A patchwork pathway for catabolism of the sulfosugar sulfofucose

Jinling Li,[a] Mihwa Lee,[a] Shuxin Yang,[a] Jessica M. Lewis,[b] Marion Herisse,[b] Sacha J. Pidot,[b] Nichollas E. Scott,[b] Spencer J. Williams[a*]

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

[b] Department of Microbiology and Immunology, University of Melbourne, at the Peter Doherty Institute for Infection and Immunity, Victoria 3000, Australia

Keywords: organosulfur, metabolism, sulfoglycolysis, sulfur cycle, structural biology

Abstract

We report the isolation of *Paracoccus onubensis* strain Merri, which can grow on the rare sulfosugar sulfofucose (6-deoxy-6-sulfo-D-galactose) as sole carbon source, leading to its breakdown and release of sulfite. Genome sequencing and comparative proteomics of samples grown on glucose versus sulfofucose identified a genetic locus containing most of the genes for a sulfoglycolytic pathway for breakdown of sulfofucose to the 3-carbon organosulfonate sulfolactaldehyde, and a biomineralization pathway for catabolism of sulfolactaldehyde to pyruvate and sulfite. A candidate sulfotagatose-1-phosphate aldolase was expressed and was active on the epimeric substrate sulfofructose-1-phosphate. Its 3D X-ray structure revealed a close homology to sulfofructose-1-phosphate aldolase from the sulfoglycolytic Embden-Meyerhof-Parnas pathway. A patchwork sulfofucose catabolism pathway is proposed involving a sulfofucolytic Embden-Meyerhof-Parnas and short-chain organosulfonate biomineralization pathway.

Significance

Sulfosugars are important compounds in the biogeochemical sulfur cycle and are produced by a wide range of photosynthetic plants, algae and cyanobacteria. While the pathways for the biosynthesis and catabolism of sulfoquinovose are well studied, much less is known about the rarer sulfosugar sulfofucose. We report a bacterium that can use sulfofucose as sole carbon source and achieve its complete breakdown to produce sulfite. We propose that sulfofucose catabolism is achieved using enzymes of the sulfoglycolytic Embden-Meyerhof-Parnas pathway that tolerate a change in stereochemistry at C4 of the sulfosugar. The sulfolactaldehyde output of this pathway is broken down to sulfite through a biomineralization pathway via sulfolactate. Microbial metabolism of this rare sulfosugar can be achieved through repurposing enzymes within known sulfosugar and short-chain organosulfonate catabolism catabolism.

Introduction

Sulfur is an elemental macronutrient that is essential for life. Within living organisms sulfur is incorporated into a wide range of organosulfur compounds (containing a carbon-sulfur bond) that function as metabolic currencies and permit the transfer of sulfur between producer and consumer organisms.[1, 2] The pathways used for the biosynthesis and catabolism of organosulfur metabolites form complex biochemical networks within the biogeochemical sulfur cycle.[3, 4] Organosulfur compounds that are significant on a global scale include the sulfur amino acids cysteine and methionine, the osmoprotectants dimethylsulfoniopropionate[5] and dimethyloxosulfonium propionate,[6] short-chain organosulfonates 2,3-dihydroxypropanesulfonate (DHPS) and sulfolactate (SL),[7, 8] and plant and cyanobacterial sulfolipid, which contains the sulfosugar sulfoquinovose (SQ, 6-deoxy-6-sulfo-D-glucose).[9]

Work over the past 30 years has revealed the pathways used by plants and cyanobacteria for the biosynthesis of SQ, and the degradation pathways that occur in bacteria.[9-12] The breakdown of SQ occurs through a variety of sulfoglycolytic and sulfolytic pathways. Sulfoglycolytic bacteria use pathways that result in cleavage of the carbon chain to allow partial consumption of SQ, followed by the excretion of short-chain C2 and C3-organosulfonates. These short-chain organosulfonates are substrates for secondary organosulfonate consumers that cleave the C-S bond, allowing metabolism of the remaining carbons and production of sulfite.[13-16] Sulfolytic organisms contain pathways that catalyze cleavage of the C-S bond, resulting in production of sulfite, and conversion of the carbon skeleton to glucose, which undergoes glycolysis.[16, 17]

Much less is known about the rarer sulfosugar sulfofucose (6-deoxy-6-sulfo-D-galactose), which has been reported as a component of glycolipids isolated from algae of the order Fucacaea,[18] and cell surface N-linked glycans from the archaeon *Thermoplasma acidophylum*.[19] No pathways are known for the biosynthesis or breakdown of sulfofucose.

Here, we report the chemical synthesis of sulfofucose and its use in the isolation of a *Paracoccus onubensis* strain that can grow on sulfofucose as sole carbon source. Using $^{13}C_6$-sulfofucose, we show that this strain utilizes all carbons of this substrate and releases sulfite as an end-product of metabolism. Comparative proteomics reveals protein abundance changes for many proteins with annotations that implicate them in organosulfur metabolism. We clone and express a candidate sulfohexulose aldolase and report its 3D X-ray structure. Collectively, these data allow

proposal of a patchwork pathway involving sulfofucolytic breakdown of sulfofucose to sulfolactaldehyde, and biomineralization to sulfite (**Figure 1**).

Results

Synthesis of sulfofucose. In order to isolate sulfofucose catabolizing bacteria, we synthesised multigram quantities of sulfofucose through a five-step route from D-galactose. Treatment of D-galactose with acetone and catalytic sulfuric acid gave diacetone-D-galactose. Conversion to the 6-iodo sugar with iodine, triphenylphosphine, and imidazole,[20] followed by nucleophilic substitution with potassium thioacetate gave the thioacetate. Oxidation with Oxone buffered with potassium acetate/acetic acid gave the protected sulfonate. Finally, treatment with trifluoroacetic acid afforded sulfofucose as the potassium salt. The overall yield from D-galactose was 52%. A similar pathway starting with $^{13}C_6$-D-galactose gave $^{13}C_6$-sulfofucose, in an overall yield of 46% (**Figure 2a, Figures S1-S11**).

Isolation and characterization of a *Paracoccus onubensis* bacterium able to grow on sulfofucose.

A soil sample from Joe's Garden, an inner-Melbourne market garden that has been continuously farmed by Chinese and Italian gardeners for 150 years, was used to inoculate a vitamin-enriched defined media containing 10 mM sulfofucose as sole carbon source. Sequential subculturing, repeated plating onto sulfofucose-containing media agar plates, followed by regrowth in sulfofucose liquid media led to isolation of strain Merri. This strain grew on glucose and sulfofucose to similar peak absorbance, but did not grow on SQ (**Figure 2b**). Genomic DNA was extracted from strain Merri, and was sequenced using the Illumina NextSeq platform. Analysis of the 16S rRNA gene revealed close similarity to *Methylarcula* and *Paracoccus* species, both members of the family Rhodobacteraceae. In silico DNA:DNA hybridization against other sequenced *Methylarcula* and *Paracoccus* genomes revealed *Paracoccus onubensis* to be the most closely related genome (94% average nucleotide identity). As such, the strain was designated as *P. onubensis* strain Merri and the key features of the draft genome are shown in **Table 1**.

We next investigated the fate of the carbon and sulfur content of sulfofucose upon metabolism of *P. onubensis* strain Merri. ^{13}C NMR analysis of spent culture medium of *P. onubensis* strain Merri grown on $^{13}C_6$-sulfofucose revealed complete consumption of the substrate, and production of ^{13}C-bicarbonate with no other distinct carbon-containing compounds, indicative of complete utilization of the complete carbon chain (**Figure 2d**). To study the fate of the sulfur within sulfofucose, we analyzed spent culture media of strain Merri grown on sulfofucose. We

followed consumption of sulfofucose with a reducing sugar assay, and measured production of sulfite and sulfate with colorimetric and turbidometric assays, respectively. We observed consumption of the majority of sulfofucose over 200 h, with concomitant production of sulfate and sulfite (**Figure 2c, Figure S12**). We saw peak sulfite levels at 160 h, and then a decline. This pattern is consistent with the known oxygen sensitivity of sulfite to oxygen,[17] suggesting that the first formed product is sulfite, which undergoes autooxidation to sulfate.

Identification of genetic loci expressed upon growth on sulfofucose

To identify possible proteins involved in the metabolism of sulfofucose, we grew samples of *P. onubensis* strain Merri in sulfofucose or glucose and examined changes in the proteome (**Figures 3a, b**). When sulfofucose was used as sole carbon source, significant increases occurred in proteins that are encoded within three regions of the genome (**Figure 4**).

Seven proteins with significantly increased abundance were mapped to a seven-gene cluster (cluster 1, locus tags 01197-01203), which encoded proteins annotated as enzymes involved in gluconeogenesis and the pentose phosphate pathway (**Figure 4**). Previous work has shown that sulfoglycolysis in *E. coli* leads to an increase in flux through gluconeogenesis, as a result of the sulfo-EMP pathway producing to the C3-phosphate DHAP.[21] Nine proteins with significantly increased abundance were mapped to an 11-gene cluster (cluster 3, locus tags 03829-03839) encoding proteins annotated with fumarate and tartrate metabolism that may be associated with the Krebs' cycle. This cluster also includes gene 03835, annotated as hypothetical protein, was homologous to the transmembrane TauE sulfite exporter.[22] Finally, eight proteins with significantly increased abundance were mapped to a 24-gene locus (cluster 2, locus tags 01699-01722), which encoded proteins annotated with a range of organosulfonate metabolism enzymes.

Within cluster 2, the proteins observed to increase in abundance in growth on sulfofucose included 01709, annotated as L-threonine 3-dehydrogenase, which is homologous to sulfolactaldehyde (SLA) dehydrogenase (catalyzing oxidation of SLA to sulfolactate), 01702, annotated as *R*-SL dehydrogenase (catalyzing oxidation of *R*-SL to sulfopyruvate), and 01703, annotated as *S*-SL dehydrogenase (catalyzing oxidation of S-SL to sulfopyruvate) (**Figure 4**). Immediately adjacent is 01701, which is annotated as SL sulfolyase (catalyzing elimination of sulfite from SL to give pyruvate). We propose that these proteins constitute a C3-organosulfonate

biomineralization pathway that allows SLA emerging from sulfoglycolysis to be oxidized to *R*-SL and then converted to pyruvate and sulfite, which is exported. This pathway could also allow scavenging of *S*-SL by conversion to *R*-SL by way of sulfopyruvate.

