
Article:
Snow, Alexander J.D, Sharma, Mahima orcid.org/0000-0003-3960-2212, Abayakoon, Palika et al. (3 more authors) (2023) Structure and mechanism of sulfofructose transaldolase, a key enzyme in sulfoquinovose metabolism. Structure. pp. 1-9. ISSN 1878-4186

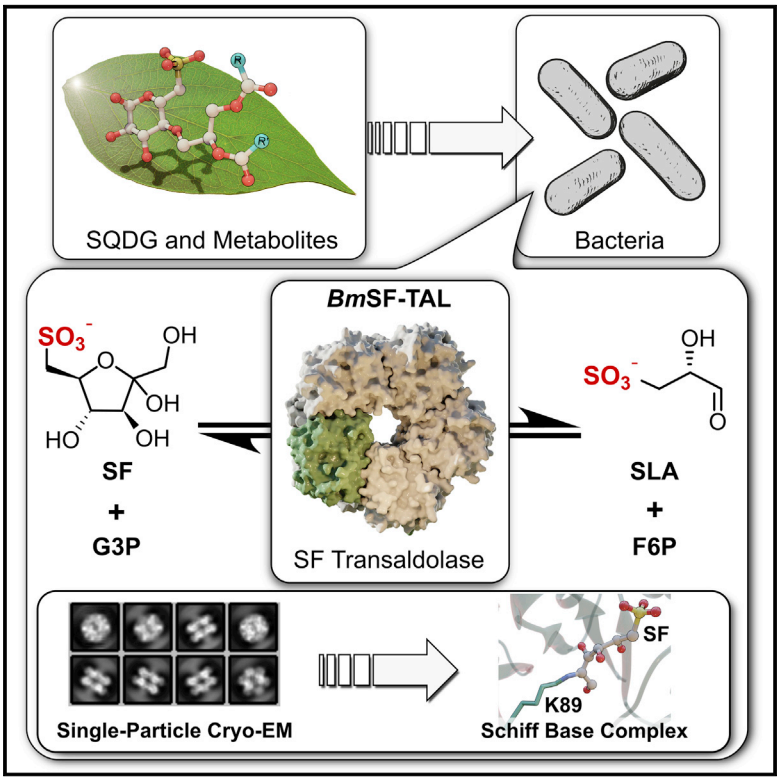
https://doi.org/10.1016/j.str.2023.01.010

Short Article

Structure

Structure and mechanism of sulfofructose transaldolase, a key enzyme in sulfoquinovose metabolism

Graphical abstract



Authors

Alexander J.D. Snow, Mahima Sharma, Palika Abayakoon, Spencer J. Williams, James N. Blaza, Gideon J. Davies

Correspondence

sjwill@unimelb.edu.au (S.J.W.), jamie.blaza@york.ac.uk (J.N.B.), gideon.davies@york.ac.uk (G.J.D.)

In brief

Snow et al. present the high-resolution cryo-EM analysis of sulfofructose transaldolase, a key component of the sulfoglycolytic pathway. The analysis unveils the Schiff-based intermediate complex with sulfofructose and highlights the key residues involved in sulfonate recognition and specificity.

Highlights

- Sulfoquinovose (SQ) accounts for up to half of the global biosulfur cycle
- The sulfoglycolytic sulfofructose transaldolase pathway enables catabolism of SQ
- 2.1-Å cryo-EM structure of the sulfofructose transaldolase Schiff base intermediate
- The catalytic mechanism involves an Asp-Lys-Glu triad

Snow et al., 2023, Structure *31*, 1–9
March 2, 2023
https://doi.org/10.1016/j.str.2023.01.010

CellPress

Please cite this article in press as: Snow et al., Structure and mechanism of sulfofructose transaldolase, a key enzyme in sulfoquinovose metabolism, Structure (2023), https://doi.org/10.1016/j.str.2023.01.010

Structure

CellPress
OPEN ACCESS

Short Article

Structure and mechanism of sulfofructose transaldolase, a key enzyme in sulfoquinovose metabolism

Alexander J.D. Snow,[1] Mahima Sharma,[1] Palika Abayakoon,[2,3] Spencer J. Williams,[2,3,*] James N. Blaza,[1,*] and Gideon J. Davies[1,4,*]

[1]Department of Chemistry, York Structural Biology Laboratory, York YO10 5DD, UK
[2]School of Chemistry, University of Melbourne, Parkville, VIC 3010, Australia
[3]Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, VIC 3010, Australia
[4]Lead contact
*Correspondence: sjwill@unimelb.edu.au (S.J.W.), jamie.blaza@york.ac.uk (J.N.B.), gideon.davies@york.ac.uk (G.J.D.)
https://doi.org/10.1016/j.str.2023.01.010

SUMMARY

Sulfoquinovose (SQ) is a key component of plant sulfolipids (sulfoquinovosyl diacylglycerols) and a major environmental reservoir of biological sulfur. Breakdown of SQ is achieved by bacteria through the pathways of sulfoglycolysis. The sulfoglycolytic sulfofructose transaldolase (sulfo-SFT) pathway is used by gut-resident firmicutes and soil saprophytes. After isomerization of SQ to sulfofructose (SF), the namesake enzyme catalyzes the transaldol reaction of SF transferring dihydroxyacetone to 3C/4C acceptors to give sulfolactaldehyde and fructose-6-phosphate or sedoheptulose-7-phosphate. We report the 3D cryo-EM structure of SF transaldolase from *Bacillus megaterium* in apo and ligand bound forms, revealing a decameric structure formed from two pentameric rings of the protomer. We demonstrate a covalent "Schiff base" intermediate formed by reaction of SF with Lys89 within a conserved Asp-Lys-Glu catalytic triad and defined by an Arg-Trp-Arg sulfonate recognition triad. The structural characterization of the signature enzyme of the sulfo-SFT pathway provides key insights into molecular recognition of the sulfonate group of sulfosugars.

INTRODUCTION

The sulfosugar sulfoquinovose (6-deoxy-6-sulfoglucose, SQ) occurs naturally as the anionic head group of sulfolipid (sulfoquinovosyl diacylglycerols, SQDGs).[1,2] SQDGs are produced by algae, plants, and cyanobacteria and are structural components of the thylakoid membranes of photosynthetic organisms. They intimately associate with the photosynthetic machinery including photosystem II.[3] Because of the ubiquity of SQDG in photosynthetic tissues, it comprises a huge sulfur reservoir. One estimate suggested that SQDG may contain up to half of the organosulfur in the biosphere, with 10 Pg of SQDG produced and degraded annually on a scale that is commensurate with the sulfur-containing amino acids cysteine and methionine.[4] In recent years, a range of microbial pathways responsible for SQDG catabolism, described as "sulfoglycolysis," have been discovered in a wide range of bacteria.[5,6] In pre-sulfoglycolytic processing, SQDG undergoes delipidation through lipase-catalyzed hydrolysis of the acyl chains, forming sulfoquinovosyl glycerol (SQGro).[6,7] SQGro is hydrolyzed to SQ through action of a dedicated family of glycoside hydrolases termed sulfoquinovosidases (SQases).[8,9] Five catabolic routes for SQ have been delineated in Gram-positive and -negative bacteria.[6] These include the sulfoglycolytic Embden-Meyerhof-Parnas (sulfo-EMP) and Entner-Doudoroff (sulfo-ED) pathways, which mirror classical EMP and ED pathways of upper glycolysis, respectively.[10,11] Other pathways have been described that involve specialized enzymes: the sulfoglycolytic sulfofructose transaldolase (sulfo-SFT; also sulfo-TAL) pathway,[12,13] the sulfoglycolytic SQ monooxygenase (sulfo-SMO) pathway,[14] and the sulfoglycolytic sulfofructose transketolase (sulfo-TK) pathway.[5,12]

The sulfo-SFT pathway is used by intestinal firmicutes such as *Clostridium symbiosum* and *Eubacterium rectale*[12] and the saprophytic soil bacterium *Bacillus megaterium* DSM1804.[13] The sulfo-SFT gene cluster from *B. megaterium* is representative and features an SQase that catalyzes hydrolysis of SQ glycosides to give SQ. SQ is acted on by an SQ isomerase to produce sulfofructose (SF). The eponymous SF transaldolase (SF-TAL) catalyzes the transfer of dihydroxyacetone (DHA) from SF to glyceraldehyde-3-phosphate (G3P) or erythrose-4-phosphate (E4P) to produce fructose-6-phosphate (F6P) or sedoheptulose-7-phosphate (S7P) and sulfolactaldehyde (SLA) (Figure 1). SLA may be oxidized to sulfolactate (SL; in *B. megaterium*) or reduced to dihydroxypropanesulfonate (DHPS; in *E. rectale*) and is then exported from the cell, while G3P/E4P are regenerated from F6P/S7P by glycolysis or the pentose phosphate pathway. The excreted C3-organosulfonate SL/DHPS is biomineralized through cross-feeding by other bacteria such as sulfite-respiring and H_2S-producing desulfovibrionaceae species[15] and the human-gut-resident *Bilophila wadsworthia*.[16,17]

