ORIGINAL PAPER

Genome sequences of *Arthrobacter* spp. that use a modified sulfoglycolytic Embden–Meyerhof–Parnas pathway

Arashdeep Kaur[1,2] · Phillip L. van der Peet[1,2] · Janice W.-Y. Mui[1,2] · Marion Herisse[3] · Sacha Pidot[3] ·
Spencer J. Williams[1,2]

Abstract

Sulfoglycolysis pathways enable the breakdown of the sulfosugar sulfoquinovose and environmental recycling of its carbon and sulfur content. The prototypical sulfoglycolytic pathway is a variant of the classical Embden–Meyerhof–Parnas (EMP) pathway that results in formation of 2,3-dihydroxypropanesulfonate and was first described in gram-negative *Escherichia coli*. We used enrichment cultures to discover new sulfoglycolytic bacteria from Australian soil samples. Two gram-positive *Arthrobacter* spp. were isolated that produced sulfolactate as the metabolic end-product. Genome sequences identified a modified sulfoglycolytic EMP gene cluster, conserved across a range of other Actinobacteria, that retained the core sulfoglycolysis genes encoding metabolic enzymes but featured the replacement of the gene encoding sulfolactaldehyde (SLA) reductase with SLA dehydrogenase, and the absence of sulfoquinovosidase and sulfoquinovose mutarotase genes. Excretion of sulfolactate by these *Arthrobacter* spp. is consistent with an aerobic saprophytic lifestyle. This work broadens our knowledge of the sulfo-EMP pathway to include soil bacteria.

Abbreviations

ABC	ATP-binding cassette
DHPS	2,3-Dihydroxypropanesulfonate
NAD(P)H	Reduced nicotinamide adenine dinucleotide (phosphate)
SF	Sulfofructose
SL	Sulfolactate
SLA	Sulfolactaldehyde
SQ	6-Deoxy-6-sulfo-D-glucose
SQase	Sulfoquinovosidase
SQDG	Sulfoquinovosyl diacylglycerol
SQGro	Sulfoquinovosyl glycerol
Sulfo-ED	Sulfoglycolytic Entner–Doudoroff
Sulfo-EMP	Sulfoglycolytic Embden–Meyerhof–Parnas
Sulfo-SFT	Sulfoglycolytic sulfofructose transaldolase
Sulfo-TK	Sulfoglycolytic transketolase
Sulfo-SMO	Sulfoglycolytic sulfoquinovose monooxygenase
TSUP	4-Toluene sulfonate uptake permease

✉ Spencer J. Williams
sjwill@unimelb.edu.au

1   School of Chemistry, University of Melbourne, Parkville, VIC 3010, Australia

2   Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, VIC 3010, Australia

3   Department of Microbiology and Immunology, University of Melbourne, Peter Doherty Institute for Infection and Immunity, Melbourne, VIC 3000, Australia

Introduction

Sulfoquinovose (SQ; 6-deoxy-6-sulfo-D-glucose) is a sulfosugar produced by photosynthetic organisms (Goddard-Borger and Williams 2017). It is primarily found as the headgroup of the sulfoglycolipid sulfoquinovosyl diacylglycerol (SQDG) in photosynthetic tissues and membranes in plants, algae and cyanobacteria (Benson et al. 1959; Goddard-Borger and Williams 2017). The annual global production of SQ is estimated at 10^{10} tons per annum (Harwood and Nicholls 1979), and thus, the degradation of this sulfosugar is an important arm of the global biogeochemical cycle. The degradation of sulfoquinovose occurs through pathways of sulfoglycolysis and provides access to its constituent carbon and generates ATP and reducing equivalents (NADH/

NADPH) (Benson and Shibuya 1961; Snow et al. 2021). The sulfoglycolytic Embden–Meyerhof–Parnas (sulfo-EMP) (Denger et al. 2014; Sharma et al. 2021), Entner–Doudoroff (sulfo-ED) (Felux et al. 2015), and sulfofructose transaldolase (sulfo-SFT) (Frommeyer et al. 2020; Liu et al. 2020) pathways cleave the 6-carbon chain of SQ into two 3-carbon fragments, one of which is utilized in primary metabolism, while the other containing the sulfonate group is converted to either sulfolactate (SL) or 2,3-dihydroxypropanesulfonate (DHPS) and excreted. The sulfoglycolytic sulfofructose transketolase (sulfo-TK) pathway uses four of the six carbons in primary metabolism while the last two are excreted as isethionate (Liu et al. 2021). The sulfoglycolytic sulfoquinovose monooxygenase (sulfo-SMO) pathway results in cleavage of the C–S bond of sulfoquinovose and leads to production of glucose, and thus enables the complete breakdown of the SQ molecule (Liu et al. 2021; Sharma et al. 2022). The genes encoding these sulfoglycolytic pathways are found within clusters that typically contain genes encoding proteins for the import of SQ or its glycosides, the export of the end-products, SL or DHPS, and a specialized glycoside hydrolase termed a sulfoquinovosidase (SQase) (Speciale et al. 2016; Abayakoon et al. 2018a; Liu et al. 2021) that can cleave SQ glycosides to release SQ that can undergo sulfoglycolysis.