Within cluster 2, the levels of 01712 (annotated as tagatose 1,6-diphosphate aldolase), and 01722 (annotated as 6-deoxy-6-sulfogluconolactonase), were significantly increased. Gene 01712 sits adjacent to 01711 (annotated as sulfofructose (SF) kinase), and elsewhere in the genome is 00740 (annotated as SQ isomerase), and while levels of these proteins were not significantly altered in the comparative proteomics analysis, peptides showing that these proteins are expressed were present in the proteomics experiments. Together, these data provide support for either a sulfo-EMP-like or sulfo-ED-like sulfofucolysis pathway. The sulfofucolytic EMP pathway could involve conversion of sulfofucose to sulfotagatose (by 00740), phosphorylation to sulfotagatose-1-phosphate (catalyzed by 01711), and retroaldol to SLA and DHAP catalyzed by 01712 (acting as a sulfotagatose-1-phosphate aldolase) (**Figure 1**). On the other hand, a sulfofucolytic ED pathway would involve oxidation of sulfofucose to sulfofuconolactone (possibly by 03336, annotated as glucose-6-phosphate dehydrogenase), hydrolysis catalyzed by 01722 to sulfogalactonate, dehydration to 3-deoxy-2-ketosulfogalactonate (possibly by 01720, annotated as D-galactonate dehydratase), and retroaldol to SLA and pyruvate (catalyzed by 01712, acting as a 3-deoxy-2-keto-sulfogalactonate aldolase). In either scenario, SLA could be degraded by the biomineralization pathway encoded within cluster 3, while DHAP or pyruvate would enter central carbon metabolism.

Locus tag 01712 encodes a sulfohexulose-1-phosphate aldolase

The above analysis of the possible sulfofucolysis pathways that may operate within *P. onubensis* strain Merri is hampered by the large number of genes that are implicated, and the lack of authentic substrates available for each of the possibilities. However, in each option we propose that the protein encoded by 01712 acts as either a sulfotagatose-1-phosphate aldolase or 3-deoxy-2-keto-sulfogalactonate aldolase. The synthesis of the substrates for these enzymes have not been reported. However, the epimeric analogues, SF-1-phosphate (SFP) and 3-deoxy-2-keto-sulfogluconate (KDSG) are available; the former by our previously reported chemoenzymatic synthesis,[23] and the latter by a chemoenzymatic synthesis we report here in which sulfogluconate[24] was treated with recombinant *Pseudomonas putida* SG dehydratase[15] and then purified.

We synthesized an *E. coli* codon-harmonized version of 01712 and heterologously expressed the protein in *E. coli*. Incubation of samples of SFP and KDSG with recombinantly produced protein followed by analysis of the products by LC-MS revealed that SFP was converted into SLA and DHAP, while KDSG was inert (**Figure 5**). These data support a tentative assignment of this protein as a sulfohexulose-1-phosphate aldolase (*Po*SHPA), and the existence of a sulfofucolytic EMP-like pathway that converts sulfofucose to SLA and DHAP.

3D structure of sulfohexulose-1-phosphate aldolase *Po*SHPA

To gain insight into the structural basis of catalysis of aldolase *Po*SHPA, recombinantly produced protein was crystallized. The crystal structure was solved and refined to 2.5 Å (**Table 2**). The structure reveals a tetramer, assembled as a dimer-of-dimers. Each dimer consists of a domain-swapped monomer where a C-terminal end from one monomer chain extends into the fold of the other monomer (**Figure 6a**). The domain-swapped C-terminal ends in each dimer also participate in formation of the tetramer, making close contact with the other dimer. Overall 20% of surface area is buried due to tetramer formation with 14% of buried surface area being contributed by dimer formation as calculated using the Proteins, Interfaces, Structures and Assemblies (PISA) server.[25]

The central domain of SHPA displays an $(\alpha/\beta)_8$-barrel fold (**Figure 6b**). A search of the structural homologues in DALI server[26] identified tagatose-1,6-bisphosphate aldolase from *Streptococcus mutans* (PDB entry 3IV3 with DALI z score of 39.7, rmsd 1.8 and 28% sequence identity) as the closest homologue of *Po*SHPA. Interestingly, the identity of the second closest structural homologue is SFP aldolase, YihT from *E. coli* (PDB entry 7AG1 with z score of 32.6, rmsd 2.3 and 26% sequence identity). *Ec*YihT belongs to Class I aldolases that utilize an active-site lysine to form a Schiff base with the substrate. We have previously reported the structure of *Ec*YihT in complex with SFP, and described the molecular details of Schiff base formation.[27]

Superposition of the structure of *Ec*YihT with that of *Po*SHPA shows the similar overall folds (**Figure 6c**). The major deviations are found in the connecting loops of the $(\alpha/\beta)_8$-barrel and C-terminal end. The extended C-terminal end of *Po*SHPA (which is absent in *Ec*YihT) is involved in domain-swapped dimerization and tetramerization as described above. Inspection of the superposed structures reveals that the residues in the active sites are well conserved including

phosphate and sulfonate-binding residues (**Figure 6d**). Lys211 in *Po*SHPA is the likely Schiff-base forming residue, equivalent to Lys193 in *Ec*YihT. Two different conformations of Lys211 are observed in the apo-*Po*SHPA structure with two tetramers in the asymmetric unit; one of the two conformations is same as that of *Ec*YihT Lys193 in the complex with SFP in a Schiff-base, as shown in Figure 6d. The other more prevalent conformation (seven out of eight monomers) is similar to that observed in the apo-*Ec*YihT (**Figure S13**) These observations suggest conformational plasticity of the side chain of this residue, in the absence of Schiff-base formation.

Since the likely substrate of *Po*SHPA is sulfotagatose-1-phosphate, which is related to SFP by inversion of stereochemistry at C4, sulfotagatose-1-phosphate was modelled into the apo-*Po*SHPA structure by switching the positions of the hydrogen and hydroxyl at C4 (**Figure 6e**). Within the model, the side chain carboxyl of Glu166 is within hydrogen-bonding distance with the hydroxyl at C4 of sulfotagatose-1-phosphate, suggesting that this enzyme can achieve equivalent interactions to that for SFP, and supporting our proposed pathway involving sulfofucolytic breakdown of sulfofucose to sulfolactaldehyde via conversion of sulfofucose to sulfotagatose-1-phosphate.

Discussion

In this work we describe a *Paracoccus onubensis* strain isolated from market garden soil that can grow on sulfofucose as sole carbon source. The type member of this genus was isolated from the wall of a show cave in Spain.[28] *Paracoccus pantotrophus* strain NKNCYSA has been reported to grow on a range of organosulfonates including cysteate, taurine, homotaurine, isethionate and SL, suggesting that this genus harbors extensive organosulfonate degradative capacity.[29] Growth on sulfofucose led to complete degradation of the carbon framework of the sulfosugar, resulting in the production of sulfite, and with ^{13}C-bicarbonate as the only discernable carbon-based product.

Growth on sulfofucose was associated with alterations in protein abundances at three main genetic loci (clusters 1-3), relative to growth on glucose. Clusters 1 and 3 encoded genes associated with aspects of central carbon metabolism (along with a likely sulfite exporter), while Cluster 2 contained 22 genes, including many that were annotated with potential organosulfonate catabolism functions, but for which a complete pathway for sulfofucose catabolism was not apparent. However, a subset of genes encode for the biomineralization of the C3-organosulfonate

SLA, an output of the EMP and ED sulfoglycolysis pathways. These genes encoded for homologues of the enzyme SLA dehydrogenase (GabD), which oxidizes SLA to SL using NAD^+/$NADP^+$ as cofactor,[15] and for SL lyase (SuyB), which catalyzes elimination of sulfite from SL to give pyruvate.[30]

Regarding the likely sulfofucolytic capacity of *P. onubensis* strain Merri, a partial set of genes annotated for steps in the sulfo-ED and sulfo-EMP pathway were identified by comparative proteomics that might utilize sulfofucose through the equivalent steps. The missing capability for the sulfo-EMP pathway, sulfofucose-sulfotagatose isomerase, was suggested to be supplied by a gene annotated as SQ-SF isomerase that is found outside of the genetic loci identified by comparative proteomics, and which is expressed under conditions of sulfofucose growth.

In both the sulfo-ED and sulfo-EMP like pathways that were considered as candidates for sulfofucolysis, a common candidate gene (locus tag 01712) annotated as tagatose-bisphosphate aldolase, could fulfil the function of either sulfotagatose-1-phosphate aldolase or 3-deoxy-2-keto-sulfogalactonate aldolase. Direct analysis of this specific activity was not possible as neither substrate was available. However, the recombinantly produced protein (*Po*SHPA) was active on the C4 epimer SFP and was inactive on 3-deoxy-2-keto-sulfogluconate, supporting its assignment as a sulfohexulose-1-phosphate aldolase and the existence of an EMP like sulfofucolytic pathway in *P. onubensis* strain Merri. This is further supported by the close structural homology of *Po*SHPA with *Ec*YihT SFP aldolase, including similar active site architecture and conserved sulfonate and phosphate-binding residues.[27]

Does *P. onubensis* strain Merri encounter sulfofucose in its environmental niche? The occurrence of sulfofucose has only been described in two publications, including within glycoconjugates produced by an alga and an archaeon.[18, 19] This limited number of reports suggests it is a rare sulfosugar, although it is possible that it may have been misidentified for SQ in mass spectrometric based studies and could have a broader distribution than the limited literature reports imply. Joe's Garden is a riparian market garden situated on the banks of Merri Creek, with a history of flooding, suggesting a possible aquatic origin of either the bacterium or its substrate. Possibly, within either the soil or aquatic niche, this bacterium may encounter sulfofucose that is liberated from these known or other as yet unknown sulfofucose-containing glycoconjugates. No candidate sulfofucosidases were identified in this work, and we note that previous work has shown

that sulfoquinovosidases from *E. coli* and *Agrobacterium tumefaciens* are inactive on α-PNP-sulfofucoside.[31]

The patchwork pathway identified in this work appears to allow the scavenging of other organosulfonate molecules. This strain also grew on DHPS, but surprisingly, did not grow on SQ itself, nor SL. The failure to utilize SQ or SL may relate to the lack of an appropriate transporter, as we showed that *Po*SHPA could utilize SFP as a substrate. SQ and sulfofucose differ in stereochemistry at C4. This results in an epimeric relationship between proposed intermediates in the sulfoglycolytic EMP pathway (SQ, SF, and SFP), and the corresponding compounds in the sulfofucolytic pathway (sulfofucose, sulfotagatose, and sulfotagatose-1-phosphate). However, aldolase catalyzed cleavage of both sulfohexulose-1-phosphates will afford the same *R*-stereoisomer of SLA, as the stereocenter in this metabolite derives from carbon-5 of both precursors and results in destruction of stereochemistry at the carbon derived from carbon-4.

