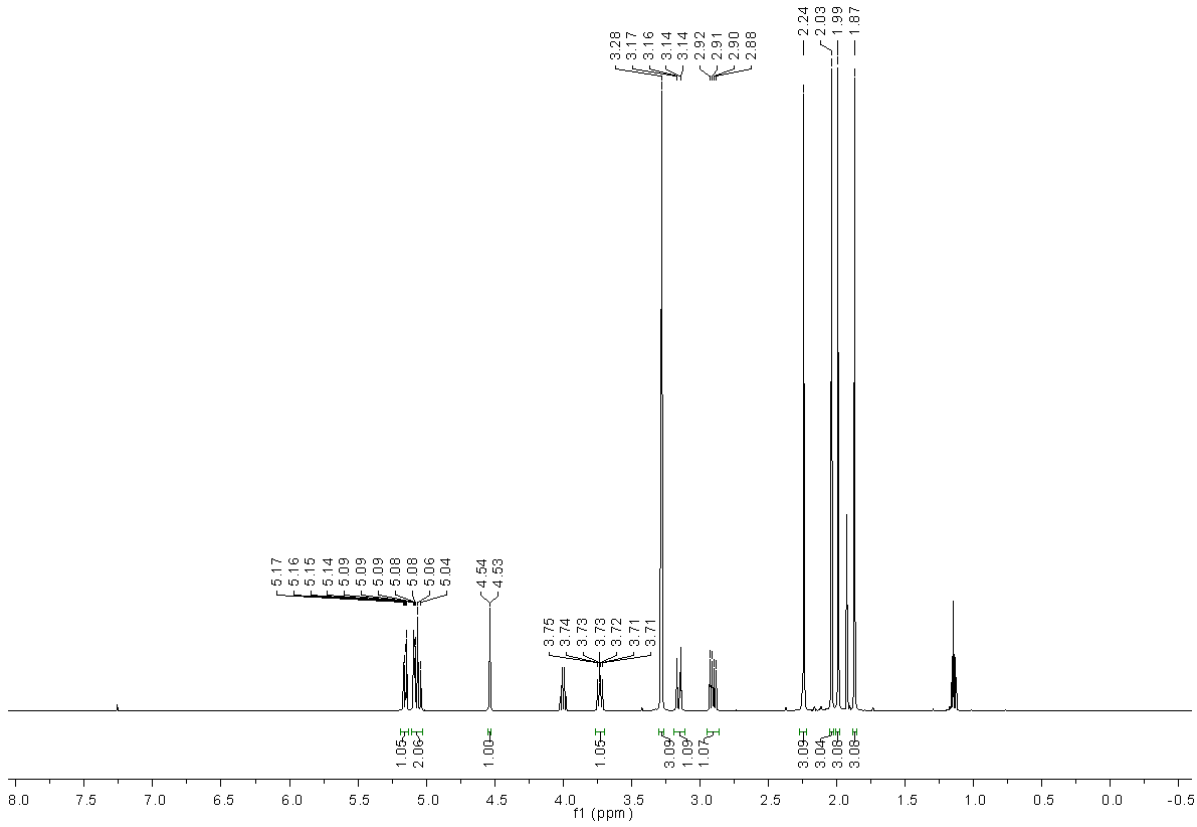


^{13}C NMR

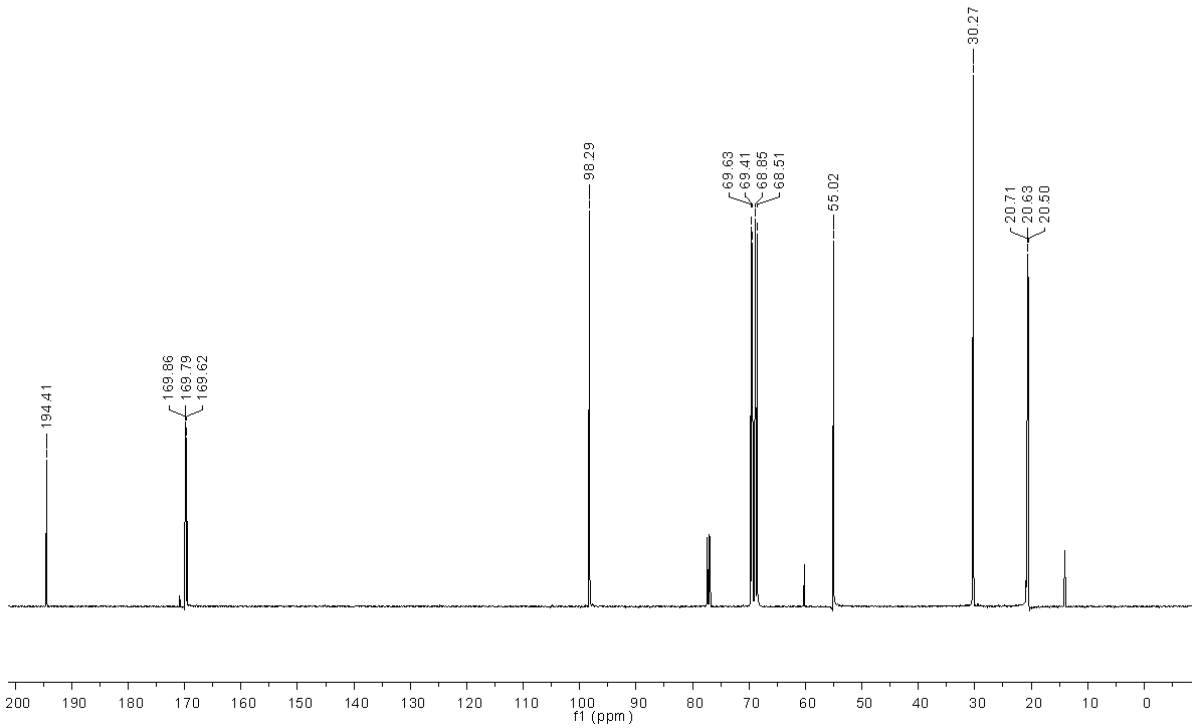


Methyl 