We report here the use of sequential enrichment culturing using minimal media containing SQ as sole carbon source to isolate new sulfoglycolytic bacteria from soil. We identify two *Arthrobacter* sp. strains, AK01 and AK04, and demonstrate that they grow on SQ and secrete SL into the growth media. We present the draft genome sequences of AK01 and AK04 that reveals that these *Arthrobacter* sp. contain a gene cluster encoding a sulfoglycolytic Embden–Meyerhof–Parnas pathway that differs from the prototypical pathway of *E. coli* through the lack of an identifiable candidate SQase, the replacement of SLA reductase with SLA dehydrogenase, and the presence of ABC transporters and TauE permeases, variations that are present within other sequenced Actinobacteria.

Materials and methods

Bacterial growth media

Growth media were prepared using M9 minimal salts media (2 mL), trace metal solution (0.1 mL), and vitamin solution (0.01 mL) and contained 5 mM sulfoquinovose (SQ) as sole carbon source, made up to a final volume of 10 mL with water. M9 minimal salts media contain 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).

Isolation of *Arthrobacter* sp.

Arthrobacter sp. strains AK01 and AK04 were isolated from enriched culture, obtained from soil of the Botany Systems Garden (University of Melbourne).

Two soil samples were collected and approximately 1 g of soil was suspended in 5 mL of sterilized growth media containing 5 mM SQ as a sole carbon source. The culture was incubated at 30 °C for 4 days with agitation at 250 rpm. A subsample (100 μL) was transferred into fresh vitamin-supplemented M9 media and grown for a further 4 days. This step was repeated four times and after outgrowth of the final culture for 4 days, cells were plated onto LB agar plates (10 g/L tryptone, 5 g/L NaCl, 5 g/L yeast, 15 g/L agar) and incubated overnight at 30 °C in dark. Single colonies were picked and inoculated into fresh vitamin-supplemented M9 media containing 5 mM SQ and incubated at 28 °C while shaking at 250 rpm using a Ratek orbital mixer incubator. Once the cultures were visibly turbid, cells were again plated onto LB agar, incubated overnight at 30 °C in dark and single colonies picked and inoculated again into fresh vitamin-supplemented M9 media containing 5 mM SQ. Once cultures were visibly turbid, frozen stocks were prepared by diluting to 10% glycerol and freezing at –80 °C. Cell morphology was examined using scanning electron microscopy. Genomic DNA was isolated using the GenElute DNA extraction kit (Sigma) with inclusion of lysozyme and RNAase.

Phenotypic assays

Bacteria were grown in vitamin-supplemented M9 minimal media containing $^{13}C_6$-SQ (7.7 mM) in culture tubes at 30 °C for 1 day with agitation at 250 rpm. After cultures became visibly turbid, the cells were sedimented by centrifugation at 9000*g* for 10 min (using Sigma laborzentrifugen model 1-15, rotor 12124) and the supernatant was diluted with 50% D_2O and transferred to a 5 mm NMR tube. ^{13}C-NMR spectra were acquired using a 500 MHz instrument and are shown in Fig. 1a.