Please cite this article in press as: Snow et al., Structure and mechanism of sulfofructose transaldolase, a key enzyme in sulfoquinovose metabolism, Structure (2023), https://doi.org/10.1016/j.str.2023.01.010

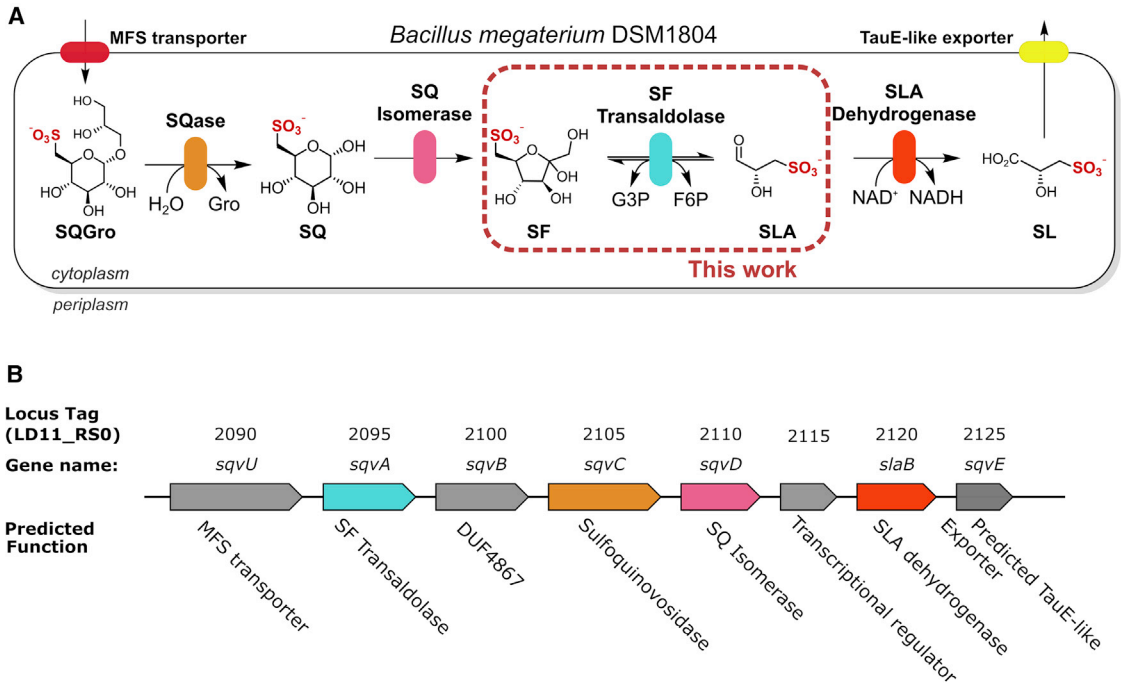


Figure 1. Sulfoglycolytic transaldolase (sulfo-TAL) pathway in *B. megaterium*

(A) SQGro, from delipidation of SQDG, enters the pathway and is hydrolyzed to release SQ, catalyzed by sulfoquinovosidase SqvC. SQ is isomerized to SF by SQ isomerase SqvD. Dihydroxyacetone transfer between SF and G3P is catalyzed by SF transaldolase SqvA (*Bm*SF-TAL), reversibly producing F6P and SLA. SqvA also catalyzes dihydroxyacetone transfer between SF and E4P to give S7P and SLA (not shown). F6P enters upper glycolysis and is phosphorylated to FBP and undergoes aldol cleavage to regenerate G3P. SLA is oxidized to SL by NADH-dependent sulfolactaldehyde dehydrogenase SlaB, producing SL.

(B) Sulfo-SFT gene cluster. SQDG, sulfoquinovosyl diacylglycerol; SQGro, α-sulfoquinovosyl glycerol; SQ, sulfoquinovose; SF, 6-deoxy-6-sulfofructose; SLA, sulfolactaldehyde; SL, sulfolactate; NADH, nicotinamide adenine dinucleotide, reduced form.

The detailed biochemistry of DHPS metabolism involving glycyl radical enzymes in sulfate and sulfite reducing bacteria and fermenting bacteria has been established.[17] The *B. megaterium* sulfo-SFT gene cluster also encodes import/export proteins, a TetR/AcrR family transcriptional factor, and a domain of unknown function (DUF4867) (Figure 1).

The function of SF transaldolase is reminiscent of transaldolases (TALs) such as fructose-6-phosphate transaldolase.[18] F6P transaldolase operates in the non-oxidative phase of the pentose phosphate pathway (PPP), where it catalyzes the reversible transfer of DHA from F6P to E4P to give S7P and G3P. TALs share strong sequence conservation and display a wide variety of oligomeric states[18] (Figure S1). They use a reversible multi-step Schiff base mechanism that is analogous to class I aldolases, which feature a conserved active site Asp-Lys-Glu catalytic triad involved in formation of the Schiff base, orientation of the substrate, and providing general acid/base catalysis.[18,19] In the first step of the TAL mechanism, condensation of the ketohexose substrate with the active site Lys forms a covalent open-chain Schiff base. In the second step, retroaldol reaction cleaves the C3-C4 bond, producing a DHA-derived Schiff base, and releasing G3P. In the third step, aldol reaction of the DHA Schiff base with a C3/C4 acceptor aldehyde generates the product ketose bound as a Schiff base. Finally, hydrolysis releases the product ketose.[20]

Here, we report the structure and mechanism of SF transaldolase from *B. megaterium*, *Bm*SF-TAL (SqvA). Using a stable-isotope-labeled substrate, we demonstrate the ability of *Bm*SF-TAL to catalyze transfer of DHA from SF to G3P or E4P to generate F6P or S7P, respectively. 3D structures determined by cryogenic electron microscopy (cryo-EM) and X-ray crystallography reveal a highly symmetrical decameric structure and provide an intimate view of the active site and define details of sulfonate/phosphate recognition through a Schiff base complex formed with the substrate, SF. This work reveals a general mechanism for SF-TALs that reversibly transfer DHA between sulfonate/phosphate intermediates within the sulfo-SFT pathway.

RESULTS

SF transaldolase catalyzes transfer of the DHA moiety from sulfofructose to G3P and E4P

The gene encoding the *B. megaterium* *Bm*SF-TAL was synthesized in codon-optimized form, heterologously expressed in *Escherichia coli*, and the resulting protein purified to homogeneity (Figure S2). To confirm SF-TAL activity, we examined its ability to catalyze transfer of the DHA moiety from SF to 3C/4C acceptor aldoses (G3P or E4P). For *in vitro* reconstitution of transaldolase activity, *Bm*SF-TAL was incubated with substrates, and the reaction mixtures as well as controls were analyzed using a ZIC-HILIC column detected by ESI-MS in the negative mode and compared with authentic samples. Incubation of *Bm*SF-TAL with SF and G3P resulted in the appearance of a new peak in the chromatogram with retention time of 16.01 min and

Please cite this article in press as: Snow et al., Structure and mechanism of sulfofructose transaldolase, a key enzyme in sulfoquinovose metabolism, Structure (2023), https://doi.org/10.1016/j.str.2023.01.010

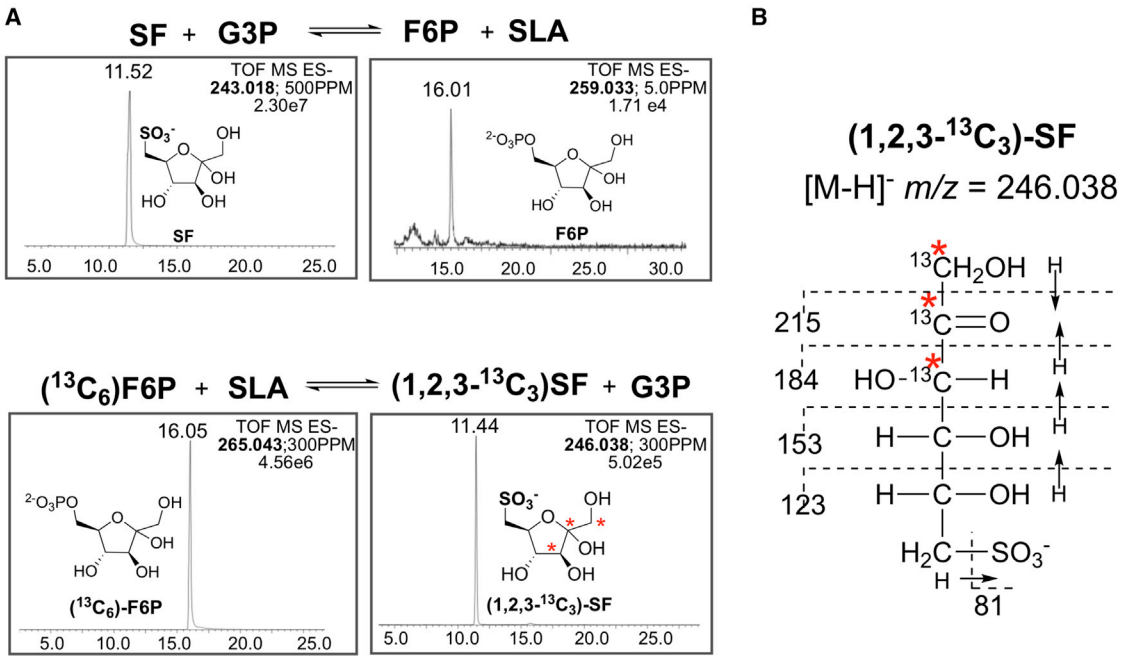


Figure 2. Mass spectrometric analysis (LC-MS/MS) of transaldolase activity of *Bm*SF-TAL on SF and G3P or ($^{13}C_6$)-F6P and SLA

(A) (Top) reaction using SF and G3P as substrates, giving a product assigned as F6P; (bottom) reaction using ($^{13}C_6$)-F6P and SLA as substrates giving a product assigned as (1,2,3-$^{13}C_3$)-SF.