2,3,4-tri-*O*-acetyl-6-*S*-acetyl-6-deoxy-6-thio-α-D-mannopyranoside

^{1}H NMR

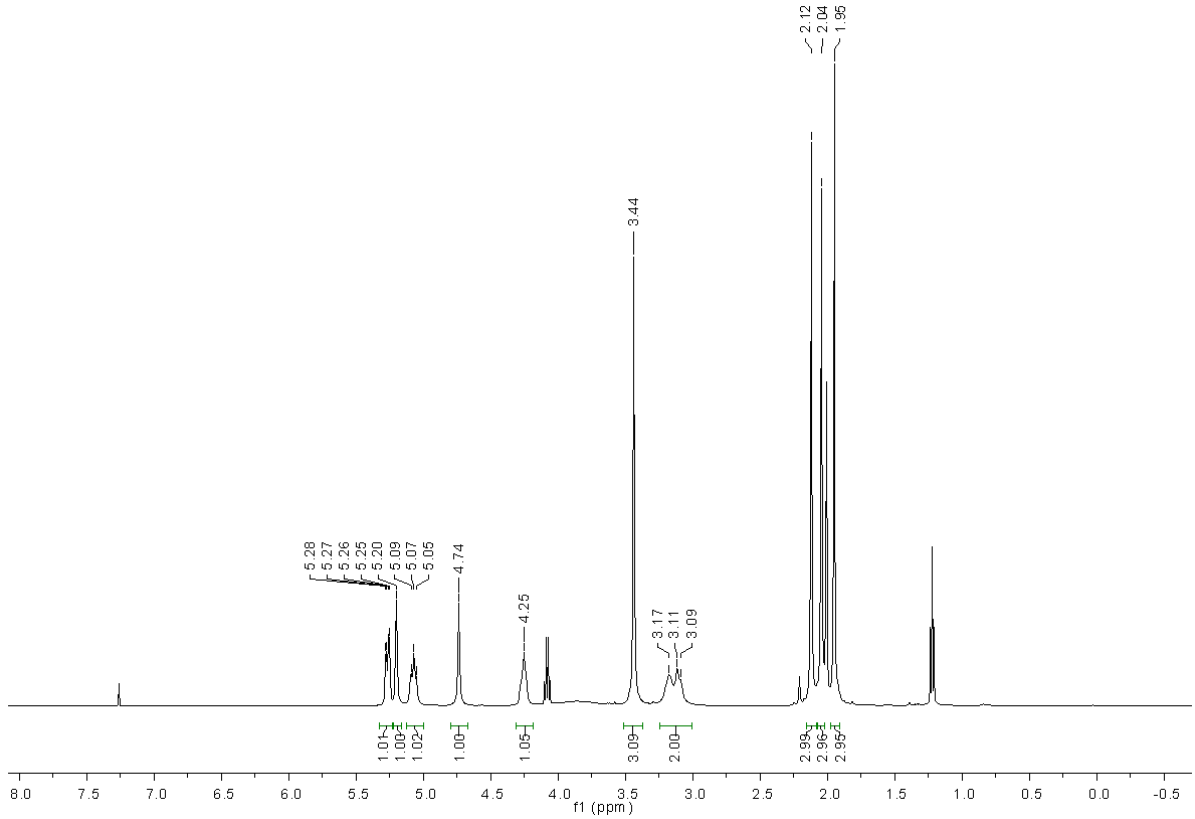