Genome sequence, assembly, and annotation

Genomic DNA was sequenced using an Illumina NextSeq at the Peter Doherty Institute for Infection and Immunity, Parkville, Victoria, Australia. DNA was prepared for
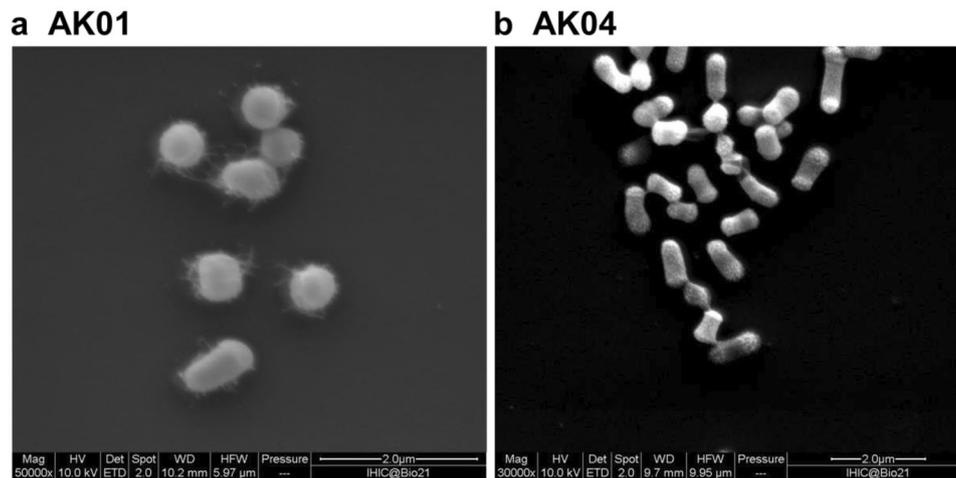
Fig. 1 Scanning electron microscopy image of *Arthrobacter* sp. **a** strain AK01, **b** strain AK04. Cell morphology was examined using a scanning electron microscope (Quanta 200 ESEM). Cells were grown in LB media for 3 days, fixed in 0.05% glutaraldehyde in 0.1 M phosphate buffer (pH 7.4), then in 2.5% glutaraldehyde in 0.1 M phosphate buffer (pH 7.4) and allowed to react for 20 min. Fixed cells were adhered onto poly-lysine-coated slides and rinsed with water 3 times, then dehydrated by soaking in an ascending ethanol gradient (20–100%). The sample was critical point dried using a Leica CPD3000 and gold coated to thickness of 5 nm using Safematic CCU-010 compact coating unit. Images are at approximately 50,000×magnification with scale bar shown

Table 1 ^{13}C-NMR (125 MHz) data of ^{13}C$_3$-SL produced as metabolite from ^{13}C$_6$-SQ

Chemical shift (d, ppm)	Multiplicity	Coupling constant (Hz)	Assignment
55.19	d	$^1J_{C1,C2} = 36.9$	C3
62.42	s	NA	–
68.85	dd	$^1J_{C2,C3} = 52.7$ $^1J_{C1,C2} = 38.0$	C2
72.01	s	NA	–
178.69	d	$^1J_{C2,C3} = 54.6$	C1

Samples contain 50% D$_2$O to allow frequency lock

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 × . Draft genomes were assembled using Shovill v1.1.0 (https://github.com/tseemann/shovill) and annotated using Prokka v1.14.5 (Seemann 2014). GC percentage and ANI calculations were performed using ANI calculator (https://www.ezbiocloud.net/tools/ani) (Yoon et al. 2017). Assembled genomes have been deposited at the NCBI (GenBank accession: AK01, SAMN23041292; AK04, SAMN23041293). General features for isolated bacteria are reported in Table 1. The protein sequences of putative sulfoglycolysis proteins in AK01 and AK04 were used to search against the NCBI non-redundant database using BLASTp. Percentage identities for key sulfo proteins are given in Table S2.

Discovery of related sulfoglycolytic operons

Sequences for *E. coli* sulfoquinovosidase (NP_418314.1, locus tag b3878), SQ mutarotase (NP_418315.3, locus tag b3879), SQ isomerase (NP_418316.4, locus tag b3880), SF kinase (NP_418319.2, locus tag b3883), SFP aldolase (NP_418317.1, locus tag b3881), SLA reductase (NP_418318.1, locus tag b3882) and sulfo-EMP regulator (NP_418320.2, locus tag b3884) were submitted separately as queries to the NCBI BLASTp tool. The database searched was the non-redundant protein sequence (nr) database, with *E. coli* (taxid: 562) sequences excluded. Standard algorithm parameters were used, except the maximum target sequences was set to 10,000. The results were filtered, with only protein sequences with E-value ≤ 5.41e-44 retained. The corresponding nucleotide accession numbers for each protein from all seven searches were extracted, and the seven lists combined and duplicates removed to give a list of candidate genome sequences. This list was converted into a MultiGeneBLAST reference library and searched using the *E. coli* sulfo-EMP gene cluster as a query. Scripts for this pipeline are available on GitHub (https://github.com/jmui-unimelb/Gene-Cluster-Search-Pipeline). Gene clusters found using this workflow were screened for clusters that contained a putative SQ isomerase, a putative SF kinase and a putative SFP aldolase, but lacked a homologous sulfoquinovosidase.

Results and discussion

Isolation and characterization of sulfoglycolytic bacteria

Bacteria able to grow on SQ as sole carbon source were selected by using soil samples (collected from the University of Melbourne, Parkville campus) to inoculate a vitamin-enriched minimal media containing SQ as sole carbon source. Sequential subculturing into fresh SQ-minimal media, followed by plating onto LB agar and then regrowth in SQ-minimal media led to isolation of strains AK01 and AK04 that possessed a short rod-like appearance (Fig. 1). Strains AK01 and AK04 grew robustly on SQ with peak growth rates of 0.129 and 0.081 A_{600}/min and achieved stationary phase after approximately 25 and

40 h, respectively (Fig. 2a, c). The absorbance at stationary phase for cultures grown on SQ were approximately half of that for cultures grown on equimolar glucose. ^{13}C-NMR analysis of culture medium of AK01 and AK04 grown on ^{13}C$_6$-SQ (7.7 mM) media gave three signals corresponding to ^{13}C$_3$-SL (Fig. 2b, d, Table 1). ^{13}C-NMR analysis of spent culture confirmed that substrate SQ is completely consumed by bacteria and metabolized to SL.