In conclusion, this work reveals a bacterium capable of the complete catabolism of the rare sulfosugar sulfofucose and provides evidence in support of a patchwork pathway involving a sulfofucolytic EMP like pathway to form C3-organosulfonate intermediates, and a classical biomineralization pathway involving a SL lyase to form sulfite. Future studies of the biochemical capacity of proposed enzymes in this organism to process sulfofucose-derived intermediates, and their specificity for this sulfosugar relative to SQ-derived intermediates will require development of methods for the synthesis of these compounds and assays for these enzymes, and could build upon previous approaches used to characterize the sulfoglycolytic EMP pathway in *E. coli*.[14, 27]

Materials and Methods

Synthesis of sulfofucose and ($^{13}C_6$)-sulfofucose

1,2:3,4-Di-*O*-isopropylidene-α-D-galactopyranose[32]: D-(+)-Galactose (10 g, 55.5 mmol) was dissolved in acetone (370 ml) and conc sulfuric acid (11 ml) was added at 0 °C. The reaction mixture was stirred at room temperature overnight, and then was neutralized by addition of saturated $NaCO_3$ solution until pH = 7. The precipitate was removed by filtration and the filtrate was combined and concentrated under reduced pressure. The crude title compound (**1**) 13.3 g (92.1%) was used without further purification.

1,2;3,4-di-*O*-isopropylidene-6-deoxy-6-iodo-α-galactopyranose[33]: Imidazole (11.1 g, 16.2 mmol), triphenylphosphine (21.0 g, 80.7 mmol) and iodine (18.3 g, 14.4 mmol) were added to a solution of 1,2:3,4-di-*O*-isopropylidene-α-D-galactopyranose (**1**) (10.3 g, 36.9 mmol) in a mixture of dry toluene (70 ml) and acetonitrile (35 ml). The mixture was stirred at 90 °C for 24 h. Then, the reaction was cooled and filtered. The filtrate was extracted with chloroform and washed with saturated $NaHCO_3$ solution. The organic layer was dried ($NaSO_4$) and purified by flash chromatography on silica gel (petroleum ether/ethyl acetate, 30:1 to 10:1) to give the title compound (**2**), 11.7 g (86%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.52 (1H, d, *J* 4 Hz), 4.59 (1H, dd, *J* 1.96, 6.28 Hz), 4.38 (1H, dd, *J* 1.40, 6.32 Hz), 4.28 (1H, dd, *J* 1.88, 3.94 Hz), 3.92 (1H, t, *J* 5.6 Hz), 3.29 (1H, dd, *J* 5.48, 7.96 Hz), 3.18 (1H, dd, *J* 5.76, 7.96 Hz), 1.52 (3H, s), 1.42 (3H, s), 1.33 (3H, s), 1.31 (3H, s); ^{13}C NMR ($CDCl_3$) δ 109.5, 108.9, 96.6, 71.5, 71.0, 70.5, 68.9, 26.0, 25.9, 24.9, 24.4, 23.8 ppm. Electrospray ionization mass spectrum (+ve mode): *m/z* 371.0348, calc (371.0350 for $C_{12}H_{20}IO_5^+$).

6-*S*-Acetyl-6-deoxy-1,2:3,4-di-*O*-isopropylidene-6-thio-α-D-galactopyranose: A mixture of 1,2;3,4-di-*O*-isopropylidene-6-deoxy-6-iodo-α-galactopyranose (**2**) (11.0 g, 29.7 mmol) and potassium thioacetate (8.2 g, 72 mmol) in acetone (100 ml) was refluxed for 24 h. Acetone was removed under reduced pressure, and the residue was extracted with ethyl acetate and saturated $KHCO_3$. The organic layer was dried ($MgSO_4$) and purified by flash chromatography (petroleum ether/ethyl acetate, 20:1 to 10:1) to give the title compound (**3**), 7.8 g (82%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.52 (1H, d, *J* 4 Hz), 4.61 (1H, dd, *J* =1.9, 4.0 Hz), 4.3 (1H, dd, *J* 1.4, 6.3 Hz), 4.27 (1H, dd, *J* =1.9, 3.9 Hz), 3.85 (1H, m), 3.17 (1H, dd, *J* 5.8, 8.0 Hz), 3.05 (1H, dd, *J* 8.7, 13.8 Hz), 2.34 (3H, s), 1.48 (3H, s), 1.46 (3H, s), 1.35 (3H, s), 1.32 (3H, s); ^{13}C NMR ($CDCl_3$) δ 195.9, 109.4, 108.8, 96.5, 72.0, 70.9, 70.5, 66.8, 30.5, 29.7, 25.92, 25.90, 25.0, 24.4 ppm. Electrospray ionization mass spectrum (+ve mode): *m/z* 319.1209, calc (319.1210 for $C_{14}H_{23}O_6S^+$).

Potassium 1,2:3,4-di-*O*-isopropylidene-6-deoxy-6-sulfonato-α-D-galactopyranose: Oxone (45 g, 73 mmol) was added to a solution of 6-*S*-acetyl-6-deoxy-1,2:3,4-di-*O*-isopropylidene-6-thio-α-D-galactopyranose (7.00 g, 24.4 mmol) and KOAc (24.0 g, 244 mmol) in glacial AcOH (80 ml). The mixture was stirred at room temperature overnight. The mixture was concentrated, and the residue was purified by flash chromatography (EtOAc/MeOH/H^2O, 15:2:1 to 5:2:1) to give the title compound (**4**) 7.9 g (90%). ^{1}H NMR (500 MHz, D_2O) δ 5.46 (1H, d, *J* 5.4 Hz), 4.64 (1H, dd, *J* 2.6,

8.3 Hz), 4.46 (1H, m), 4.40 (1H, t, *J* 7.0 Hz), 4.35 (1H, dd, *J* 2.6, 5.3 Hz), 3.70 (1H, s), 3.43–3.21 (2H, m), 2.95 (1H, dd, *J* 5.1, 14.4 Hz), 1.56 (3H, s), 1.41 (3H, s), 1.35 (3H, s) 1.34 (3H, s). Electrospray ionization mass spectrum (-ve mode): *m/z* 323.0800, calc (323.0806 for $C_{12}H_{19}O_8S^-$).

Potassium 6-deoxy-6-sulfonato-D-galactose

A mixture of potassium 1,2:3,4-di-*O*-isopropylidene-6-deoxy-6-sulfonato-α-D-galactopyranose (7.5 g, 20.7 mmol) was dissolved in trifluoroacetic acid (10 ml) and the mixture was stirred at room temperature for 30 min. The solvent was evaporated to dryness under reduced pressure. The residue was purified by flash chromatography (EtOAc/MeOH/H2O, 10:2:1 to 2:2:1) and give the title compound (**5**) 5.2 g (89%). ^{1}H NMR (500 MHz, D_2O) δ 5.15 (1H, d, *J* 3.9 Hz), 4.53 (1H, d, *J* 7.9 Hz), 4.40 (1H, dd, *J* 4.5, 7.5 Hz), 3.95 (1H, m), 3.88 (2H, dd, *J* 3.2, 14.7 Hz), 3.86 (1H, d, J 3.4 Hz), 3.81 (1H, dd, *J* 3.5, 10.4 Hz), 3.69 (1H, dd, *J* 4.0, 10.3 Hz), 3.60 (1H, dd, *J* 3.5, 10.0 Hz), 3.39 (1H, dd, *J* 8.0, 9.9 Hz), 3.20–3.03 (4H, m); ^{13}C NMR (D_2O) δ 96.53, 92.40, 72.75, 71.60, 71.18, 71.12, 70.05, 69.20, 68.12, 66.66, 52.19, 52.04 ppm. Electrospray ionization mass spectrum (-ve mode): *m/z* 243.0174, calc (243.0180 for $C_6H_{11}O_8S^-$).

Potassium 6-deoxy-6-sulfonato-D-($^{13}C_6$)-galactose

Potassium 6-deoxy-6-sulfonato-D-($^{13}C_6$) galactose was synthesized by using the same route of potassium 6-deoxy-6-sulfonato-D-galactose, but using D-($^{13}C_6$)-galactose as starting material. ^{1}H NMR (500 MHz, D_2O) δ 5.24 (1H, d, *J* 6.3 Hz), 4.91 (1H, d, *J* 5.5 Hz), 4.61 (1H, d, *J* 7.6 Hz), 4.47 (1H, s), 4.29 (1H, d, *J* 7.9 Hz), 4.18 (1H, s), 4.04 (1H, s), 3.94 (2H, m), 3.87 (1H, m), 3.76 (2H, m), 3.62 (3H, m), 3.43 (3H, m), 3.18 (5H, m), 2.90 (4H, m); ^{13}C NMR (500 MHz, D_2O) δ 96.3 (1C, d, *J* 45.3 Hz), 92.2 (1C, d, *J* 44.6 Hz), 72.5 (1C, t, *J* 37.2 Hz), 71.8–69.7 (4C, m), 69.0 (1C, t, *J* 38.1 Hz), 67.8 (1C, t, *J* 44.8 Hz), 66.4 (1C, t, *J* 39.2 Hz), 51.9 (1C, dd, *J* 20.4, 41.5 Hz). Electrospray ionization mass spectrum (-ve mode): *m/z* 249.0373, calc (249.0381 for $^{13}C_6H_{11}O_8S^-$)