^{13}C NMR

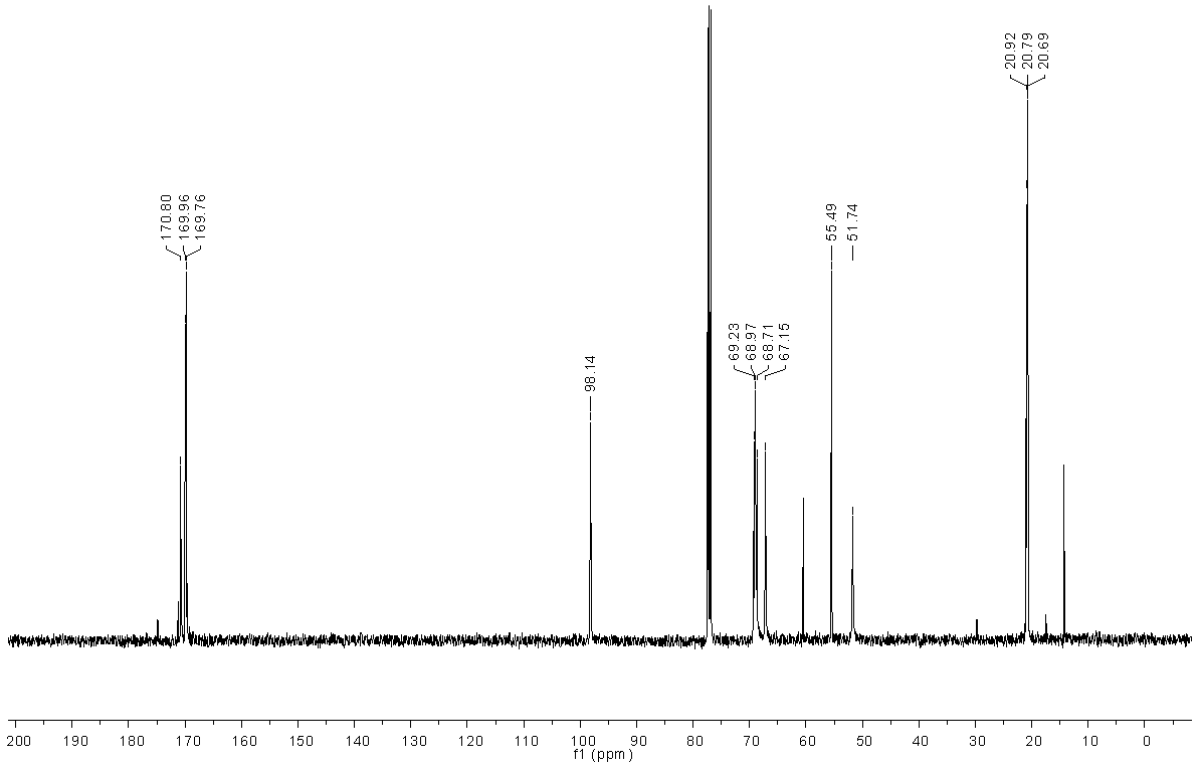


Potassium methyl 2,3,4-tri-*O*-acetyl-6-deoxy-6-sulfonato-α-D-mannopyranoside