To investigate the genetic basis for SQ consumption by AK01 and AK04, DNA extracted from these bacteria were sequenced using the Illumina NextSeq platform. Table 2 shows the key features of the two draft genomes. On the basis of 16S rRNA gene sequence analysis the organisms were assigned as *Arthrobacter* sp. The 16S rRNA genes of the two strains were 98.8% identical over 1520 bp, which suggests that these organisms are of the same species (Kim et al. 2014). However, average nucleotide identity (ANI)
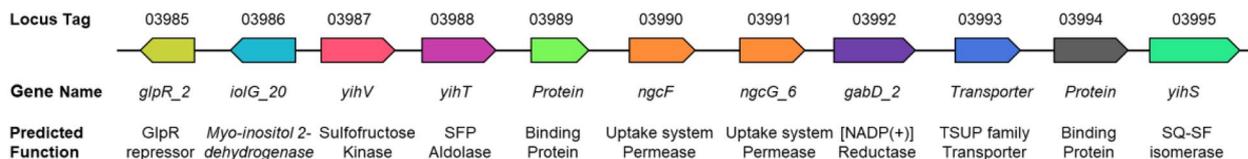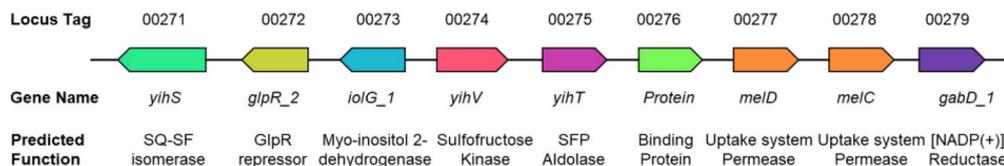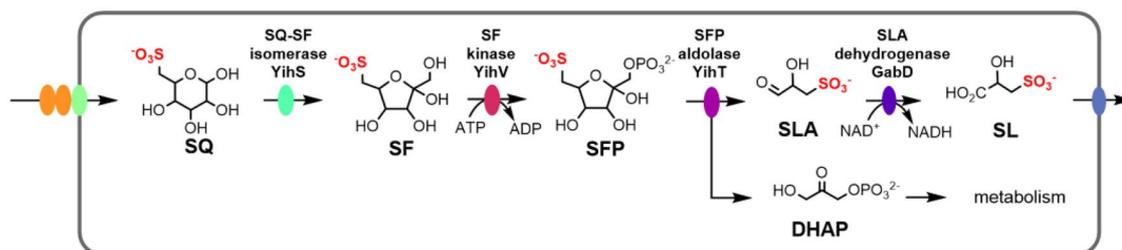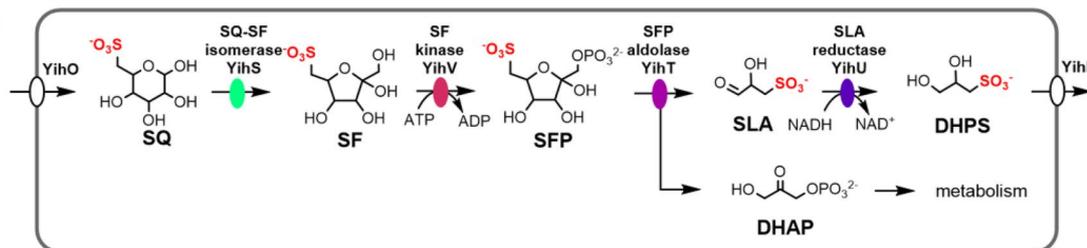

Fig. 2 Proposed sulfoglycolytic Embden–Meyerhof–Parnas (sulfo-EMP) pathway for *Arthrobacter* spp. **a** Gene cluster encoding the sulfo-EMP pathway for *Arthrobacter* sp. AK04. **b** Gene cluster encoding the sulfo-EMP pathway for *Arthrobacter* sp. AK01. **c** Proposed sulfo-EMP pathway for *Arthrobacter* spp. **d** Comparison with EMP pathway for *E. coli* K12

Table 2 Key features of the AK01 (GenBank accession: SAMN23041292) and AK04 (GenBank accession: AK01, SAMN23041293) draft genome assemblies

	AK01	AK04
Genome size	5,105,913	4,700,363
Number of contigs	169	130
GC %	62.9	65.5
Number of ORFs	4734	4318
Number of putative genes	4669	4261
Number of putative tRNA	59	53
Number of putative rRNA	5	3
Number of putative tmRNA	1	1
Number of genes assigned a function (%)	2478 (53%)	2219 (52%)

between the two organisms is only 76.5%, conflicting with the 16S rDNA gene results and suggesting that these two organisms differ sufficiently to be considered separate species (Jain et al. 2018).