2-Keto-3,6-dideoxy-6-sulfogluconate, disodium salt

Four separate reaction mixtures were prepared as follows. Solutions of 6-deoxy-6-sulfogluconate (1 ml of 10 mM stock; 0.0304 g, 0.10 mmol), Tris-buffer pH 8.0 (5 ml of 50 mM stock), $MnCl_2$ (0.5 ml of 50 mM stock), and deionized water (Milli-Q, 3.5 ml) were combined. The final volume was 10 mL; the final concentrations were 6-deoxy-6-sulfogluconate (1 mM), Tris-buffer pH 8.0 (25 mM), and $MnCl_2$ (2.5 mM). SG dehydratase (1 mg/mL, 10 μL) was added to each reaction solution to

start the reaction, and the solutions were kept at 30 °C for 24 h. The solutions were then heated at 80 °C for 5 min, then were freeze-dried to give a light brown residue. The four separate residues were combined and purified by flash chromatography on silica gel (20-60% EtOAc in MeOH/H_2O = 2:1). The appropriate fractions were combined and concentrated under vacuum to give a white foam. The residue was dissolved in water and applied to an ion-exchange column (5 ml of AmberChrom 50W x 2 resin, 100-200 mesh, Na^+ form) and eluted with water. Freeze-drying gave the title compound, 0.092 g (81%). Based on NMR analysis, the compound was assigned as a mixture of approx. 80% α,β-furanose anomers (in a 9:10 ratio), 15% acyclic keto compound, 5% 2,3-enol, as well as traces of other unidentified enol isomers. ^{1}H NMR (600 MHz, D_2O) δ 2.04 (0.9H, dd, *J* 14.2, 4.1 Hz, H3-α or β), 2.32 (1H, dd, *J* 14.0, 6.2 Hz, H3-α or β), 2.32–2.39 (0.4H, m, H3-*keto*), 2.41 (1H, dd, *J* 13.8, 6.4 Hz, H3-α or β), 2.56 (0.4H, dd, *J* 15.2, 3.4 Hz, H3-*keto*), 2.62 (0.9H, dd, *J* 14.2, 7.4 Hz, H3-α or β), 2.92–3.12 (m, H6,6-*keto*, H6,6-*enol*), 3.22–3.30 (m, 5.5H, H6,6-α,β, H6,6-*keto*, H6,6-*enol*), 3.67–3.84, 3.99–4.03, 4.04–4.08, 4.13–4.17 (4 m, H4,5-*keto*, H4,5-*enol*), 4.34–4.38, 4.42–4.47 (2 m, 2.3H and 2.3H, H4,4-α,β, H5,5-α,β, H4,5-*keto*, H4,5-*enol*), 5.21 (d, *J* 3.7 Hz, H3-*enol*); ^{13}C NMR (151 MHz, D_2O) δ 39.7 and 42.1 (1C, C3-*keto*), 42.5, 42.6 (2C, C3-α,β), 53.2, 53.3 (2C, C6-keto, C6-*enol*), 53.8, 55.0 (2C, C6-α,β), 69.5, 70.5 and 71.5 (2C, C5-*keto*, C5-*enol*), 74.47, 74.60 (2C, C4-α,β), 74.42, 74.55 (1C, C4-*keto*, C4-*enol*), 81.2, 82.2 (2C, C5-α,β), 103.5, 103.7 (2C, C2-α,β), 112.0, 116.1 (C2-*enol*, C3-*enol*), 169.4, 176.3, 176.7 (2C, C1-α,β), 179.6 (2C, C1-*keto*, C1-*enol*), 203.6 (1C, C2-*keto*). Electrospray ionization mass spectrum (-ve mode): *m/z* 241.0021, calc (241.0024 for $C_6H_9O_8S^-$).

Bacterial growth media

Growth media was prepared using M9 minimal salts media (2 ml), trace metal solution (0.1 ml), and vitamin solution (0.01 ml) and contained 10 mM sulfofucose as sole carbon source, made up to a final volume of 10 ml with deionized water. M9 minimal salts media contains 0.45 M Na_2HPO_4, 0.11 M KH_2PO_4, 0.09 M NH_4Cl, 0.04 M NaCl, 0.1 M $MgSO_4$, 0.1 M $CaCl_2$. Trace metal solution contains 0.4 mM $FeCl_3$, 0.08 mM $CoCl_2$, 0.08 mM $CuCl_2$, 0.08 mM $MnCl_2$, 0.55 mM $ZnCl_2$, 0.01 mM $NiCl_2$, 0.05 mM Na_2MoO_4 and 0.5 mM H_3BO_3. Vitamin mixture contains 0.04 mM biotin, 0.05 mM calcium pantothenate, 0.15 mM thiamine hydrochloride, 0.36 mM p-aminobenzoic acid, 0.81 mM

nicotinic acid, 1.49 mM pyridoxamine dihydrochloride, 0.01 B12 (cyanocobalamin). Growth curves were measured using a custom built MicrobeMeter portable spectrophotometer.[34]

Isolation of *Paracoccus onubensis* strain Merri

Paracoccus onubensis strain Merri was isolated from enriched culture, obtained from soil of Joe's Market Garden (34 Edna Grove, Coburg, Victoria 3058; 37.7451° S, 144.9813° E). A soil sample was collected and approximately 1 g of soil was suspended in 5 ml of sterilized M9 media (supplemented with trace metals and vitamins) containing 10 mM sulfofucose as sole carbon source. The culture was incubated at 30 ℃ for 4 d with agitation at 250 rpm. A subsample (100 µL) was transferred into fresh M9 media (supplemented with trace metals and vitamins) and grown for a further 4 d. This step was repeated five times and after outgrowth of the final culture for 4 d, cells were plated onto sulfofucose-agar plates (20 mM sulfofucose, 15 g/l agar) and incubated at 30 °C in the dark. After 3 d, a single colony was picked and inoculated into fresh M9 media (supplemented with trace metals and vitamins) containing 10 mM sulfofucose and incubated at 30 ℃ while shaking at 250 rpm. Once the culture was visibly turbid, cells were again plated onto sulfofucose-agar plates, incubated at 30 °C in dark and single colony picked and inoculated again into fresh vitamin supplemented M9 media containing 10 mM sulfofucose. Once the culture was visibly turbid, frozen stocks were prepared by diluting to 5% DMSO and freezing at -80 °C. This strain could grow on racemic DHPS (5 mM), but did not grow on SQ (5 mM), or *R*-SL (5 mM).

Genome sequence, assembly and annotation

P. onubensis strain Merri was grown in 10 ml 10 mM sulfofucose for 5 days and then the cell pellet was harvested by centrifugation. Genomic DNA was isolated using the GenElute DNA extraction kit (Sigma) with inclusion of lysozyme and RNAase, then it was sequenced using an Illumina NextSeq at the Peter Doherty Institute for Infection and Immunity, Parkville, Victoria, Australia. DNA was prepared for sequencing on the Illumina NextSeq platform using the Nextera XT DNA preparation kit (Illumina) with ×150 bp paired end chemistry and with a targeted sequencing depth of >50×. A draft genome was assembled using Shovill v1.0.9 (https://github.com/tseemann/shovill) and annotated using PGAP as part of the NCBI Genome submission pipeline. GC percentage and ANI

calculations were performed using ANI calculator (https://www.ezbiocloud.net/tools/ani).[35] Assembled genomes have been deposited at the NCBI (GenBank accession: Bioproject PRJNA993841). General features for the isolated bacterium are reported in Table 1.

^{13}C-NMR analysis of spent culture media from growth of *P. onubensis* strain Merri on ($^{13}C_6$)-sulfofucose

M9 minimal media (supplemented with trace metals and vitamins) (4 ml) containing 10 mM ($^{13}C_6$)-sulfofucose was inoculated with *P. onubensis* strain Merri and grown to mid-log phase (5 days) and stationary phase (14 days) at 30 °C (250 rpm). The final A_{600} was 0.3 (5 days) and 0.52 (14 days). Cells were sedimented by centrifugation at 5,000 × *g* for 3 min, and the supernatant was collected. 540 µL of culture supernatant was diluted with 60 µL D_2O and ^{13}C-NMR spectra acquired using a 500 MHz spectrometer. A control sample contained 10 mM ($^{13}C_6$)-sulfofucose in M9 minimal media prepared in the same way for NMR analysis.

Analysis of sulfofucose, sulfite and sulfate in *P. onubensis* culture

M9 minimal media (supplemented with trace metals and vitamins) 50 mL containing 10 mM sulfofucose was inoculated with *P. onubensis* strain Merri and grown at 30 °C at 250 rpm in a 500 mL baffled conical flask. Samples were withdrawn at various time points, cells were pelleted by centrifugation for 5 min at 5,000 *g*, and samples were frozen and stored at -80 °C prior to analysis for reducing sugar, sulfate or sulfite content.

Reducing sugar assay: The reducing sugar assay was performed according to the procedure of Blakeney and Mutton.[36] Alkaline diluent was prepared by the addition of sodium hydroxide (20 g, 0.50 mol) to a solution of 0.10 M trisodium citrate (50 mmol, 500 ml) and 0.02 M calcium chloride (13 mmol, 500 ml). PAHBAH working solution was prepared by dissolving 4-hydroxybenzhydrazide (PAHBAH) (0.25 g, 1.6 mmol) in alkaline diluent (50 ml). The PAHBAH working solution should be made fresh shortly before use. To determine reducing sugar concentration, 90 µL of PAHBAH working solution and 10 µL of culture supernatant were combined in a 96-well half-area microplate. The mixture was heated at 98 °C for 5 min, and the absorbance read at 415 nm using a UV/visible

spectrophotometer. Concentrations of sulfofucose were determined by reference to a standard curve constructed using sulfofucose. Analyses were performed in duplicate and errors are reported as standard deviation.

Turbidometric sulfate assay: The sulfate assay was performed according to the procedure of Sörbo.[37] This assay uses a barium sulfate–polyethylene glycol (Ba-PEG) reagent, which contains polyethylene glycol (PEG) to stabilize $BaSO_4$ crystals and a small amount of preformed $BaSO_4$ seed crystals to improve the reproducibility and linearity of the assay. Ba-PEG reagent was prepared by dissolving $BaCl_2$ (42 mg, 0.20 mmol) and polyethylene glycol 6000 (0.75 g) in deionized water (5.0 ml). A small amount of Na_2SO_4 (10 µl, 50 mM) was added to this solution, with efficient magnetic stirring to generate preformed $BaSO_4$ seed crystals. The Ba-PEG reagent should be prepared fresh before use. Individual sulfate assays were conducted as follows: In an Eppendorf tube, 50 µl of 0.5 M HCl was added to 50 µl culture supernatant, followed by Ba-PEG reagent (50 µl). The mixture was mixed vigorously and then 90 µl was transferred to the 96-well half-area microplate. The absorbance of the sample at 400 nm was measured using a UV/visible spectrophotometer. Concentrations of sulfate were determined by reference to a standard curve constructed using 0-10 mM Na_2SO_4 . Analyses were performed in duplicate and errors are reported as standard deviation.

Colorimetric fuchsin sulfite assay: The fuchsin sulfite assay was performed according to the procedures of Brychkova et al.[38] and Kurmanbayeva et al.[39] This method requires three freshly pre-prepared solutions: Reagents A, B, and C. Reagent A was prepared by adding 1 ml Schiff's reagent to 8 ml deionised water, followed by addition of 5 M HCl until the solution turned colorless. Reagent B was prepared by diluting formaldehyde (36% in H_2O, 0.32 ml) in deionised water (9.68 ml) at 0 °C. Reagent C was prepared by dilution of Reagent A (1 ml) in deionised water (7 ml), prior to the addition of solution reagent B (1 ml). Individual sulfite assays were performed by adding Reagent C (100 µl) to the cell supernatant (10 µl). The sample was incubated at 25 °C for 10 min and then transferred to the 96 well half area microplate. The absorbance of the sample at 570 nm was determined by using a UV/visible spectrophotometer. Concentrations of sulfite were determined by reference to a standard curve constructed using Na_2SO_3. Analyses were performed in duplicate and errors are reported as standard deviation.