^{1}H NMR

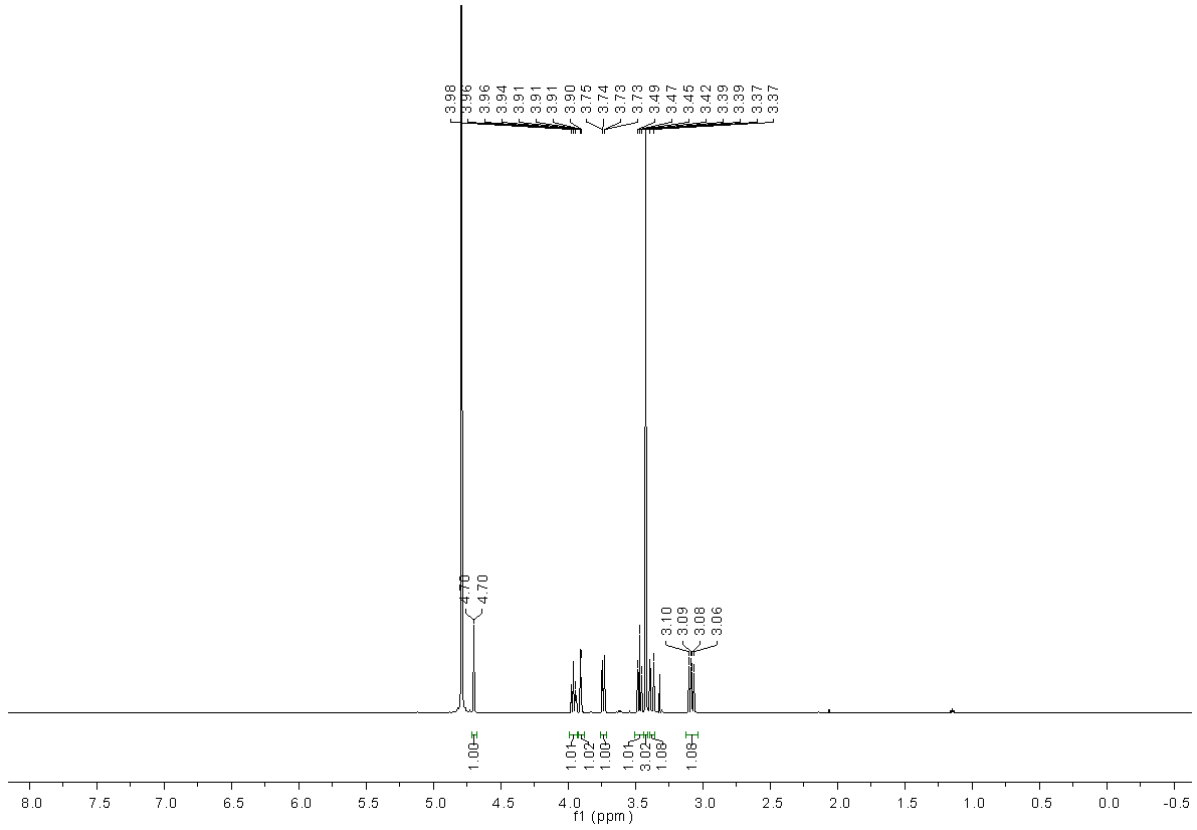


^{13}C NMR

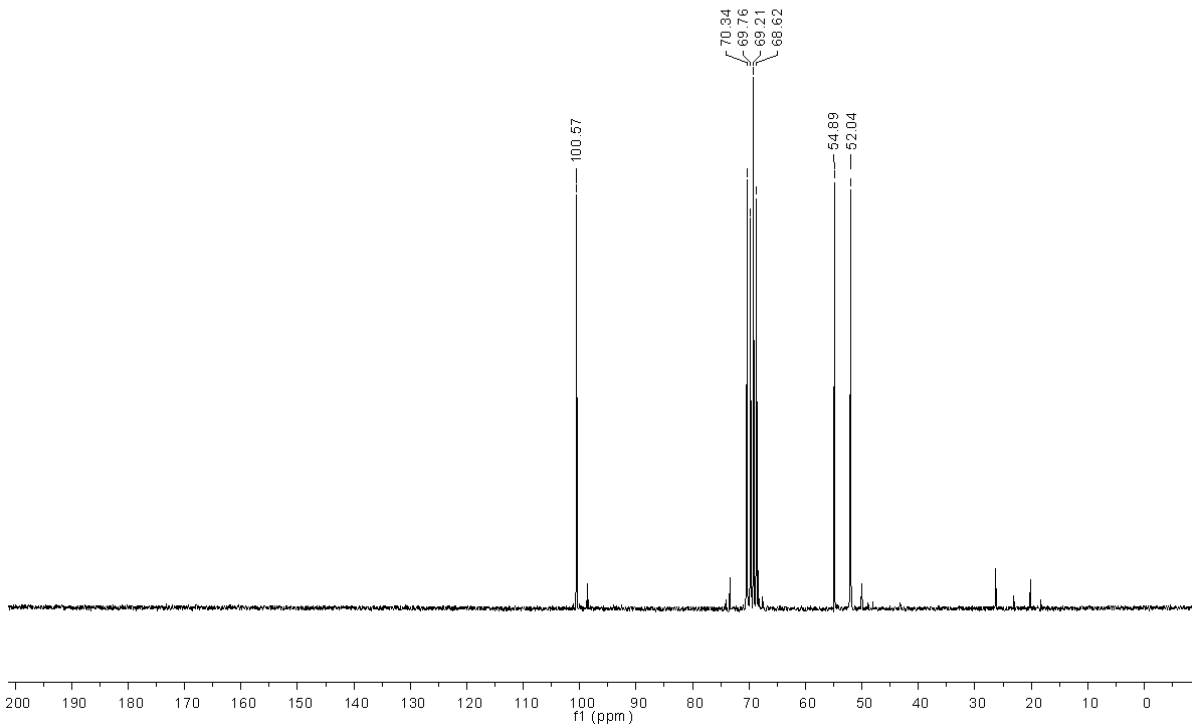