Genomic features related to SQ metabolism

Genome analysis of *Arthrobacter* spp. AK01 and AK04 revealed a cluster of genes that were assigned as encoding SQ degradation through the sulfo-EMP pathway (Fig. 3a, b). The prototypical sulfo-EMP pathway was identified in *E. coli* and involves a 10-gene cluster (*yihOPQRSTUV*, *csqR*) (Denger et al. 2014). In *E. coli* these genes encode a transcription factor (CsqR) (Shimada et al. 2019), putative transmembrane proteins for the import of SQ and export of the end-product of the pathway, DHPS (YihO, YihP). The enzymatic steps involve a sulfoquinovosidase (YihQ) for cleavage of SQ glycosides (Speciale et al. 2016), sulfoquinovose mutarotase (YihR) for interconversion of SQ anomers (Abayakoon et al. 2018b), SQ-sulfofructose (SF) isomerase (YihS) that interconverts SQ, SF and sulforhamnose (Sharma et al. 2021), an ATP-dependent sulfofructose kinase (YihV) that converts SF to SF-1-phosphate (SFP) (Sharma et al. 2021), SFP aldolase (YihT) which converts SFP to SLA and dihydroxyacetone phosphate (Sharma et al. 2021), and an NADH-dependent SLA reductase (YihU) to convert SLA to DHPS (Sharma et al. 2020), which is excreted into the growth media (Fig. 3d).

Both *Arthrobacter* sp. contained genes encoding SQ-SF isomerase (YihS), SF kinase (YihV), and SFP aldolase (YihT). Consistent with the excretion of SL into the growth media, both strains lacked an SLA reductase homologue, but instead contained an SLA dehydrogenase homologue, annotated as GabD. GabD homologues within sulfoglycolytic gene clusters have been identified for bacteria that utilize the sulfo-ED (Felux et al. 2015; Li et al. 2020) and sulfo-SFT (Frommeyer et al. 2020; Liu et al. 2020) pathways, and

in *Pseudomonas putida* SQ1 there is an $NAD^+/NADP^+$-dependent SLA dehydrogenase (Felux et al. 2015). Recently, a sulfo-EMP pathway in *Bacillus urumqiensis* was identified that contained a GabD homologue, SlaB, however, it was unable to be recombinantly expressed and its activity is therefore unknown (Liu et al. 2021); we note that SlaB from *Desulfovibrio* sp strain DF1 could be recombinantly expressed and was confirmed to be an NADH-dependent, SL-producing SLA dehydrogenase (Burrichter et al. 2018). A proposed pathway for SQ metabolism in these *Arthrobacter* spp. is shown in Fig. 3c. AK01 and AK04 represent the first characterized examples of sulfoglycolytic bacteria that use a sulfo-EMP pathway but with an SLA dehydrogenase instead of SLA reductase, as described for the prototypical sulfo-EMP pathway in *E. coli* (Denger et al. 2014). *E. coli* and related Enterobacteriacae that contain SLA reductases are facultative anaerobes, and the presence of a reducing pathway for excretion of DHPS may support their anaerobic lifestyle. On the other hand, *Arthrobacter* are normally considered aerobes (Jones and Keddie 2006) (although anaerobic *Arthrobacter* have been described (Eschbach et al. 2003)), which is consistent with an oxidative pathway for excretion of SL.

SQ import and DHPS/SL export is poorly characterized, but several different strategies have been identified across diverse sulfoglycolytic bacteria. Both AK01 and AK04 contain genes annotated as ABC transporter cassettes and solute binding proteins. ABC transporter systems have been identified in *Agrobacterium tumefaciens* C58 (which uses the sulfo-SMO pathway) (Sharma et al. 2022) and *R. leguminosarum* SRDI858 (which utilizes a sulfo-ED pathway) (Li et al. 2020). The *A. tumefaciens* solute binding protein binds SQGro with high affinity (Sharma et al. 2022). We, therefore, propose that these *Arthrobacter* isolates utilize the solute binding protein and an ABC transporter system to import SQ or its glycosides. Strain AK04 also contains a TauE permease (a member of the 4-toluene sulfonate uptake permease (TSUP) system) (Shlykov et al. 2012). TSUP proteins are poorly characterized permeases that are suggested to be involved with the transport of sulfur-containing organic compounds. TauE of *Cupriavidus necator* H16 is proposed to be involved in the export of sulfolactate (Weinitschke et al. 2007), a function that is consistent with the excretion of SL by AK04. The absence of an obvious permease candidate in the AK01 gene cluster suggests that another protein may adopt this function in this strain.