Comparative proteomics

Six independent cultures of *P. onubensis* strain Merri were grown in 10 mM sulfofucose M9 media and another six independent cultures were grown in 10 mM glucose M9 media (both supplemented with trace metals and vitamins). The cultures were harvested at mid-log phase (A_{600} approx 0.25), and bacteria collected by centrifugation at 4000 × *g* for 10 min. The pelleted cells were washed by resuspension with fresh 10 ml ice-cold phosphate-buffered saline (PBS) and re-pelleting three times, with the supernatant discarded each time. Cells were resuspended and transferred to a 1.5 ml tube using ice-cold PBS (1 ml), and collected by centrifugation for 2 min at 10,000 × *g*, and then snap-frozen in liquid nitrogen.

Frozen whole bacterial pellets were prepared using the in-StageTip preparation approach as previously described.[40] Cells were resuspended in 4% sodium deoxycholate (SDC), 100 mM Tris pH 8.0 and boiled at 95 °C with shaking for 10 min to aid solubilisation. Samples were allowed to cool for 10 min and then boiled for a further 10 min before the protein concentration was determined by bicinchoninic acid assays (Thermo Fisher Scientific). 100 µg of protein equivalent for each sample was reduced/alkylated with the addition of tris-2-carboxyethylphosphine hydrochloride and iodoacetamide (final concentrations 20 mM and 60 mM, respectively), by incubating in the dark for 1 h at 45 °C. Following reduction/alkylation, samples were digested overnight with trypsin (1/33 w/w Solu-trypsin, Sigma) at 37 °C with shaking at 1000 rpm. Digests were then halted the addition of isopropanol and trifluoroacetic acid (TFA) (50% and 1%, respectively) and samples were cleaned up using home-made SDB-RPS StageTips prepared according to previously described protocols.[40-42] SDB-RPS StageTips were placed in a Spin96 tip holder 2 to enable batch-based spinning of samples and tips conditioned with 100% acetonitrile; followed by 30% methanol, 1% TFA followed by 90% isopropanol, 1% TFA with each wash spun through the column at 1000 *g* for 3 min. Acidified isopropanol / peptide mixtures were loaded onto the SDB-RPS columns and spun through, and tips were washed with 90% isopropanol, 1% TFA followed by 1% TFA in deionised water. Peptide samples were eluted with 80% acetonitrile, 5% ammonium hydroxide and dried by vacuum centrifugation at room temperature, then were stored at -20 °C.

Prepared digested proteome samples were re-suspended in Buffer A* (2% acetonitrile, 0.01% TFA) and separated using a two-column chromatography setup composed of a PepMap100 C18 20-mm by 75-µm trap and a PepMap C18 500-mm by 75-µm analytical column (Thermo

Fisher Scientific). Samples were concentrated onto the trap column at 5 µl/min for 5 min with Buffer A (0.1% formic acid, 2% DMSO) and then infused into a Orbitrap Fusion Eclipse at 300 nl/minute via the analytical columns using a Dionex Ultimate 3000 UPLCs (Thermo Fisher Scientific). 125-minute analytical runs were undertaken by altering the buffer composition from 2% Buffer B (0.1% formic acid, 77.9% acetonitrile, 2% DMSO) to 22% B over 95 min, then from 22% B to 40% B over 10 min, then from 40% B to 80% B over 5 min. The composition was held at 80% B for 5 min, and then dropped to 2% B over 2 min then was held at 2% B for another 8 min. The Fusion Eclipse Mass Spectrometer was operated in a data-dependent mode with a single Orbitrap MS scan (375-1500 *m/z* and a resolution of 120k) was acquired every 3 seconds followed by Orbitrap MS/MS higher energy collisional dissociation (HCD) scans of precursors (stepped NCE 25,30,40%, with a maximal injection time of 80 ms with the Automatic Gain Control set to 400% and the resolution to 30k).

Identification and LFQ analysis were accomplished using MaxQuant (v1.6.17.0)[43] using the in-house generated proteome of *P. onubensis* strain Merri (RES21-00833.faa) allowing for oxidation on methionine. The LFQ and "Match Between Run" options were enabled to allow comparison between samples. The resulting data files were processed using Perseus (v1.4.0.6)[44] to compare the growth conditions using Student's t-tests as well as Pearson correlation analyses. For LFQ comparisons independent replicates were grouped and missing values imputed based on the observed total peptide intensities with a range of 0.3σ and a downshift of 1.8σ.

The mass spectrometry proteomics data has been deposited in the Proteome Xchange Consortium[45] via the PRIDE partner repository with the data set identifier: PXD043488 Username: reviewer_pxd043488@ebi.ac.uk Password: DXxRoG1Y.

Cloning, protein expression and purification of *Po*SHPA

The codon-harmonized gene (locus tag 01712) of sulfohexulose-1-phosphate aldolase (SHPA) was synthesized and cloned between *Nco*I and *Xho*I in pET-28a with an N-terminal hexahistidine tag, followed by the TEV protease recognition site. For protein production, pET28a-SHPA was transformed and expressed in *Esherichia coli* BL21(DE3). Cells were grown in LB media with 50 µg/ml kanamycin at 37 °C and induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside at an A_{600} of 0.8. After further incubation at 18 °C for 16 h, the cells were harvested by centrifugation at

6000 *g*. The cell pellet was resuspended in buffer A (20 mM Tris–HCl (pH 7.5), 500 mM NaCl, 25 mM imidazole), disrupted using a TS Series Benchtop cell-disruptor and clarified by centrifugation at 40000 *g*. The soluble fraction was then applied to a chelating column (Hi-Trap HP, Cytiva) charged with Ni^{2+}. SHPA was eluted using a gradient of 25–500 mM imidazole in buffer A. Pooled fractions were applied to a size-exclusion column (HiLoad 16/600 Superdex 200 pg) equilibrated with 20 mM Tris–HCl (pH 7.5), 200 mM NaCl. Final protein samples were concentrated to 28 mg/ml, snap-frozen using liquid nitrogen and stored at -80°C until further use.

Detection of sulfohexulose aldolase activity using HPLC–LC–MS/MS

HPLC–LC–MS/MS analysis was performed as follows using solvent A: 20 mM NH_4OAc in 1% acetonitrile; solvent B: acetonitrile. The HPLC conditions were: from 90% B to 50% B over 7 min; then 50% B for 2 min; back to 90% B over 3 min and then 90% B for 1 min; flow rate, 0.50 ml min^{-1}; injection volume, 5 µL. The mass spectrometer was operated in negative ionization mode. 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 DHAP: retention time: 7.78 min. ESI–MS/MS *m/z* of $[M-H]^-$ 169, product ions 151, 97, 79.

For SFP: retention time: 6.83 min. ESI–MS/MS *m/z* of $[M-H]^-$ 323, product ions 225, 207,153.

Protein crystallization and X-ray diffraction data collection

Crystals of *Po*SHPA were grown by hanging-drop vapor diffusion experiments at 20°C. 2 µl of *Po*SHPA (7 mg/ml) was mixed with an equal volume of reservoir solution (0.1 M HEPES (pH 7.5), 0.2 M L-proline, 14–18% (w/v) PEG3350) and equilibrated against 0.5 ml of the reservoir solution. Prior to cryocooling, crystals were successively transferred to artificial reservoir solutions containing 20% ethylene glycol in 10% increments. Diffraction data were recorded on beamline MX2 at the Australian Synchrotron[46] at a wavelength of 0.954 Å at 100 K. The data were processed with XDS,[47] and merged and scaled with AIMLESS.[48] Data collection and merging statistics are summarized in **Table 2**.

Crystal structure solution and refinement

The crystal structure of *Po*SHPA was solved by molecular replacement using PHASER[49] within the CCP4 suite.[50] The AlphaFold model of the closest sequence homologue, annotated as DUF2090 domain-containing protein from *Paracoccus onubensis* (Uniprot ID A0A418T8G4; 99.1% sequence identity; AlphaFold model AF-A0A418T8G4-F1),[51, 52] was used as initial search model. The AlphaFold model found a solution with eight copies of monomer, locating two tetramers in the asymmetric unit. Iterative model building with COOT[53] and refinement with autoBUSTER[54] was carried out. The quality of the model was validated using MolProbity.[55] The refinement statistics are included in Table 2. The atomic coordinates have been deposited in the Protein Data Bank as entry 8TNF.

Acknowledgements

This work was supported by the Australian Research Council (DP210100233, DP210100233). Aspects of this research were undertaken on the Macromolecular Crystallography beamlines at the Australian Synchrotron (Victoria, Australia), part of ANSTO, and made use of the Australian Cancer Research Foundation (ACRF) detector. We thank the beamline staff for their professional support. J.L. is supported by a Ph.D. scholarship from the China Scholarship Council and the Norma Hilda Schuster Scholarship. Ethan Goddard is thanked for valuable discussions, and Alexandre Tolotchov is thanked for assistance with the construction and use of the MicrobeMeter. S.J.P is supported by and National Health and Medical Research Council of Australia (NHMRC) Ideas grant (GNT2021638) and ARC Discovery Project Grant (DP230102668). N.E.S is supported by an Australian Research Council Future Fellowship (FT200100270) and an ARC Discovery Project Grant (DP210100362). We thank the Melbourne Mass Spectrometry and Proteomics Facility of The Bio21 Molecular Science and Biotechnology Institute for access to MS instrumentation.

References

1. Durham, B.P. Deciphering Metabolic Currencies That Support Marine Microbial Networks. *mSystems* **6**, e0076321 (2021).
2. Durham, B.P., Boysen, A.K., Carlson, L.T., Groussman, R.D., Heal, K.R., Cain, K.R., Morales, R.L., Coesel, S.N., Morris, R.M., Ingalls, A.E. & Armbrust, E.V. Sulfonate-based networks between eukaryotic phytoplankton and heterotrophic bacteria in the surface ocean. *Nat. Microbiol.* **4**, 1706-1715 (2019).
3. Moran, M.A. & Durham, B.P. Sulfur metabolites in the pelagic ocean. *Nat. Rev. Microbiol.* **17**, 665-678 (2019).
4. Chaudhary, S., Sindhu, S.S., Dhanker, R. & Kumari, A. Microbes-mediated sulphur cycling in soil: Impact on soil fertility, crop production and environmental sustainability. *Microbiol. Res.* **271**, 127340 (2023).
5. Shaw, D.K., Sekar, J. & Ramalingam, P.V. Recent insights into oceanic dimethylsulfoniopropionate biosynthesis and catabolism. *Environ. Microbiol.* **24**, 2669-2700 (2022).
6. Thume, K., Gebser, B., Chen, L., Meyer, N., Kieber, D.J. & Pohnert, G. The metabolite dimethylsulfoxonium propionate extends the marine organosulfur cycle. *Nature* **563**, 412-415 (2018).
7. Celik, E., Maczka, M., Bergen, N., Brinkhoff, T., Schulz, S. & Dickschat, J.S. Metabolism of 2,3-dihydroxypropane-1-sulfonate by marine bacteria. *Org. Biomol. Chem.* **15**, 2919-2922 (2017).
8. Cook, A.M., Denger, K. & Smits, T.H. Dissimilation of C3-sulfonates. *Arch. Microbiol.* **185**, 83-90 (2006).
9. Goddard-Borger, E.D. & Williams, S.J. Sulfoquinovose in the biosphere: occurrence, metabolism and functions. *Biochem. J.* **474**, 827–849 (2017).
10. Benning, C. Biosynthesis and function of the sulfolipid sulfoquinovosyl diacylglycerol. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **49**, 53-75 (1998).