Potassium methyl 6-deoxy-6-sulfonato-α-D-mannopyranoside

^{1}H NMR

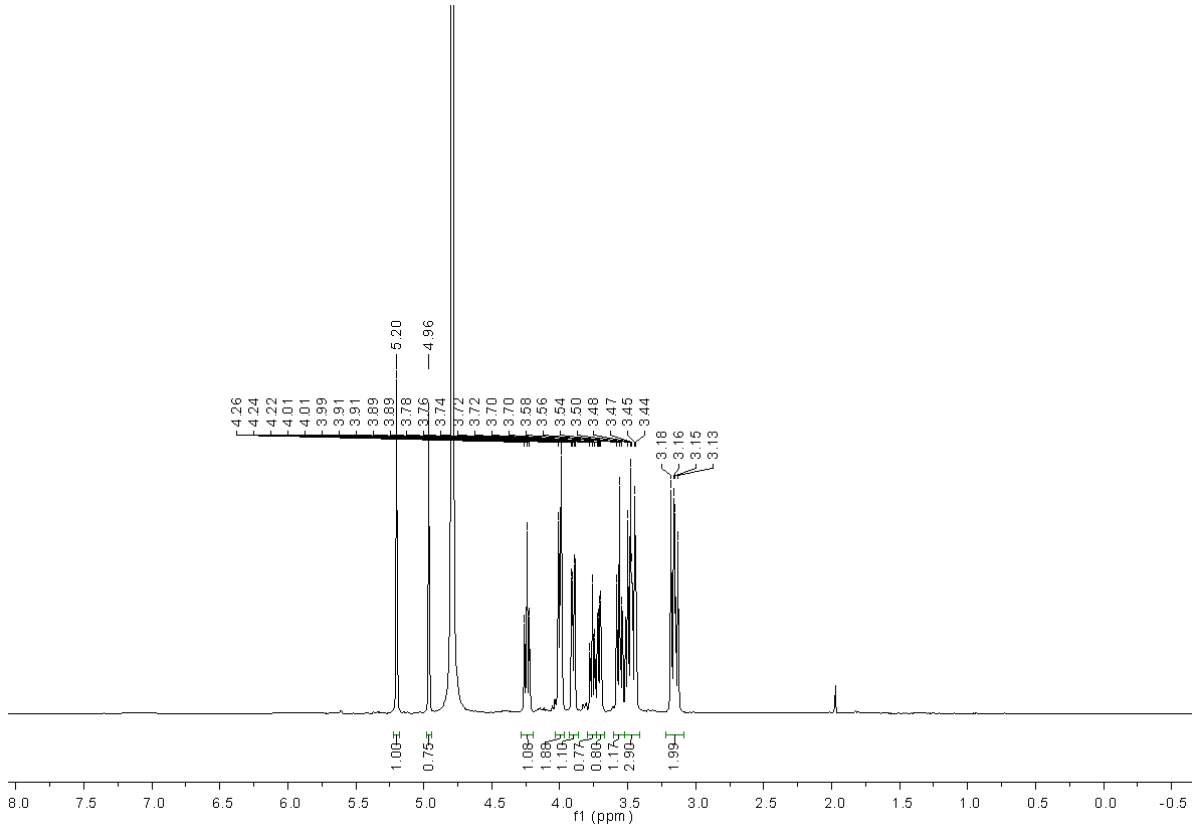