Other key differences with *E. coli* includes the lack of identifiable SQ mutarotase and SQase encoding genes. Interestingly, the isolated strains could grow on the simple SQ glycoside, methyl α-sulfoquinovoside. This suggests that they may harbor an unidentified SQase without homology to known SQases. Finally, both organisms contained a gene encoding an IolG homologue, of unknown function. IolG proteins are
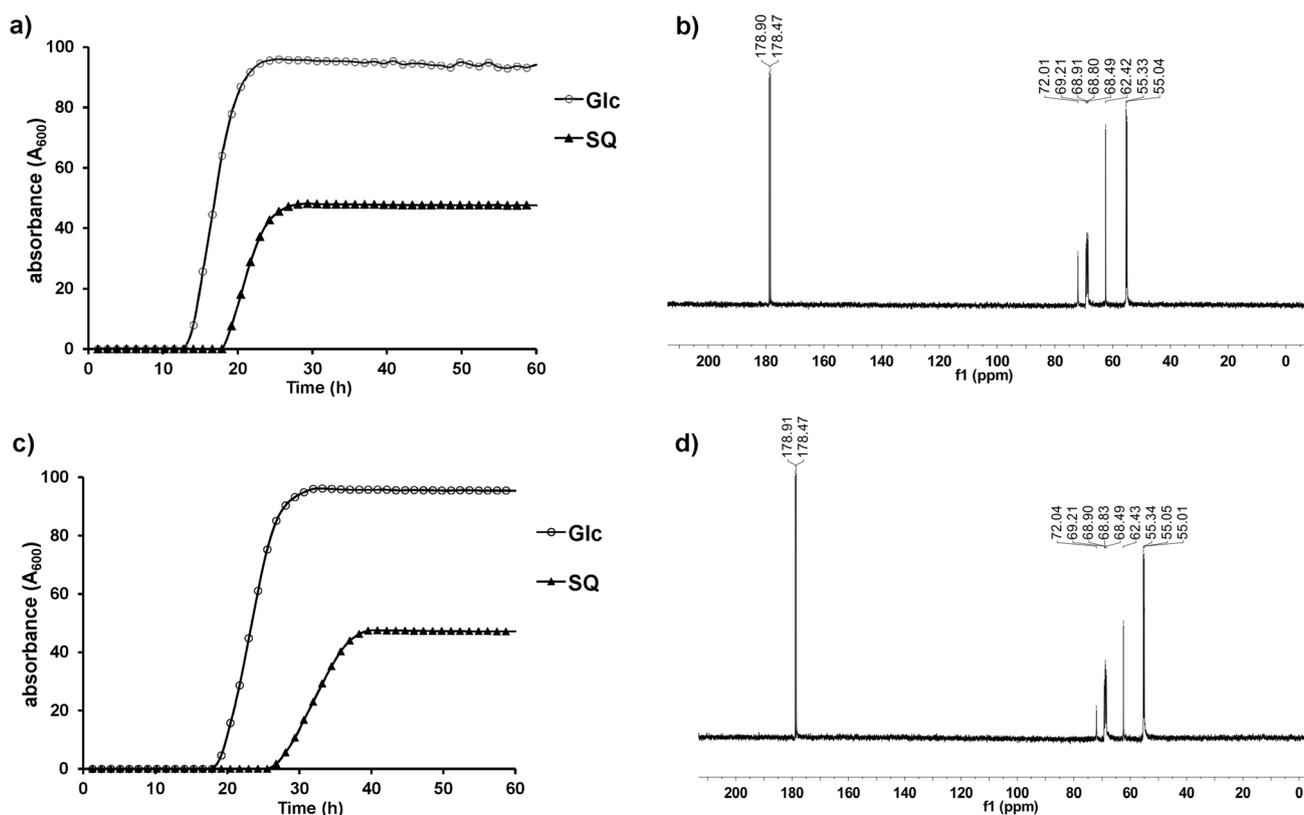
Fig. 3 Growth curves of *Arthrobacter* strains **a** AK01 and **c** AK04 grown on minimal salts media containing 5 mM glucose or SQ. ^{13}C-NMR (500 MHz) spectra of spent culture media of *Arthrobacter* strains **b** AK01 and **d** AK04 grown on ^{13}C$_6$-SQ (7.7 mM)