11. Snow, A.J.D., Burchill, L., Sharma, M., Davies, G.J. & Williams, S.J. Sulfoglycolysis: catabolic pathways for metabolism of sulfoquinovose. *Chem. Soc. Rev.* **50**, 13628–13645 (2021).
12. Wei, Y., Tong, Y. & Zhang, Y. New mechanisms for bacterial degradation of sulfoquinovose. *Biosci. Rep.* **42** (2022).
13. Frommeyer, B., Fiedler, A.W., Oehler, S.R., Hanson, B.T., Loy, A., Franchini, P., Spiteller, D. & Schleheck, D. Environmental and Intestinal Phylum Firmicutes Bacteria Metabolize the Plant Sugar Sulfoquinovose via a 6-Deoxy-6-sulfofructose Transaldolase Pathway. *iScience* **23**, 101510 (2020).
14. Denger, K., Weiss, M., Felux, A.K., Schneider, A., Mayer, C., Spiteller, D., Huhn, T., Cook, A.M. & Schleheck, D. Sulphoglycolysis in *Escherichia coli* K-12 closes a gap in the biogeochemical sulphur cycle. *Nature* **507**, 114-117 (2014).
15. Felux, A.K., Spiteller, D., Klebensberger, J. & Schleheck, D. Entner-Doudoroff pathway for sulfoquinovose degradation in *Pseudomonas putida* SQ1. *Proc. Natl. Acad. Sci. USA* **112**, E4298-4305 (2015).
16. Liu, J., Wei, Y., Ma, K., An, J., Liu, X., Liu, Y., Ang, E.L., Zhao, H. & Zhang, Y. Mechanistically Diverse Pathways for Sulfoquinovose Degradation in Bacteria. *ACS Catalysis*, 14740-14750 (2021).
17. Sharma, M., Lingford, J.P., Petricevic, M., Snow, A.J.P., Zhang, Y., Jarva, M., Mui, J.W.-Y.., Scott, N.E., Saunders, E.C., Mao, R., Epa, R., da Silva, B.M., Pires, D.E.V., Ascher, D.B., McConville, M.J., Davies, G.J., Williams, S.J., and Goddard-Borger, E.D. Oxidative desulfurization pathway for complete catabolism of sulfoquinovose by bacteria. *Proc. Natl. Acad. Sci. USA* **119**, e2116022119 (2022).
18. Liem, P.Q. & Laur, M.-H. Structures, teneurs et compositions des esters sulfuriques, sulfoniques, phosphoriques des glycosyldiglycérides de trois fucacées. *Biochimie* **58**, 1367-1380 (1976).
19. Vinogradov, E., Deschatelets, L., Lamoureux, M., Patel, G.B., Tremblay, T.L., Robotham, A., Goneau, M.F., Cummings-Lorbetskie, C., Watson, D.C., Brisson, J.R., Kelly, J.F. & Gilbert, M. Cell surface glycoproteins from *Thermoplasma acidophilum* are modified with an N-linked glycan containing 6-C-sulfofucose. *Glycobiology* **22**, 1256-1267 (2012).
20. Garegg, P.J. & Samuelsson, B. Novel reagent system for converting a hydroxy-group into an iodo-group in carbohydrates with inversion of configuration. *J. Chem. Soc., Chem. Commun.*, 978-980 (1979).
21. Mui, J.W.-Y., Souza, D.P.D., Saunders, E.C., McConville, M.J. & Williams, S.J. Remodeling of Carbon Metabolism during Sulfoglycolysis in Escherichia coli. *Appl. Environ. Microbiol.* **89**, e02016-02022 (2023).
22. Weinitschke, S., Denger, K., Cook, A.M. & Smits, T.H.M. The DUF81 protein TauE in *Cupriavidus necator* H16, a sulfite exporter in the metabolism of C2 sulfonates. **153**, 3055-3060 (2007).
23. 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.* **84**, 2901-2910 (2019).
24. 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. & Williams, S.J. A Sulfoglycolytic Entner-Doudoroff Pathway in *Rhizobium leguminosarum* bv. *trifolii* SRDI565. *Appl. Environ. Microbiol.* **86**, e00750-00720 (2020).
25. Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372**, 774-797 (2007).
26. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545-W549 (2010).

27. 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., Davies, G.J. & Williams, S.J. Molecular basis of sulfosugar selectivity in sulfoglycolysis. *ACS Cent. Sci.* **7**, 476-487 (2021).
28. Gutierrez-Patricio, S., Gonzalez-Pimentel, J.L., Miller, A.Z., Hermosin, B., Saiz-Jimenez, C. & Jurado, V. *Paracoccus onubensis* sp. nov., a novel alphaproteobacterium isolated from the wall of a show cave. *Int. J. Syst. Evol. Microbiol.* **71** (2021).
29. Mikosch, C.A.R.M., Denger, K., Schäfer, E.-M. & Cook, A.M. Anaerobic oxidations of cysteate: degradation via L-cysteate: 2-oxoglutarate aminotransferase in *Paracoccus pantotrophus*. *Microbiology (Reading, England)* **145**, 1153-1160 (1999).
30. Rein, U., Gueta, R., Denger, K., Ruff, J., Hollemeyer, K. & Cook, A.M. Dissimilation of cysteate via 3-sulfolactate sulfo-lyase and a sulfate exporter in *Paracoccus pantotrophus* NKNCYSA. *Microbiology (Reading, England)* **151**, 737-747 (2005).
31. 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. & Williams, S.J. Structural and Biochemical Insights into the Function and Evolution of Sulfoquinovosidases. *ACS Cent. Sci.* **4**, 1266-1273 (2018).
32. Ji, P., Zhang, Y., Wei, Y., Huang, H., Hu, W., Mariano, P.A. & Wang, W. Visible-Light-Mediated, Chemo- and Stereoselective Radical Process for the Synthesis of C-Glycoamino Acids. *Org Lett* **21**, 3086-3092 (2019).
33. Koketsu, M., Kuberan, B. & Linhardt, R.J. Stereoselective synthesis of the alpha-glycoside of a KDO "C"-disaccharide. *Org Lett* **2**, 3361-3363 (2000).
34. Sasidharan, K., Martinez-Vernon, A.S., Chen, J., Fu, T. & Soyer, O.S. A low-cost DIY device for high resolution, continuous measurement of microbial growth dynamics. *bioRxiv*, 407742 (2018).
35. Yoon, S.H., Ha, S.M., Lim, J., Kwon, S. & Chun, J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek* **110**, 1281-1286 (2017).
36. Blakeney, A.B. & Mutton, L.L. A simple colorimetric method for the determination of sugars in fruit and vegetables. *J. Sci. Food Agric.* **31**, 889-897 (1980).
37. Sörbo, B. in Methods Enzymol., Vol. 143 3-6 (Academic Press, 1987).
38. Brychkova, G., Yarmolinsky, D., Fluhr, R. & Sagi, M. The determination of sulfite levels and its oxidation in plant leaves. *Plant Sci.* **190**, 123-130 (2012).
39. Kurmanbayeva, A., Brychkova, G., Bekturova, A., Khozin, I., Standing, D., Yarmolinsky, D. & Sagi, M. in Plant Stress Tolerance: Methods and Protocols. (ed. R. Sunkar) 253-271 (Springer New York, New York, NY; 2017).
40. Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N. & Mann, M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* **11**, 319-324 (2014).
41. Harney, D.J., Hutchison, A.T., Hatchwell, L., Humphrey, S.J., James, D.E., Hocking, S., Heilbronn, L.K. & Larance, M. Proteomic Analysis of Human Plasma during Intermittent Fasting. *J. Proteome Res.* **18**, 2228-2240 (2019).
42. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896-1906 (2007).
43. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367-1372 (2008).
44. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M. & Cox, J. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731 (2016).

45. Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., Pérez, E., Uszkoreit, J., Pfeuffer, J., Sachsenberg, T., Yilmaz, S., Tiwary, S., Cox, J., Audain, E., Walzer, M., Jarnuczak, A.F., Ternent, T., Brazma, A. & Vizcaíno, J.A. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**, D442-d450 (2019).
46. Aragao, D., Aishima, J., Cherukuvada, H., Clarken, R., Clift, M., Cowieson, N.P., Ericsson, D.J., Gee, C.L., Macedo, S., Mudie, N., Panjikar, S., Price, J.R., Riboldi-Tunnicliffe, A., Rostan, R., Williamson, R. & Caradoc-Davies, T.T. MX2: a high-flux undulator microfocus beamline serving both the chemical and macromolecular crystallography communities at the Australian Synchrotron. *J. Synchrotron Radiation* **25**, 885-891 (2018).
47. Kabsch, W. Xds. *Acta Crystallogr., Section D: Biol. Crystallogr.* **66**, 125-132 (2010).
48. Evans, P.R. & Murshudov, G.N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D* **69**, 1204-1214 (2013).
49. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. & Read, R.J. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674 (2007).
50. Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A. & Wilson, K.S. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235-242 (2011).
51. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P. & Hassabis, D. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
52. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Žídek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., Figurnov, M., Cowie, A., Hobbs, N., Kohli, P., Kleywegt, G., Birney, E., Hassabis, D. & Velankar, S. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439-D444 (2022).
53. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D* **66**, 486-501 (2010).
54. Bricogne G., B.E., Brandl & M., F.C., Keller P., Paciorek W., Roversi P., Sharff A., Smart O.S., Vonrhein C., Womack T.O., Edn. Version 2.10.4. (Cambridge, United Kingdom: Global Phasing Ltd, 2017).
55. 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., Jain, S., Lewis, S.M., Arendall Iii, W.B., Snoeyink, J., Adams, P.D., Lovell, S.C., Richardson, J.S. & Richardson, D.C. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293-315 (2018).

Table 1. Key features of the *Paracoccus onubensis* (GenBank accession: Bioproject PRJNA993841) draft genome assembly.