^{13}C NMR

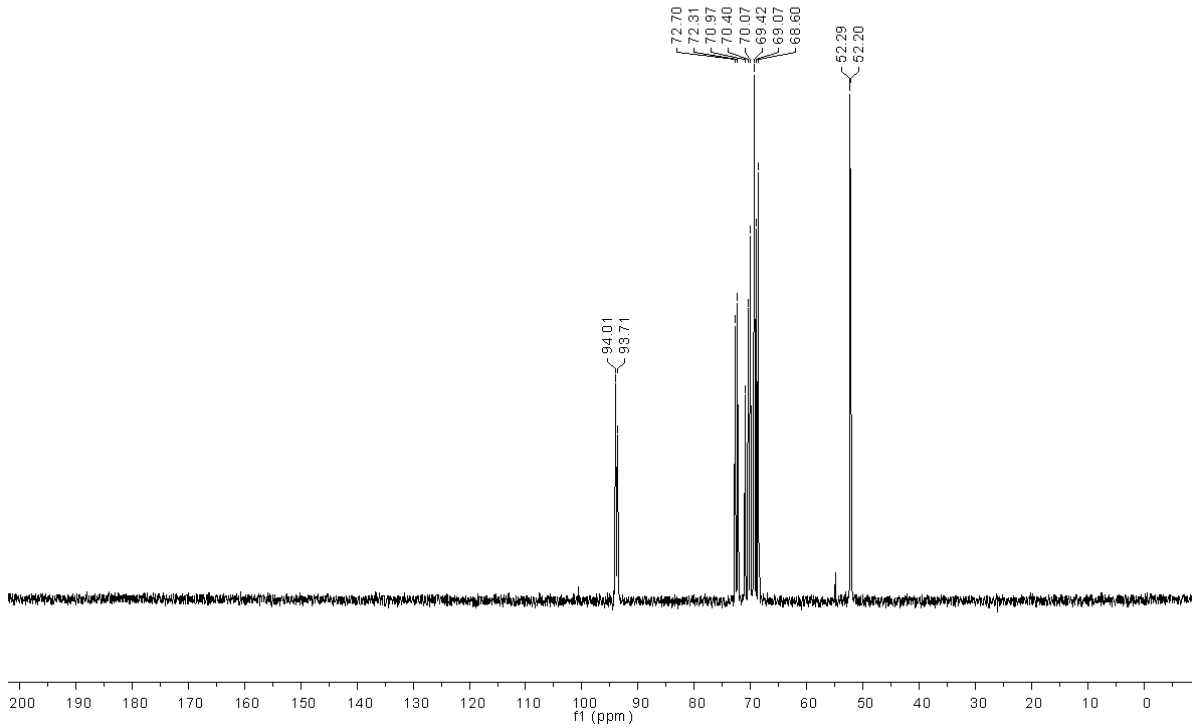


Potassium 6-deoxy-6-sulfonato-D-mannopyranose

^{1}H NMR