(B) MS/MS analysis of (1,2,3-$^{13}C_3$)-SF.

m/z = 259.033, corresponding to F6P (Figure 2A). To demonstrate the reversibility of the reaction, we incubated stable-isotope-labeled ($^{13}C_6$)-F6P with chemically synthesized SLA[21] in the presence of *Bm*SF-TAL, which resulted in the formation of a product with a retention time of 11.44 min and *m/z* = 246.038 corresponding to (1,2,3-$^{13}C_3$)-SF, indicating transfer of the ^{13}C-labeled DHA moiety derived from carbons 1–3 of F6P to SLA. MS/MS analysis of the (1,2,3-$^{13}C_3$)-SF formed revealed fragments supporting the proposed labeling pattern (Figures 2B and S3).

Schleheck and co-workers have demonstrated that the SF-TAL from *Bacillus aryabhattai* SOS1 catalyzes transfer of the DHA moiety from SF to E4P resulting in the formation of the ketoheptose S7P. To examine if the *B. megaterium* enzyme could also transfer to a four-carbon acceptor, we incubated *Bm*SF-TAL with SF and E4P, which resulted in formation of SLA and a new peak with *m/z* = 289.043 corresponding to S7P. Similarly, incubation of *Bm*SF-TAL with S7P and SLA revealed formation of SF and E4P, demonstrating the ability of *Bm*SF-TAL to catalyze the reverse reaction. The identity of the sulfonate product SF was confirmed by ESI-MS and MS/MS analyses (Figure S4). Satisfied that the *B. megaterium* enzyme was a bonafide SF-TAL, we next sought to determine its 3D structure.

B. megaterium SF transaldolase is a decamer composed of two cyclic pentamers

Size-exclusion chromatography-multi-angle laser light scattering (SEC-MALLS) analysis of *Bm*SF-TAL revealed that it exists as a decamer with a molecular weight of ~250 kDa in solution, making it a suitable candidate for structural analysis using cryo-EM methods (Figure S5).

Samples for EM analysis were prepared using gold grids with a protein concentration of 2.5 mg mL^{-1} (seeStar Methods). Data collection utilized a Glacios 200 kV microscope while image processing and 3D reconstruction were performed using Relion v.3.0.[22] Early 2D and 3D reconstructions of the ligand-free structure showed a clear pair of pentameric rings, facing away from one another and offset by about 36° from one another on the *z* axis (Figure S6). This allowed assignment to symmetry group D5. Combined with the high contrast and particle count in the micrographs, this allowed for reconstruction of the ligand-free complex to 2.6 Å.

The *Bm*SF-TAL decamer assembles as a dimer-of-pentamers (Figures 3A and 3B). The *Bm*SF-TAL monomer adopts a triose phosphate isomerase $(\alpha\beta)_8$ barrel fold. This is adorned with an extended C-terminal α-helix, which is observed in all TALs, and which projects from the main fold in an extended conformation in a manner consistent with other decameric TALs[23] (Figure 3C). Pentamer formation involves the C-terminal α-helix on one monomer inserting into the adjacent monomer within each pentamer to form a ring (Figure 3C). The adjacent *Bm*SF-TAL monomer contains a complementary hydrophobic groove along the top of the $(\alpha\beta)_8$ barrel that accommodates the C-terminal α-helix (Figure 3D). The C-terminal α-helix is amphipathic, with the hydrophobic face oriented inward and the hydrophilic face oriented toward the solvent. This interaction places the C-terminal α-helix near the active site; however, no residue from it directly interacts with the substrate (*vide infra*). The interaction of the C-terminal α-helix with the adjacent monomer involves hydrophobic interfaces along helix α2 and sheet β6. Consistent with the 5-fold symmetry, the angle

Please cite this article in press as: Snow et al., Structure and mechanism of sulfofructose transaldolase, a key enzyme in sulfoquinovose metabolism, Structure (2023), https://doi.org/10.1016/j.str.2023.01.010

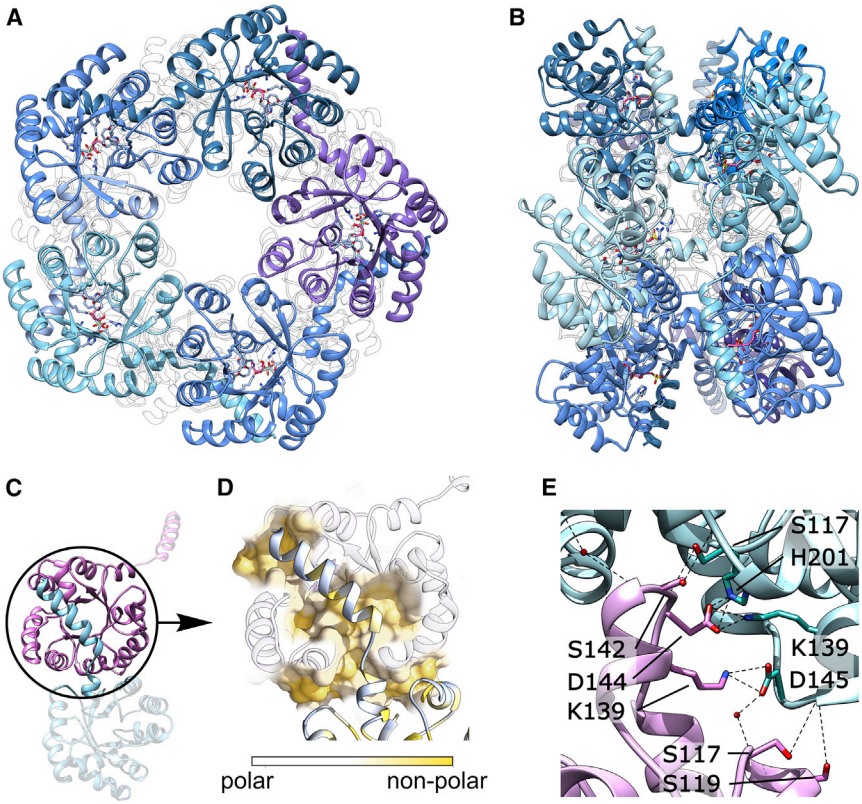


Figure 3. Overall structure and quaternary organization of *B. megaterium* SF transaldolase

(A) Front and (B) side view of "dimer-of-pentamers." In both cases the cryo-EM structure of BmSF-TAL complexed with SF is shown. SF is in pink, and the BmSF-TAL monomers are shaded to highlight each separately.

(C) C-terminal α-helix interaction between two BmSF-TAL monomers (purple, cyan).

(D) Heatmap of C-terminal α-helix/groove interaction. Non-polar surfaces are depicted in yellow.

(E) View of the BmSF-TAL apo crystal structure showing ordered water molecules and hydrogen bonding interactions between the two pentameric rings.

between the interface in adjacent monomers is 72°, allowing *Bm*SF-TAL to form a pentameric ring (Figures 3A and 3D). The decamer is formed from two pentameric rings that associate by facing toward one another, with a *z* axis rotation of 36°. Interactions between to the two pentamers involve mainly polar amino acid side chains of helix α7 on each monomer. These face inward and form a network of hydrogen bonds to an acceptor region between helices α7 and strand β6 and altogether result in a total of 10 α7/α7β6 interactions.

Small conformational variations in protomers within the decameric complex are averaged out in cryo-EM reconstructions, and unless the changes are large enough to contribute to a signal above the noise in the particles, these changes cannot be classified through symmetry expansion approaches. To gain insight into possible conformational differences that may exist in each protomer, we solved the ligand-free X-ray crystal structure of *Bm*SF-TAL in space group P 2_1, which refined to an R/R_{free} of 0.23/0.28, at 3.2 Å resolution (Table S1). This structure contains one decamer in the asymmetric unit and contains a range of monomer-specific differences such as an additional six C-terminal residues of two protomers. Notably, strand β3, which harbors catalytic residue Glu61, shows greater conformational variation across individual protomers compared with the rest of the protein fold (Figure S7A). Further, two arginine residues, Arg30 and Arg172, which line the substrate pocket, inhabit a range of conformations in the absence of ligand. This flexibility may account for the sulfonate-phosphate dual-specificity of the pocket that accommodates SF or F6P (Figure S7B).