	Strain Merri
Genome size	5,082,950
Number of contigs	74
GC %	60.1
Number of ORFs	4913
Number of putative genes	4966
Number of putative tRNA	49
Number of putative rRNA	3
Number of putative tmRNA	1
Number of genes assigned a function (%)	2847 (57%)

Table 2. Diffraction data and refinement statistics for *Po*SHPA.[a]

Data collection	
Space group	*P*1
Unit cell parameters (Å, °)	78.7, 82.5, 144.8 96.1, 90.0, 97.4
Resolution (Å)	47.98–2.50 (2.54–2.50)
No. of observations	219,370 (9,812)
No. of unique reflections	121,327 (5,943)
Completeness (%)	97.7 (95.9)
Redundancy	1.8 (1.7)
R_{merge} (%)	11.0 (67.0)
R_{pim} (%)	7.7 (61.2)
$CC_{1/2}$	0.991 (0.538)
Average $I/\sigma(I)$	4.9 (1.1)
Refinement	
R (%)	23.7 (32.5)
R_{free} (%)	28.2 (35.4)
No. (%) of reflections in test set	6,050 (5.0)
No. of protein molecules per asu	8
R.m.s.d bond length (Å)	0.01
R.m.s.d bond angle (°)	0.89
Average *B*-factors (Å^2)[b]	52.6
Protein molecules	52.8
Water molecules	44.9
Ramachandran plot[c]	
Residues other than Gly and Pro in:	
Most favored regions (%)	97.3
Additionally allowed regions (%)	2.4
Outliers (%)	0.3
PDB code	8TNF

[a] Values in parentheses are for the highest-resolution shell.

[b] Calculated by BAVERAGE in CCP4 Suite.[50]

[c] Calculated using MolProbity.[55]

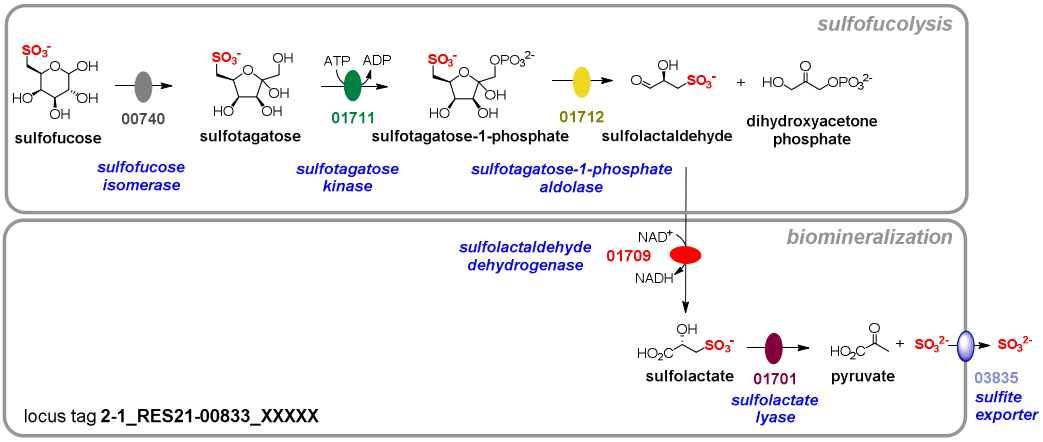

Fig 1. Proposed pathway for catabolism of sulfofucose in *Paracoccus onubensis* strain Merri.

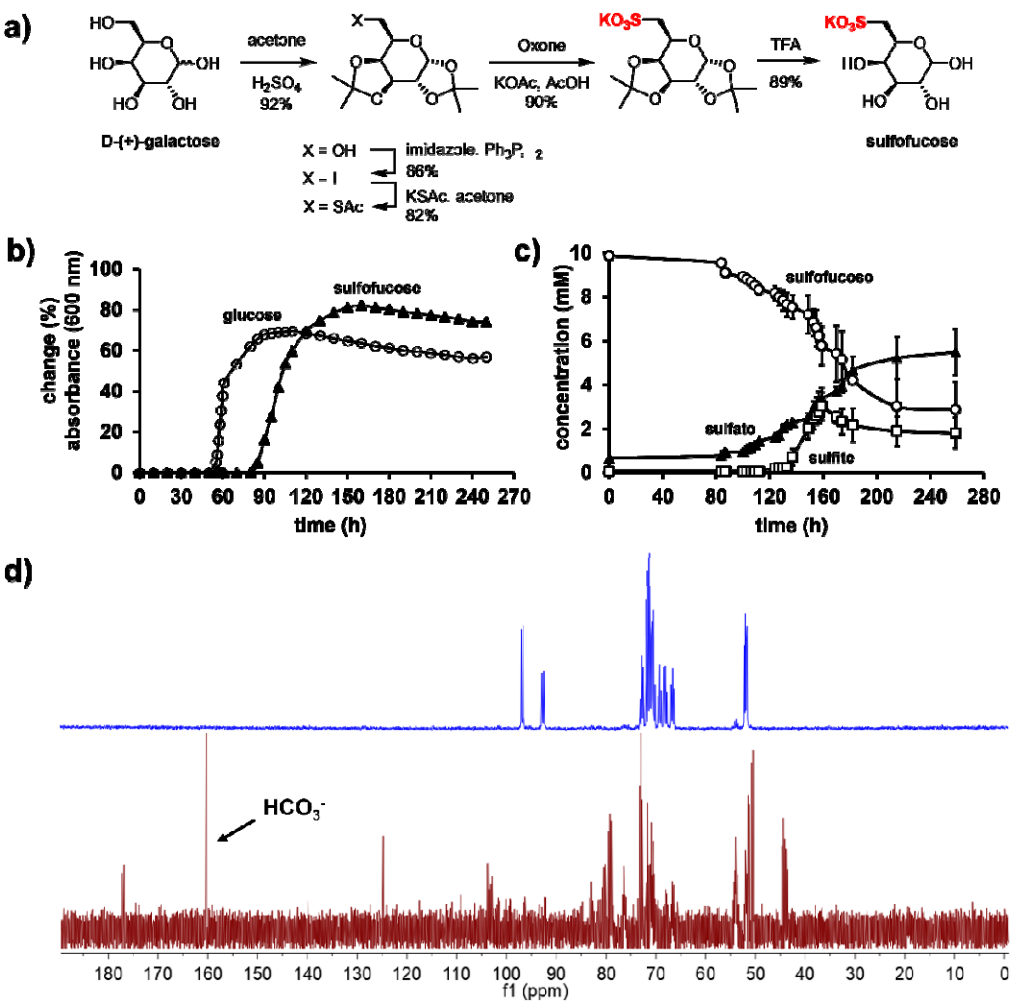

Fig 2. Characterization of sulfofucose-utilizing *Paracoccus onubensis* strain Merri. a) Chemical synthesis of sulfofucose. **b)** Growth of *P. onubensis* Merri on 10 mM sulfofucose or glucose in M9 media at 30 °C. **c)** *P. onubensis* Merri uses sulfofucose as a carbon source and produces sulfite. Plot showing concentration of sulfofucose (open circles), change in concentration of sulfite (open squares) and of sulfate (closed triangles). Analyses were performed in duplicate and errors are reported as standard deviations. **d)** ^{13}C NMR spectra (500 MHz) of (top) $^{13}C_6$-sulfofucose and (bottom) spent growth media from *P. onubensis* Merri grown on $^{13}C_6$-sulfofucose (10 mM).

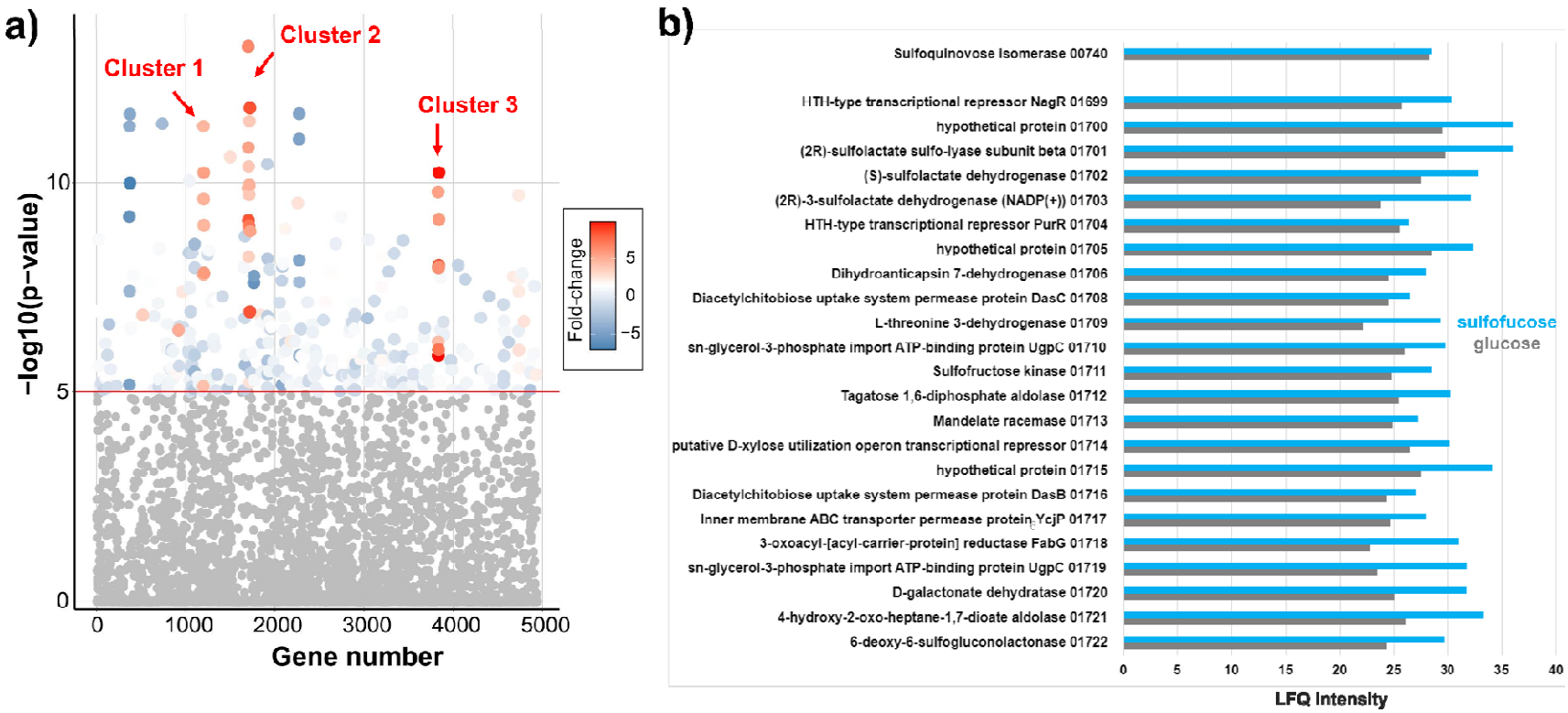

Fig 3. Comparative proteomics of *P. onubensis* Merri grown on sulfofucose or glucose as sole carbon source. a) Manhattan plot identified three differentially expressed clusters. **b)** Average LFQ intensity differences of three clusters' proteins in a) between the sulfofucose groups (blue) and the glucose groups (grey).