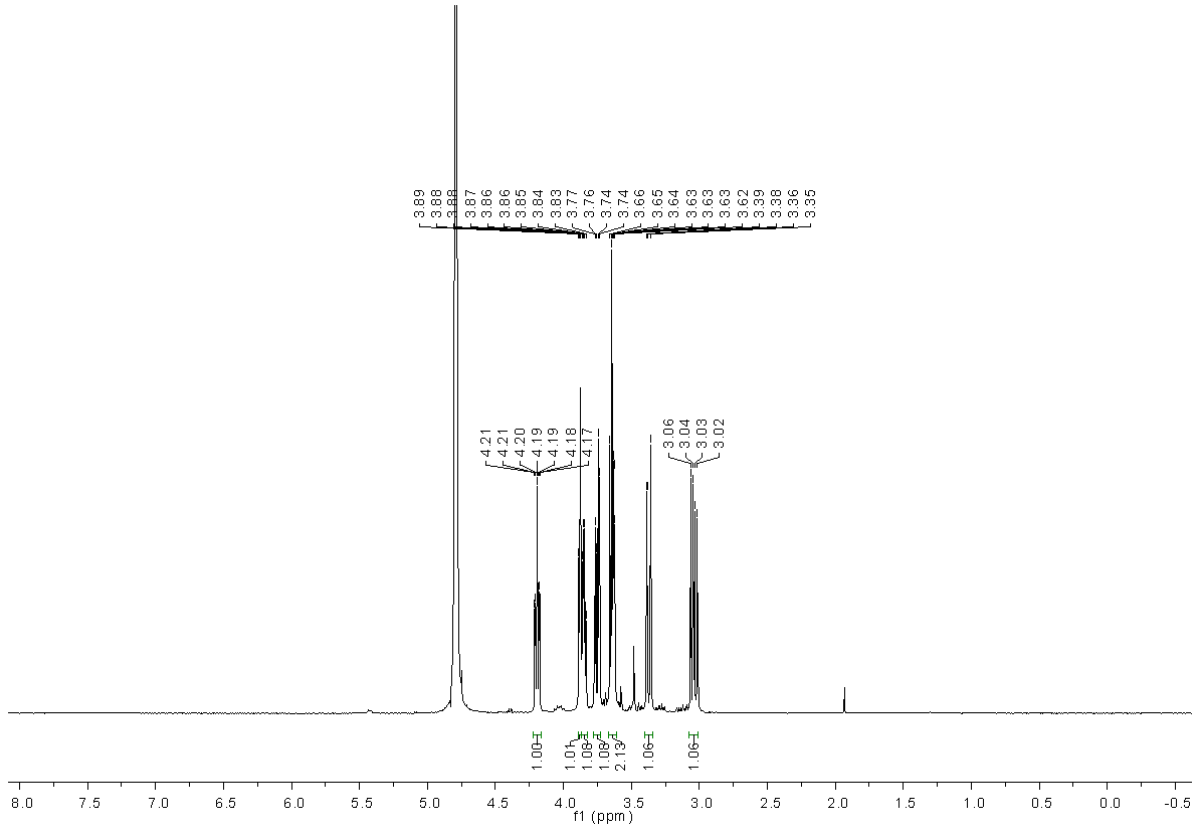


^{13}C NMR



Sorbitol-6-sulfonate, sodium salt

^{1}H NMR



^{13}C NMR