The SF transaldolase forms a Schiff base adduct with sulfofructose

To define an approach to obtain a protein-ligand complex, we applied a thermal shift assay. Using nano differential scanning fluorimetry, we observed a large increase in melting temperature (ΔTm of 5.6°C) of *Bm*SF-TAL in the presence of SF. Armed with this insight, we incubated *Bm*SF-TAL with SF (2 mM), and subsequent grid preparation and optimization resulted in a cryo-EM structure at resolutions up to 2.1 Å (Figures S8 and S10). SF-bound maps showed clear and well-resolved side-chain density and a continuous and traceable series of peptide backbones for each monomer. Comparison of active site features with well-characterized F6P transaldolases from *E. coli Ec*TAL-B (PDB: 4S2C) or from *T. acidophilum* (*Tac*TAL, PDB: 3S1V) and sequence alignment revealed a conserved catalytic triad of Asp6-Lys89-Glu61 within the barrel of each monomer (Figure 4A).

Reconstruction and model building of *Bm*SF-TAL with open-chain SF showed a clear extension of density beyond Nε of Lys89 and included a well-defined six-carbon chain and six-sulfonate head group (Figures 4A and 4B). Contiguous density from Lys89 to the C2 carbon of SF enabled an iminium adduct to be modeled. When reprocessed without application of symmetry, SF was present in every binding site. The density allowed modeling of the entire SF chain in 9 of 10 monomers with weaker density in the remaining protomer. The average B factors for *Bm*SF-TAL·SF increased from 56 $Å^2$ in all chains and 30 $Å^2$ for ligands in D5 symmetry to an average 65 $Å^2$ and 69 $Å^2$, respectively, in C1 symmetry ($\Delta B_{chain} = 9\ Å^2$, $\Delta B_{ligand} = 39\ Å^2$). Otherwise, there is little conformational change between ligand-free and SF-bound forms (RMSD = 0.1 Å; Figure S11).

SF is covalently attached to Lys89 in a polar active site and forms a range of hydrogen bonding interactions in the substrate binding pocket. C1 hydroxyl is H-bonded to Ser133 and Asn111 at distances of 2.7 Å and 3.0 Å, respectively. The C3- and C5-hydroxyls of SF H-bond with the side-chain carboxylate of Asp6 at distances of 2.5 Å and 2.5 Å, respectively. The C4 hydroxyl interacts with the carboxylate of Glu61, via an ordered

Please cite this article in press as: Snow et al., Structure and mechanism of sulfofructose transaldolase, a key enzyme in sulfoquinovose metabolism, Structure (2023), https://doi.org/10.1016/j.str.2023.01.010

CellPress
OPEN ACCESS

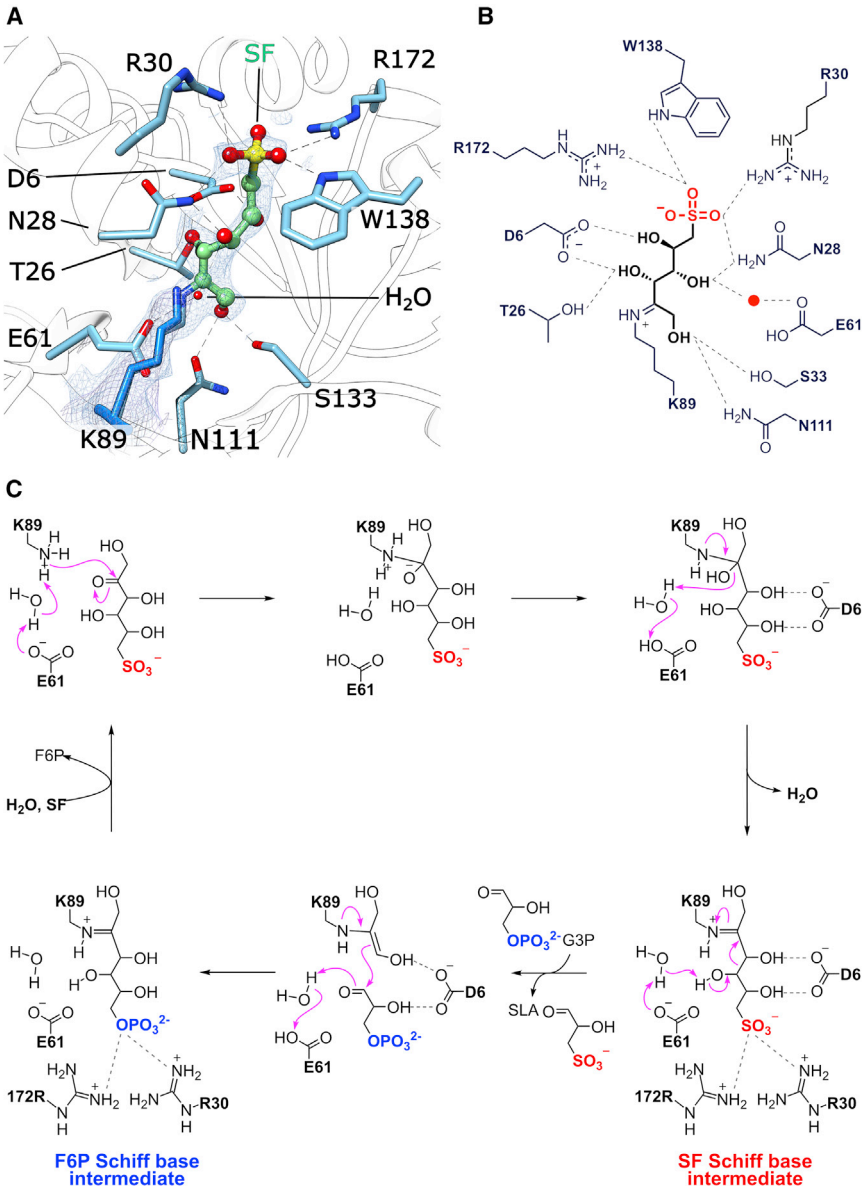


Figure 4. Structural basis of SF recognition by BmSF-TAL and proposed reaction mechanism

(A) Active site details of SF-bound BmSF-TAL. SF is in green and yellow. Electron density in both cases is contoured to 1.0 σ (0.29 e/Å^3).

(B) Binding site schematic of BmSF-TAL with SF bound.

(C) Mechanism of SF transaldolase depicting catalytic triad D6-E61-K89 and sulfonate binding arginines: R30, R172.

water at a distance of 3.2 Å (Figures 4A and 4B). The anionic sulfonate of SF is recognized by an Arg30-Trp138-Arg172 triad. One sulfonate oxygen is engaged in an H-bond with the indole amine of Trp138 (2.8 Å) and Arg172 (3.0 Å), and a second sulfonate oxygen interacts with Arg30 and Asn28 at 2.7Å and 2.9 Å, respectively (Figures 4A and 4B). The binding mode of the hexose donor is similar to that seen in complexes of classical TALs like F6P transaldolase from *E. coli* *Ec*TAL-B or *Tac*TAL with F6P, with the exception of the sulfonate binding interactions.

Even though sequence and structural comparison of *Bm*SF-TAL with other bacterial F6P transaldolases reveals a high degree of sequence and spatial conservation in the residues lining the active site, there are key differences in the phosphate/sulfonate binding pocket. Examination of *Bm*SF-TAL·SF complex versus TAL·F6P complexes (structural homologs were identified using a DALI search) shows that the phosphate moiety of the substrate within classical TALs is recognized by a conserved Arg135-Ser167-Arg169 motif (*Tac*TAL numbering) (Figure S12). However, this RSR motif is not conserved in *Bm*SF-TAL: the residue corresponding to *Tac*TAL Arg135 in *Bm*SF-TAL is replaced by tryptophan (Trp138), and directly across the active site, a serine residue that projects into the active site of *Tac*TAL is replaced by Arg30, which in turn recognizes the sulfosugar. Arg172 in *Bm*SF-TAL (corresponding to Arg160 in *Tac*TAL) is conserved and makes electrostatic interactions with SF, and these three residues result in a unique Arg30-Trp138-Arg172 "RWR" sulfonate binding pocket for SF transaldolases (Figures S13 and S14). The conserved Ser167 (*Tac*TAL) seen in classical TALs is replaced by an Ala residue in the equivalent position in *Bm*SF-TAL (Ser170). Zhao and co-workers applied sequence similarity network analysis to 1,000 close homologs and proposed a Trp138-Ala170-Arg172 "WAR" sulfonate-binding motif for *Bm*SF-TAL, but they did not identify the Arg30 residue we identified by structural methods.[13]