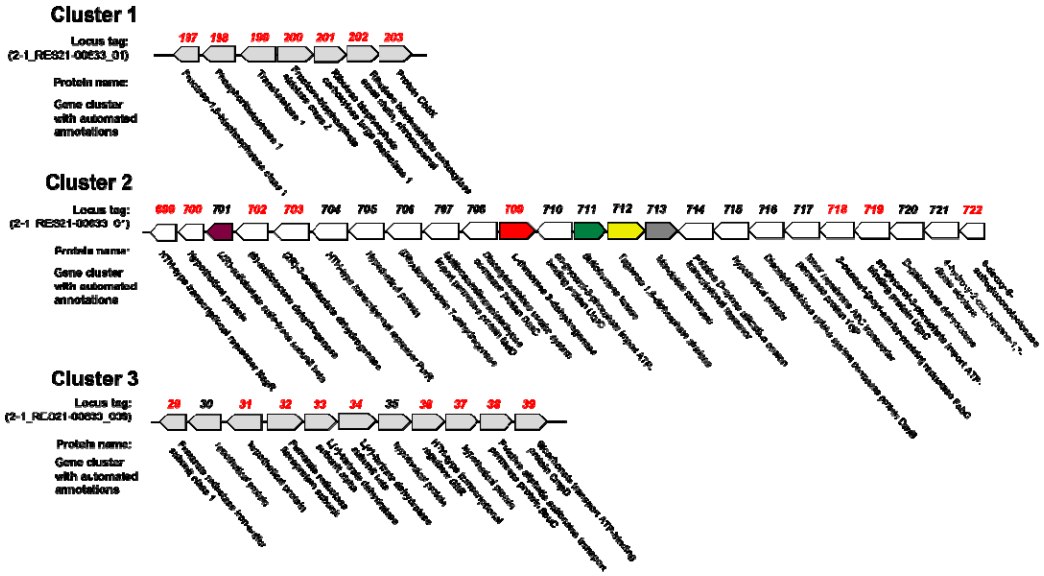

Fig 4. Genomic regions corresponding to three identified clusters in Fig 3a.

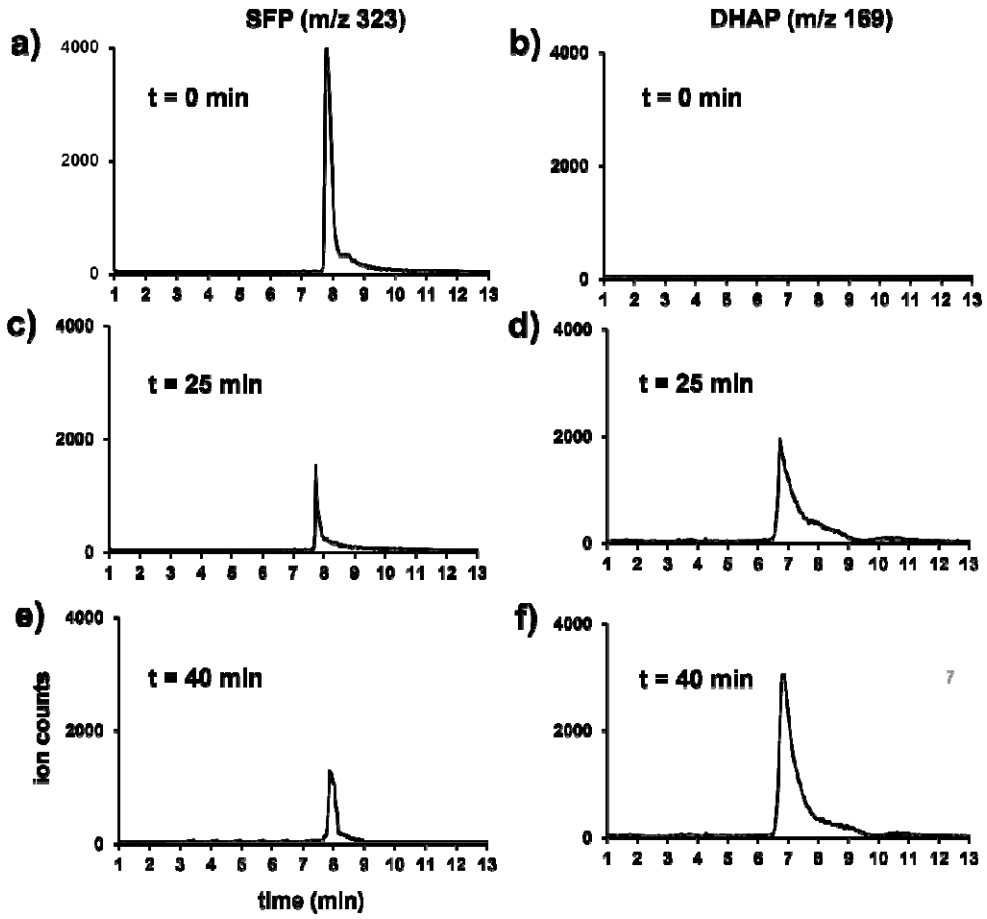

Fig 5. Cleavage of sulfofructose-1-phosphate by *Po*SHPA. HPLC-MS/MS elution profiles of the consumption of **(a,c,e)** sulfofructose-1-phosphate (SFP, *m/z* 323) catalyzed by aldolase (*Po*SHPA) to produce **(b,d,f)** dihydroxyacetone phosphate (DHAP, *m/z* 169), monitoring parent ions in ESI negative ion mode.

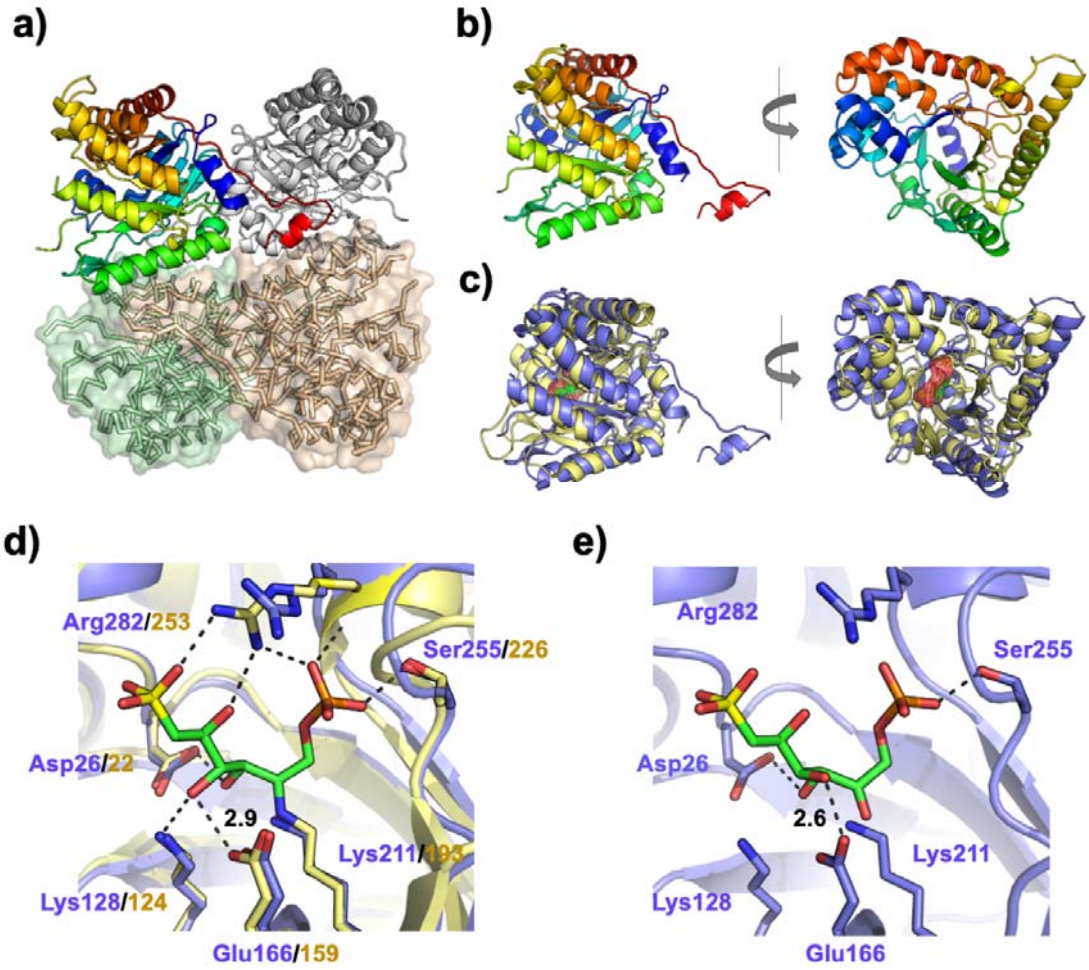

Fig 6. Crystal structure of *Po*SHPA and comparison with *Ec*YihT. a) Quaternary structure of *Po*SHPA. Chains A (rainbow) and B (grey) are displayed in cartoon representation while chains C (salmon) and D (light green) in ribbon and surface representation. **b)** Monomer of *Po*SHPA is shown in cartoon representation with rainbow color scheme, highlighting the extended C-terminus (red). **c)** Superposition of *Po*SHPA (slate) with *Ec*YihT (PDB entry 7NE2; pale yellow) shows the similar $(\alpha/\beta)_8$ fold with the main deviation being the extended C-terminal end of *Po*SHPA that contributes to the domain-swapped dimerization in *Po*SHPA. The SFP Schiff base is shown in the active site of *Ec*YihT. **d)** The superposed active sites of *Po*SHPA and *Ec*YihT reveal the similar active site architecture with the conserved residues binding phosphate and sulfonate groups. Lys211 in *Po*SHPA is predicted to be a Schiff-base forming residue. The distance between C4-OH of SFP and Oε1 of Glu159 (*Ec*YihT) is shown in Å. **e)** Model of sulfotagatose-1-phosphate, the likely substrate of *Po*SHPA, is shown in the active site of *Po*SHPA, displaying a possible hydrogen-bond interaction with Glu166. The distance between C4-OH of sulfotagatose-1-phosphate and Oε2 of Glu166 in the model is shown in Å.