NAD(P)-dependent oxidoreductases of the Gfo/Idh/MocA family, and catalyze oxidation of the hydroxyl groups of pyranose and inositol rings (Taberman et al. 2016). Members of this family include inositol dehydrogenase (Idh), which forms 2-keto-*myo*-inositol (2-inosose) from *myo*-inositol (Ramaley et al. 1979; Yoshida et al. 2006), glucose-6-phosphate dehydrogenase (G6PD), which forms 6-phosphogluconolactone from glucose-6-phosphate (Rowland et al. 1994), and levoglucosan dehydrogenase, which forms 3-keto-levoglucosan from levoglucosan (1,6-anhydro-β-D-glucose) (Sugiura et al. 2018; Kuritani et al. 2020). Inositol dehydrogenase IolG from *Bacillus subtilis* has maximal activity on *myo*-inositol and possesses activity on D-glucose and D-xylose, and produces D-gluconolactone from the former (Ramaley et al. 1979). Possibly, the *Arthobacter* spp. IolG homologues may convert SQ to SGL, which coincidentally is an intermediate in the sulfo-ED pathway (Felux et al. 2015), although the downstream sulfo-ED genes are missing in these organisms. An alternative possibility is suggested by the Gfo/Idh/MocA family member DgpA from the intestinal bacterium PUE. DgpA catalyzes the oxidation of the 3-hydroxyl group of the C-glycoside puerarin, facilitating the elimination of the aglycon and formation of 3-keto-2-hydroxyglucal (1,5-anhydro-D-*erythro*-hex-1-en-3-ulose) (Nakamura et al. 2020). A similar process applied to an SQ glycoside could facilitate the cleavage of the glycoside in the absence of an SQase and would give rise to the corresponding 6-sulfo-3-keto-2-hydroxyglucal, the fate of which is uncertain. However, we were unable to identify possible companion genes encoding proteins that would be required to enable the elimination/isomerization/reduction to SQ.

A search for organisms with gene clusters related to strains AK01 and AK04 led to identification of other *Arthrobacter* strains with syntenic or closely related sulfo-EMP gene clusters (Fig. 4). Other *Arthrobacter* spp. were identified that contained gene clusters with architectures identical to AK01; and both *Arthrobacter* spp. and *Pseudarthrobacter* spp. were identified with gene clusters identical to AK04. Non-identical but closely syntenic sulfo-EMP gene clusters were observed in select Actinobacteria including other *Arthrobacter* spp. and *Microbacterium* sp. No. 7 (both order Micrococcales), and a more distantly related cluster in the actinobacteria *Acrocarpospora corrugata* (order Streptosporangiales) and *Phytohabitans houttuyneae* (order Micromonosporales).
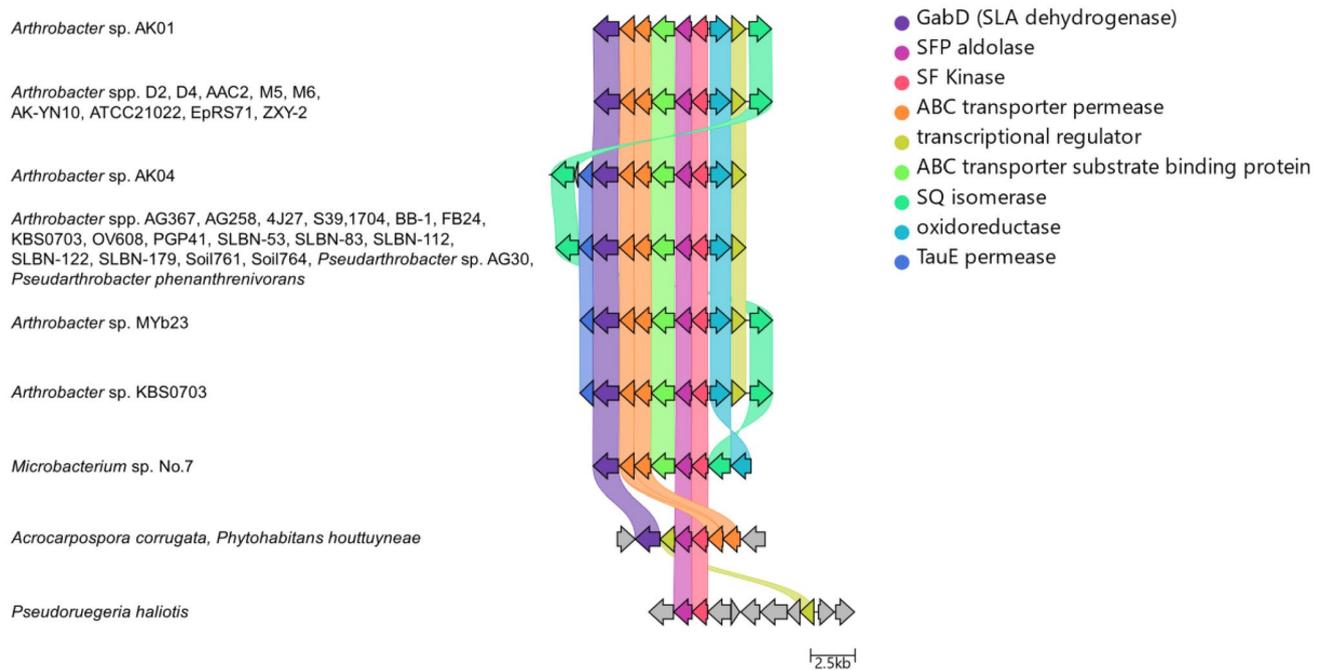
Fig. 4 Distribution and architecture of sulfo-EMP gene clusters in *Arthrobacter* and related organisms. Syntenic relationship of sulfo-EMP gene clusters in *Arthrobacter* sp. AK01 and AK04 with homologous gene clusters. Colored links indicate ≥ 30% protein sequence similarity. Genome accession codes: *Arthrobacter* sp. D2 (LUKB01000109.1), *Arthrobacter* sp. D4 (LUKC01000078.1), *Arthrobacter* sp. AAC2 (JAAGBD010000014.1), *Arthrobacter* sp. M5 (LVCB01000107.1), *Arthrobacter* sp. M6 (LVCC01000103.1), *Arthrobacter* sp. AK-YN10 (AVPD02000157.1), *Arthrobacter* sp. ATCC 21,022 (CP014196.1) *Arthrobacter* sp. EpRS71 (LNUV01000003.1), *Arthrobacter* sp. ZXY-2 (CP017421.1), *Arthrobacter* sp. AG367 (VIVE01000010.1), *Arthrobacter* sp. AG258 (SOBI01000009.1), *Arthrobacter* sp. 4J27 (CAQI01000048.1), *Arthrobacter* sp. S39 (SIHX01000007.1), *Arthrobacter* sp.1704 (SOBD01000016.1), *Arthrobacter* sp. BB-1 (VDEV01000010.1), *Arthrobacter* sp. FB24 (CP000454.1),