Because *Bm*SF-TAL also binds F6P (as product; or substrate for the reverse reaction) the RWR triad within the sulfonate binding pocket must therefore also bind phosphate. We propose that the phosphate binds in the same pocket as the sulfonate groups, and that the side-chain rearrangements of two arginine residues observed in individual monomers in the ligand-free 3D X-ray structure in this substrate binding pocket allow binding of F6P as the phosphate dianion. Our activity assays show that *Bm*SF-TAL not only binds six-carbon donors (SF or F6P) but also S7P, which shows that a seven-carbon skeleton is accommodated within the active site. Despite multiple attempts, we could not trap complexes of *Bm*SF-TAL with F6P, S7P, G3P, or SLA. An overlay of *Bm*SF-TAL complexed with sulfonate substrate SF versus the complexes of *Tac*TAL with bound F6P and S7P indicates rearrangements of the flexible arginine side chains (Figure S15). These side-chain rearrangements of Arg residues in the substrate binding pocket allow provision of steric and electrostatic

Please cite this article in press as: Snow et al., Structure and mechanism of sulfofructose transaldolase, a key enzyme in sulfoquinovose metabolism, Structure (2023), https://doi.org/10.1016/j.str.2023.01.010

Table 1. Data collection, refinement, and validation statistics for cryo-EM datasets of *Bm*SF-TAL

	*Bm*SF-TAL apo	*Bm*SF-TAL·SF (D5)	*Bm*SF-TAL·SF (C1)
Data collection and processing			
Magnification	240,000	310,000	310,000
Voltage (kV)	200	200	200
Electron fluence (e^{-}/Å^2)	50	50	50
Pixel size (Å^2)	0.574	0.450	0.450
Symmetry	D5	D5	C1
Initial particles	153,063	77,489	77,489
Final particles	59,063	53,450	53,450
Map resolution (Å)	2.7	2.1	2.7
FSC threshold	0.143	0.143	0.143
Map resolution range (Å)	34–2.4	44–1.9	31–2.5
Refinement			
Initial model	N/D	*Bm*SF-TAL apo	*Bm*SF-TAL·SF (D5)
Map sharpening *B* factor (Å^2)	−67	−39	−44
Model composition			
Non-H atoms	16,770	17,399	16,448
Residues	2,180	2,180	2,160
Ligands	0	10	9
Water	460	500	273
Mean *B* factors (Å^2)			
Protein	84.2	60.7	63.6
Ligand	N/D	113.1	116.6
Water	82.8	71.1	62.2
RMS deviations			
Bond length (Å)	0.011	0.014	0.014
Bond angles (°)	1.634	2.080	2.143
Validation			
MolProbity score	2.2	1.9	2.3
Clashscore	2.73	1.6	4.1
Poor rotamers %	0	0	0
Ramachandran plot			
Favored %	95.83	96.76	95.82
Allowed %	4.17	3.24	4.13
Outliers %	0	0	0.05
PDB code	8BC2	8BC3	8BC4

complementarity to accommodate C6/C7 ketoses as well as the sulfonate/phosphate dual-specificity observed in *Bm*SF-TAL.

*Bm*SF-TAL uses a Schiff base mechanism

Identification of the active site machinery in the *Bm*SF-TAL·SF complex and the observation of a covalent Schiff base with the acyclic SF molecule allows proposal of a reaction mechanism for the transaldolase reaction. Like other classical transaldolases, *Bm*SF-TAL contains a conserved Asp6-Lys81-Glu61 catalytic triad. Formation of the Schiff base complex with SF will involve nucleophilic attack by Nε of Lys89 on the ketone of acyclic SF, assisted by general base Glu61, via a carbinolamine intermediate. Hydrogen bonding of Asp6 to the C3 and C5 hydroxyls helps to ensure the correct orientation of the extended chain of the donor ketose. Retroaldol of the Schiff base occurs by Glu61 acting as general base to deprotonate the C4-OH via a bridging water molecule, triggering the scission of the C3-C4 bond, releasing SLA, and generating a new Schiff base of Lys89 (as enamine tautomer) with dihydroxyacetone. DHA remains anchored as a Schiff base with Lys89 allowing aldol condensation and transfer to the incoming acceptor G3P (or E4P) through the reverse process, followed by the hydrolysis of the Schiff base and release of the transaldol product (Figure S4C). Arg30 and Arg172 are involved in sulfonate/phosphate recognition and provide the ability to bind mono- and dianionic species. Broadly, this mechanism is conserved with that proposed for F6P transaldolase.[24,25]

DISCUSSION

*Bm*SF-TAL is a transaldolase that reversibly transfers DHA between sulfofructose and G3P (or E4P) to give SLA (or S7P). F6P can enter the PPP or central carbon metabolism (via lower glycolysis), or be converted to aminosugars supporting cell wall biogenesis, or to G1P for generation of storage polysaccharides such as bacterial glycogen. Through the action of F6P transaldolase, S7P can be converted to ribose 5-phosphate (R5P), where it can be used in the synthesis of nucleotides and nucleic acids and regenerate F6P. Thus, *Bm*SF-TAL provides a pathway for direct injection of a C3 fragment from the sulfosugar SQ into key pathways in hexose/heptose metabolism.

Classical TALs exhibit a variety of multimeric states, including monomers, dimers, tetramers to decameric assemblies.[18] *Bm*SF-TAL has a decameric structure composed of two cyclic pentamers. This is similar to the decameric assembly observed for F6P transaldolases from several species including *T. acidophilum* (PDB: 3S1V). The *Bm*SF-TAL cyclic pentamers are formed by interposition of the C-terminal α-helix of one monomer into a well-defined binding cleft in the next, granting the overall homodecameric assembly high thermodynamic stability, as reflected in a melting temperature of 80°C.

The structure of the *Bm*SF-TAL·SF complex shows a Schiff base adduct formed with the active site Lys89 and highlights the involvement of Asp6 in ensuring correct orientation of the extended sugar chain, while Glu61 acts as general acid-base in the catalytic mechanism. The conservation of this catalytic triad (Asp6-Glu61-Lys89) with F6P transaldolases supports a common Schiff base retroaldol/aldol mechanism. This similarity extends to the observation of an ordered water in the active site, as seen in *Tac*TAL (PDB: 3S1V).

*Bm*SF-TAL performs a step analogous to F6P transaldolase within the oxidative branch of the PPP, which can interconvert S7P and G3P with F6P and E4P through transfer of a dihydroxyacetone group. However, F6P transaldolases only bind phosphorylated substrates/products, whereas *Bm*SF-TAL can bind phosphate and sulfonate groups. The high specificity of phosphorylated substrates for F6P transaldolases is due to a

conserved RSR motif, as observed in several complexes with F6P (e.g., 3S1V, 4S2C, 3TKF) or S7P (3TNO, 3S1U).

SF sulfonate binding interactions occur with the triad Arg30-Trp138-Arg172, which is conserved among a range of SF-TALs. The two arginine residues present within this RWR motif may allow both monoanionic sulfonate and dianionic phosphate residues to bind. Within individual monomers of the apo form of *Bm*SF-TAL, some conformational mobility of these residues was observed. This conformational mobility may allow accommodation of the difference in chain lengths imposed by switching between a sulfonate ($-SO_3^-$) and a phosphate ($-OPO_3^{2-}$) in hexoketoses SF and F6P, as well as the additional carbon in the heptoketose S7P. Overlays of the structures of the F6P and S7P complexes of F6P transaldolases with the *Bm*SF-TAL·SF structure suggests that these chain length differences may be accomodated through a combination of conformational flexibility from Arg30 and Arg172 and by rotation of the phosphate relative to the sulfonate.

The sulfo-EMP pathway contains an aldolase (YihT in *E. coli* and *Salmonella enterica*), which catalyzes the retroaldol reaction of SF-1-phosphate to give DHAP and SLA.[26] Like *Bm*SF-TAL, SFP aldolase operates through a class I aldolase mechanism, involving a Schiff base formed between an active site lysine with SFP. Upon retroaldol cleavage and release of SLA, this leads to a DHAP moiety bound as a Schiff base, which hydrolyzes to release free DHAP. Kinetic studies reveal that *S. enterica* YihT is highly selective for SFP, with no detectible activity on FBP. 3D X-ray structures of Schiff base complexes of *S. enterica* YihT in complex with SFP and DHAP have been reported (PDB: 7NE2). Key differences between *S. enterica* SFP aldolase and *Bm*SF-TAL include the following: the Schiff base is formed with DHAP in SFP aldolase rather than DHA in *Bm*SF-TAL, and selectivity for SFP over FBP for YihT, versus activity on both SF and F6P in *Bm*SF-TAL. The sulfonate binding pocket of YihT contains only one Arg group, while that of *Bm*SF-TAL contains two, suggesting that the selectivity for a sulfonate substrate in the former, and dual activity on sulfo/phosphosugars in the latter may be driven by these residues.

Sulfoglycolytic pathways feature sulfonated intermediates (arising from catabolism of SQ), and consequently these pathways usually contain enzymes with specific sulfonate binding pockets. By contrast, SF-TAL, the namesake enzyme of the sulfo-SFT pathway, acts on both sulfonate and phosphate substrates. Our structural studies reveal a conserved "RWR" motif that we propose imparts dual phosphate/sulfonate binding capabilities to this central enzyme of the sulfo-SFT pathway and distinguishes SF-TALs from F6P TALs. Elucidation of the signature sulfonate-binding motif in the SF-TAL enzymes will enhance the power of bioinformatic annotation, support the discovery of new organisms that use this pathway, and help illuminate the occurrence and distribution of the sulfo-SFT pathway for SQ catabolism.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - BL21(DE3) competent E. coli
- METHOD DETAILS
  - Gene expression and protein purification
  - SEC-MALLS analysis
  - NanoDSF analysis
  - LC-MS detection of substrates and products of transaldolase activity
  - Electron cryomicroscopy (cryo-EM)
  - Bioinformatics
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.str.2023.01.010.

ACKNOWLEDGMENTS

This work was supported by the Leverhulme Trust (RPG-2017-190), Biotechnology and Biological Sciences Research Council (BB/W003805/1), the Australian Research Council (DP210100233, DP210100233), the UKRI Future Leader Fellowship Program (MR/T040742/1), and the Royal Society for the Ken Murray Research Professorship to G.J.D. We thank the Wellcome Trust for funding the Glacios electron microscope (grant number 206161/Z/17/Z). We acknowledge Diamond Light Source for access and support of beamline i-03 (under proposal mx24948), and Dr Johan Turkenburg and Sam Hart for assistance with XRD data collection. We also acknowledge Dr. Andrew Leech at the University of York Bioscience Technology Facility for assistance with SEC-MALLs, and the Metabolomics & Proteomics Laboratory, Department of Biology, University of York for MS facilities and data acquisition.

AUTHOR CONTRIBUTIONS

G.J.D. and S.J.W. conceived the project. M.S. and A.J.D.S. designed the plasmid construct and over-expressed and purified *Bm*SF-TAL. A.J.D.S. prepared the cryo-EM grids and performed the image acquisition and processing. P.A. synthesized SF and C3-sulfonate intermediates. M.S. carried out activity tests and LC-MS analysis. J.N.B. helped with the image acquisition, processing, and model building. A.J.D.S., M.S., S.J.W., and G.J.D. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: October 25, 2022
Revised: December 15, 2022
Accepted: January 25, 2023
Published: February 17, 2023

REFERENCES

1. Benson, A.A., Daniel, H., and Wiser, R. (1959). A sulfolipid in plants. Proc. Natl. Acad. Sci. USA *45*, 1582–1587. https://doi.org/10.1073/pnas.45.11.1582.
2. Goddard-Borger, E.D., and Williams, S.J. (2017). Sulfoquinovose in the biosphere: occurrence, metabolism and functions. Biochem. J. *474*, 827–849. https://doi.org/10.1042/BCJ20160508.

3. Loll, B., Kern, J., Saenger, W., Zouni, A., and Biesiadka, J. (2005). Towards complete cofactor arrangement in the 3.0 Å resolution structure of photosystem II. Nature *438*, 1040–1044. https://doi.org/10.1038/nature04224.

4. Harwood, J.L., and Nicholls, R.G. (1979). The plant sulpholipid– a major component of the sulphur cycle. Biochem. Soc. Trans. *7*, 440–447. https://doi.org/10.1042/bst0070440.

5. Liu, J., Wei, Y., Ma, K., An, J., Liu, X., Liu, Y., Ang, E.L., Zhao, H., and Zhang, Y. (2021). Mechanistically diverse pathways for sulfoquinovose degradation in bacteria. ACS Catal. *11*, 14740–14750. https://doi.org/10.1021/acscatal.1c04321.

6. Snow, A.J.D., Burchill, L., Sharma, M., Davies, G.J., and Williams, S.J. (2021). Sulfoglycolysis: catabolic pathways for metabolism of sulfoquinovose. Chem. Soc. Rev. *50*, 13628–13645.

7. Yagi, T., and Benson, A.A. (1962). Plant sulfolipid. V. Lysosulfolipid formation. Biochim. Biophys. Acta *57*, 601–603. https://doi.org/10.1016/0006-3002(62)91171-x.

8. 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., et al. (2018). Structural and biochemical insights into the function and evolution of sulfoquinovosidases. ACS Cent. Sci. *4*, 1266–1273. https://doi.org/10.1021/acscentsci.8b00453.

9. 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. https://doi.org/10.1038/nchembio.2023.

10. 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. https://doi.org/10.1038/nature12947.

11. 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–E4305. https://doi.org/10.1073/pnas.1507049112.

12. Frommeyer, B., Fiedler, A.W., Oehler, S.R., Hanson, B.T., Loy, A., Franchini, P., Spiteller, D., Schleheck, D., Hanson, S.R., and Loy, B.T. (2020). Environmental and intestinal phylum Firmicutes bacteria metabolize the plant sugar sulfoquinovose via a 6-deoxy-6-sulfofructose transaldolase pathway. iScience *23*, 101510. https://doi.org/10.1016/j.isci.2020.101510.

13. Liu, Y., Wei, Y., Zhou, Y., Ang, E.L., Zhao, H., and Zhang, Y. (2020). A transaldolase-dependent sulfoglycolysis pathway in Bacillus megaterium DSM 1804. Biochem. Biophys. Res. Commun. *533*, 1109–1114. https://doi.org/10.1016/j.bbrc.2020.09.124.

14. Sharma, M., Lingford, J.P., Petricevic, M., Snow, A.J.D., Zhang, Y., Järvå, M.A., Mui, J.W.-Y., Scott, N.E., Saunders, E.C., Mao, R., et al. (2022). Oxidative desulfurization pathway for complete catabolism of sulfoquinovose by bacteria. Proc. Natl. Acad. Sci. USA *119*. e2116022119. https://doi.org/10.2185/jjrm.23.202.

15. Burrichter, A., Denger, K., Franchini, P., Huhn, T., Müller, N., Spiteller, D., and Schleheck, D. (2018). Anaerobic degradation of the plant sugar sulfoquinovose concomitant with H2S production: Escherichia coli k-12 and desulfovibrio sp. Strain DF1 as co-culture model. Front. Microbiol. *9*, 2792. https://doi.org/10.3389/fmicb.2018.02792.

16. Hanson, B.T., Dimitri Kits, K., Löffler, J., Burrichter, A.G., Fiedler, A., Denger, K., Frommeyer, B., Herbold, C.W., Rattei, T., Karcher, N., et al. (2021). Sulfoquinovose is a select nutrient of prominent bacteria and a source of hydrogen sulfide in the human gut. ISME J. *15*, 2779–2791. https://doi.org/10.1038/s41396-021-00968-0.

17. Xing, M., Wei, Y., Zhou, Y., Zhang, J., Lin, L., Hu, Y., Hua, G., N Nanjaraj Urs, A., Liu, D., Wang, F., et al. (2019). Radical-mediated C-S bond cleavage in C2 sulfonate degradation by anaerobic bacteria. Nat. Commun. *10*, 1609. https://doi.org/10.1038/s41467-019-09618-8.

18. Samland, A.K., and Sprenger, G.A. (2009). Transaldolase: from biochemistry to human disease. Int. J. Biochem. Cell Biol. *41*, 1482–1494. https://doi.org/10.1016/j.biocel.2009.02.001.

19. Samland, A.K., Baier, S., Schürmann, M., Inoue, T., Huf, S., Schneider, G., Sprenger, G.A., and Sandalova, T. (2012). Conservation of structure and mechanism within the transaldolase enzyme family. FEBS J. *279*, 766–778. https://doi.org/10.1111/j.1742-4658.2011.08467.x.

20. Lehwess-Litzmann, A., Neumann, P., Parthier, C., Lüdtke, S., Golbik, R., Ficner, R., and Tittmann, K. (2011). Twisted Schiff base intermediates and substrate locale revise transaldolase mechanism. Nat. Chem. Biol. *7*, 678–684. https://doi.org/10.1038/nchembio.633.

21. 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*, 2901–2910. https://doi.org/10.1021/acs.joc.9b00055.

22. Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J.H., Lindahl, E., and Scheres, S.H.W. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife *7*, e42166. https://doi.org/10.7554/eLife.42166.

23. Schörken, U., Jia, J., Sahm, H., Sprenger, G.A., and Schneider, G. (1998). Disruption of Escherichia coli transaldolase into catalytically active monomers: evidence against half-of-the-sites mechanism. FEBS Lett. *441*, 247–250. https://doi.org/10.1016/S0014-5793(98)01521-X.

24. Jia, J., Schörken, U., Lindqvist, Y., Sprenger, G.A., and Schneider, G. (1997). Crystal structure of the reduced Schiff-base intermediate complex of transaldolase B from Escherichia coli: mechanistic implications for class I aldolases. Protein Sci. *6*, 119–124. https://doi.org/10.1002/pro.5560060113.

25. Schörken, U., Thorell, S., Schürmann, M., Jia, J., Sprenger, G.A., and Schneider, G. (2001). Identification of catalytically important residues in the active site of Escherichia coli transaldolase. Eur. J. Biochem. *268*, 2408–2415. https://doi.org/10.1046/j.1432-1327.2001.02128.x.

26. Sharma, M., Abayakoon, P., Epa, R., Jin, Y., Lingford, J.P., Shimada, T., Nakano, M., Mui, J.W.Y., Ishihama, A., Goddard-Borger, E.D., et al. (2021). Molecular basis of sulfosugar selectivity in sulfoglycolysis. ACS Cent. Sci. *7*, 476–487. https://doi.org/10.1021/acscentsci.0c01285.

27. Winter, G. (2010). Xia2: an expert system for macromolecular crystallography data reduction. J. Appl. Crystallogr. *43*, 186–190. https://doi.org/10.1107/S0021889809045701.

28. Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? Acta Crystallogr. D Biol. Crystallogr. *69*, 1204–1214. https://doi.org/10.1107/S0907444913000061.

29. Vagin, A., and Teplyakov, A. (2010). Molecular replacement with MOLREP. Acta Crystallogr. D Biol. Crystallogr. *66*, 22–25. https://doi.org/10.1107/S0907444909042589.

30. Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr. D Biol. Crystallogr. *62*, 1002–1011. https://doi.org/10.1107/S0907444906022116.

31. McNicholas, S., Potterton, E., Wilson, K.S., and Noble, M.E.M. (2011). Presenting your structures: the CCP4mg molecular-graphics software. Acta Crystallogr. D Biol. Crystallogr. *67*, 386–394. https://doi.org/10.1107/S0907444911007281.

32. Schrödinger, L.L.C. (2015). The PyMol Molecular Graphics System.

33. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera - a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612. https://doi.org/10.1002/jcc.20084.

34. Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132. https://doi.org/10.1107/S0907444904019158.

35. Beilsten-Edmands, J., Winter, G., Gildea, R., Parkhurst, J., Waterman, D., and Evans, G. (2020). Scaling diffraction data in the DIALS software package: algorithms and new approaches for multi-crystal scaling. Acta Crystallogr. D Struct. Biol. *76*, 385–399. https://doi.org/10.1107/S2059798320003198.

36. Zhang, K. (2016). Gctf: real-time CTF determination and correction. J. Struct. Biol. *193*, 1–12. https://doi.org/10.1016/j.jsb.2015.11.003.

37. Terwilliger, T.C., Sobolev, O.V., Afonine, P.V., and Adams, P.D. (2018). Automated map sharpening by maximization of detail and connectivity. Acta Crystallogr. D Struct. Biol. *74*, 545–559. https://doi.org/10.1107/S2059798318004655.

38. Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., et al. (2018). MolProbity: more and better reference data for improved all-atom structure validation. Protein Sci. *27*, 293–315. https://doi.org/10.1002/pro.3330.

39. Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. D Struct. Biol. *75*, 861–877. https://doi.org/10.1107/S2059798319011471.

40. Holm, L. (2020). DALI and the persistence of protein shape. Protein Sci. *29*, 128–140. https://doi.org/10.1002/PRO.3749.

41. Madeira, F., Pearce, M., Tivey, A.R.N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., and Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Res. *50*, W276–W279. https://doi.org/10.1093/NAR/GKAC240.

42. Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009). Jalview Version 2-A multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189–1191. https://doi.org/10.1093/bioinformatics/btp033.

Please cite this article in press as: Snow et al., Structure and mechanism of sulfofructose transaldolase, a key enzyme in sulfoquinovose metabolism, Structure (2023), https://doi.org/10.1016/j.str.2023.01.010

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
BL21(DE3) competent *E.coli*	New England Biolabs	Cat#2527H
Chemicals, peptides, and recombinant proteins		
*Bm*SF-TAL gene	*Bacillus megaterium* DSM1909	WP_013058335.1
Critical commercial assays		
NanoTemper Prometheus NT.48	NanoTemper	N/A
Deposited data		
*Bm*SF-TAL ligand-free (XRD)	This Paper	PDB 8C4I
*Bm*SF-TAL ligand-free EM Model	This Paper	PDB 8BC2
*Bm*SF-TAL ligand-free EM Map	This Paper	EMDB-15960
*Bm*SF-TAL SF complex (D5) EM Model	This Paper	PDB 8BC3
*Bm*SF-TAL SF complex (D5) EM Map	This Paper	EMDB-15961
*Bm*SF-TAL SF complex (C1) EM Model	This Paper	PDB 8BC4
*Bm*SF-TAL SF complex (C1) EM Map	This Paper	EMDB-15962
Experimental models: Cell lines		
BL21(DE3) competent *E.coli*	New England Biolabs	Cat#2527H
Recombinant DNA		
BmSF-TAL gene	NCBI database	WP_013058335.1
Software and algorithms		
CCP4i2	CCP4 project	https://www.ccp4.ac.uk/
Relion	Zivanov et al.[22]	https://www3.mrc-lmb.cam.ac.uk/relion/index.php/Main_Page
LC solutions	Shimadzu	https://www.shimadzu.com/an/products/liquid-chromatography/hplc-software/index.html
Astra V	Wyatt Technology	https://wyatt-technology-astra.software.informer.com/5.3/
MassLynx V.4.2	Waters	https://www.waters.com/waters/en_US/MassLynx-MS-Software/nav.htm?cid=513662&locale=en_US
Xia2	Winter,[27]	https://xia2.github.io/
AIMLESS	Evans and Murshudov,[28]	https://www.ccp4.ac.uk/
MOLREP	Vagin and Teplyakov[29]	https://www.ccp4.ac.uk/
BUCCANEER	Cowtan,[30]	https://www.ccp4.ac.uk/
CCP4MG	McNicholas et al.[31]	https://www.ccp4.ac.uk/
Pymol	Schrödinger,[32]	https://pymol.org/2/
UCSF Chimera	Pettersen et al.[33]	https://www.cgl.ucsf.edu/chimera/
Coot	Emsley et al.[34]	https://www.ccp4.ac.uk/
Inkscape	inkscape.org	https://inkscape.org/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Professor Gideon Davies (gideon.davies@york.ac.uk).

Please cite this article in press as: Snow et al., Structure and mechanism of sulfofructose transaldolase, a key enzyme in sulfoquinovose metabolism, Structure (2023), https://doi.org/10.1016/j.str.2023.01.010

Materials availability

A map for the *Bm*SF-TAL gene in pET28a can be made available upon request.

Data and code availability

- All cryo-EM data have been deposited in the PDB and EM Databank (EMDB) and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. All X-ray diffraction data have been deposited in the PDB and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

BL21(DE3) competent E. coli

In this study commercially available BL21(DE3) competent *E. coli* were used (product code C2527H). Unless otherwise stated these were grown in LB supplemented with 25 μg/mL kanamycin, at 37°C and shaken at 220rpm.

METHOD DETAILS

Gene expression and protein purification

The gene for *Bm*SF-TAL (WP_013058335.1) was codon optimised for *E. coli* expression and synthesized by GenScript Biotech (Netherlands) B.V., and sub-cloned into a standard pET28a vector that encodes an N-terminal hexa-His tag. The resulting plasmid was transformed into BL21(DE3) competent *E. coli* (New England Biolabs) using the heat shock method following the supplier's protocol, then plated onto LB-agar containing 35 μg mL^{-1} kanamycin and grown at 37°C overnight. Isolated colonies were picked and used to form liquid starter cultures in 20mL LB with 35 μg mL^{-1} kanamycin, then grown at 37°C, 180 rpm overnight. 5 mL of starter culture was used to inoculate 1 L of expression media (2x YT + 35 μg mL^{-1} kanamycin). Expression cultures were grown at 37°C, 225 rpm until an OD_{600} of 0.6–0.8 was achieved. Isopropyl β-D-thiogalactopyranoside (IPTG) was then added to a final concentration of 1 mM and induction took place over 18 h, at 18°C, shaken at 180 rpm. Cells were harvested by centrifugation at 4,445 rcf, 20 min, 4°C. The supernatant was discarded, and the pellet resuspended in 40 mL binding buffer (50 mM TRIS/HCl. 300 mM NaCl, 30 mM imidazole, pH 7.4) per 1 L culture. 1 tablet of EDTA-free protease inhibitor (Roche) was also added per 1 L culture. Cells were lysed either by sonication (5 pulses of 100% amplitude, 60s on, 60s off while on ice) or using a cell disruptor (25 kPSI, 4°C, eluted into an ice-cooled flask). Cell debris was pelleted by centrifugation at 35,199 rcf, 40 min, 4°C. The pellet was discarded, and the soluble fraction was loaded to a 5 mL HisTrap Nickel-nitrilotriacetic acid crude Fast flow (Ni-NTA crude FF) column (Amersham) at 2 mL min^{-1}, that had been pre-equilibrated with binding buffer. A linear gradient with elution buffer (50 mM TRIS/HCl, 300 mM NaCl, 500 mM imidazole, pH 7.4) was used to elute the His6-tagged protein into a 2 mL 96-well deep block, and protein purity and size were assessed by SDS-PAGE. Protein-containing fractions were pooled and buffer-exchanged into binding buffer. The pooled fractions were then concentrated to a 2 mL final volume and loaded to a HiLoad Superdex 16/600 S200 column (GE Healthcare) for size-exclusion chromatography. The column was pre-equilibrated into SEC buffer (50 mM TRIS/HCl, 300 mM NaCl, pH 7.4). Elution took place using SEC buffer at 1 mL min^{-1} into a 96-well deep block, and protein purity and size were assessed by SDS-PAGE. *Bm*SF-TAL-containing fractions were pooled and concentrated to a working concentration of 38 mg mL^{-1}, confirmed by A_{280} (extinction coefficient = 29,910 M^{-1} cm^{-1}) then snap-frozen with liquid nitrogen, and stored at −70°C (Figure S2).

SEC-MALLS analysis

Size-Exclusion Chromatography with Multiple Angle Laser Light Scattering (SEC-MALLS) was performed using 120 μL samples of *Bm*SF-TAL at 2 mg mL^{-1}, in 50 mM TRIS/HCl, 300 mM NaCl, pH 7.4. The experiments were conducted on a Wyatt HELEOS-II multi-angle light scattering detector with 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). Experiments were 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 2 column volumes 50 mM TRIS/HCl, 300 mM NaCl, pH 7.4 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. LC Solutions (Shimadzu) was used to control the HPLC and Astra V was used for the HELEOS-II and rEX detectors. The Astra data collection was 1 min shorter than the LC solutions run to maintain synchronisation. TM Data were analyzed 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).

NanoDSF analysis

*Bm*SF-TAL samples were at 1 mg mL^{-1} in 50 mM TRIS/HCl, 150 mM NaCl, pH 7.5. Ligands were weighed dry and dissolved in 50 mM TRIS/HCl, 150 mM NaCl, pH 7.4 to a stock concentration of 50 mM. Ligands and cofactors were added 10 min prior to loading to a final concentration of 2 mM, and three repeats were run of each sample, with each replicate loaded to a standard glass capillary. The

temperature was increased from 20°C to 95°C at 1°C min^{-1} in a linear gradient and W fluorescence was measured by a NanoTemper Prometheus NT.48. Data processing and graphing was performed in Origin.

LC-MS detection of substrates and products of transaldolase activity

A 50 μL reaction mixture containing 50 mM TRIS/HCl, pH 7.5, 10 mM C6 or C7 ketose (SF, S7P or F6P), 20 mM acceptor (GAP, SLA, E4P), 2 mM *Bm*SF-TAL, and without *Bm*SF-TAL were also performed. Completed reaction samples were diluted with organic solvents to a final composition of water:isopropanol:acetonitrile 10:10:80 (v/v), and were analyzed by LC-MS, which consisted of an Acquity I-Class LC connected to a Synapt G2-Si time-of-flight mass spectrometer(both Waters, Elstree, UK). A SeQuant ZIC-HILIC PEEK column (5μm, 200Å, 100 × 4.6 mm; Merck, Watford, UK) including a guard column (Phenomenex, Macclesfield, UK) was used, maintained at 40°C. Sample temperature was 14°C, injection volume: 10 μL.

The mobile phase consisted of A) 20 mM ammonium acetate in 1% acetonitrile, and B) acetonitrile. LC-MS gradient was 90% B at 0 min, which decreased linearly to 40% B over 30 min where it stayed until 40 min; increased to 90% B at 43 min and re-equilibrated until 56 min. The flow rate was 0.5 mL/min. The MS operated in negative electrospray ionization resolution mode (capillary voltage 2.0 kV; sampling cone voltage 60 V; source offset 80 V; source and desolvation temperatures 150°C and 220°C; cone gas 100 L/h; desolvation gas 600 L/h; nebuliser 6.5 bar; mass range 50–600 *m/z*; scan time 0.75 s for MS1 and 0.5 s for MS2 acquisitions). Precursors in the targeted MS2 scans were *m/z* 259.0 (F6P); 265.0 (13C 6 -F6P); 243.0 (SF); 246.0 (13C 3 -SF); 289.0 (S7P). A transfer cell collision energy ramp from 15 to 45 eV was used. Data was analyzed using MassLynx V.4.2 (Waters) for chromatogram visualization and Skyline 20.2.0 (Adams et al., 2020) for peak integration.

*Bm*SF-TAL crystals were grown at 20°C in 0.2 M $NaSO_4$, 0.1 M bis-TRIS propane, 20% PEG 3350, with a 3:2 protein:mother liquor ratio and 1 μL drop size. *Bm*SF-TAL was used in 50 mM TRIS/HCl, 300 mM NaCl pH 7.4 at 23 mg mL^{-1}. Crystals were harvested into liquid nitrogen, using nylon CryoLoops™ (Hampton). Cryoprotection took place over 5 s and used 10% glycerol, or 10% ethylene glycol in mother liquor. All crystals were tested using the in-house Rigaku Micro-Max 007HF X-ray generator with an RAXIS IV++ imaging plate detector. Data was collected at 120 K using a 700 Series Cryostream (Oxford Cryosystems). Diffraction pattern quality assessment and resolution estimate performed using ADXV. X-ray data collection took place at Diamond Light Source on beamline I-03. Data collection statistics are available in the results section. Data indexing and initial processing was performed using either DIALS or 3dii pipelines from the Xia2 package.[27,35] Data reduction was performed with AIMLESS, and resolution was cut until CC1/2 = 0.5.[28] Molecular replacement was performed using MOLREP, using a monomer of the ligand-free EM structure of *Bm*SF-TAL as a search model.[29] Automated model building was performed using BUCCANEER.[30] Diagram preparation for molecular models was performed using CCP4MG, Chimera or Pymol, depending on the desired outcome.[31–33] Data collection and refinement statistics are available in Table S1.

Electron cryomicroscopy (cryo-EM)

The York Glacios electron cryomicroscope (ThermoFisher Scientific), equipped with a Falcon IV detector, was used for imaging. The instrument was operated at 200 kV. Automated data collection was performed using EPU (ThermoFisher Scientific). Experiment-specific data collection parameters are available in Table 1. Image processing and 3D reconstruction was performed using Relion 3.[22] Motion correction was performed using MotionCorr v.2.0, using *Relions'* own implementation, initial CTF estimation was performed by GCTF v.1.06.[36] Particle picking for the ligand-free dataset was achieved through an initial round of Laplacian of Gaussian reference-free picking and 2D classification, after which classes resembling the particles of interest were used as the basis for particle selection. Particles for the SF dataset were picked based on the final 3D reconstruction for the ligand free, lowpass-filtered to 15 Å resolution. Final 2D classification in all cases was performed reference-free and at 3.6 Å/px. As the holozyme was expected to form the 2-layer, 5-member ring common to decameric TALs, broken particles and noisy classes that did not resemble the expected shape were removed. All classes showing high resolution features consistent with the expected arrangement were selected for production of an initial model, which was used as a reference for higher-resolution refinement and reconstruction. Particles used in this initial model were then re-extracted at 0.9 Å/px. In all cases final particle selection was achieved using 4 3D classes. These models were coarse, with angular sampling limited to 7.5°. This was performed in C1 for the ligand-free dataset and in D5 for the SF complex. Any classes not clearly exhibiting the correct overall fold were discarded. 3D auto-refinement then produced datasets of higher resolution. For resolution estimation and sharpening, solvent masking was achieved using a lowpass-filtered map (20 Å), expanded by 10px with a 6px soft edge. To improve the resolution, refinement of anisotropic magnification, trefoil and fourth order aberrations, astigmatism and defocus were carried out, with the latter two applied per-particle. Bayesian polishing based on a 10,000 particle training dataset was also performed. All model validation was performed using Molprobity.[37–39] Reconstruction details and Fourier shell correlation plots are in Figures S7., S8. and S9. Model building and validation was performed using Coot[34] and diagram preparation was performed identically to XRD structures.

Bioinformatics

Generation of *Bm*SF-TAL structural homologues was performed using the DALI web server.[40] All alignment and diagram preparation was performed using Jalview 11.1.4 and the T-COFFEE server.[41,42]

Please cite this article in press as: Snow et al., Structure and mechanism of sulfofructose transaldolase, a key enzyme in sulfoquinovose metabolism, Structure (2023), https://doi.org/10.1016/j.str.2023.01.010

QUANTIFICATION AND STATISTICAL ANALYSIS

Data collection and refinement statistics for all structures described in this work are available in Tables 1 and S1, as well as the specific PDB, EMDB entries for each dataset. Calculation of SD was performed to assess uncertainty in nanoDSF binding experiments, which were performed with 3 repeats in all cases.