Arthrobacter sp. KBS0703 (MVDG02000001.1), *Arthrobacter* sp. OV608 (FOEZ01000003.1), *Arthrobacter* sp. PGP41 (CP026514.1), *Arthrobacter* sp. SLBN-53 (VFMZ01000001.1), *Arthrobacter* sp. SLBN-83 (VFMX01000001.1), *Arthrobacter* sp. SLBN-112 (VFMU01000001.1), *Arthrobacter* sp. SLBN-122 (VFMS01000001.1), *Arthrobacter* sp. SLBN-179 (VFNR01000001.1), *Arthrobacter* sp. Soil761 (LMSF01000007.1), *Arthrobacter* sp. Soil764 (LMSI01000008.1), *Pseudarthrobacter phenanthrenivorans* (CP002379.1), *Pseudarthrobacter phenanthrenivorans* (RBNH01000003.1), *Pseudarthrobacter phenanthrenivorans* (VHJD01000009.1), *Pseudarthrobacter* sp. AG30 (QEHL01000024.1), *Arthrobacter* sp. MYb23 (PCPR01000010.1), *Arthrobacter* sp. KBS0703 (MVDG02000001.1), *Microbacterium* sp. No. 7 (CP012697.1), *Acrocarpospora corrugata* (BLAD01000050.1), *Phytohabitans houttuyneae* (BLPF01000004.1), *Pseudoruegeria haliotis* (PVTD01000003.1)

Conclusions

Two sulfoglycolytic soil bacteria belonging to the *Arthrobacter* genus (strains AK01 and AK04) were isolated from soil by enrichment culture involving growth on SQ. The stationary phase optical density of these bacteria when grown on SQ was approximately half that of growth on glucose, consistent with utilization of only three of the six carbons of SQ. Both possessed a variant of the sulfo-EMP pathway that uses an SLA dehydrogenase to produce SL that is secreted into the growth media, and which is proposed to arise from an SLA dehydrogenase (GabD). SL in turn becomes available for other members of the microbial community that specialize in its metabolism (Rein et al. 2005; Cook et al. 2006; Denger et al. 2009, 2012; Denger and Cook 2010). Prior to this work, the only sulfo-EMP pathway bacteria that have been characterized were a range of

E. coli strains, which produce DHPS through the action of SLA reductase (Denger et al. 2014). Notably, strains AK01 and AK04 lack genes assigned as encoding an SQase, but nonetheless could grow on MeSQ, indicating the presence of a non-specific SQase, a novel SQase that is not readily identified by sequence homology, or a novel lyase pathway. This study highlights the core genes required for sulfoglycolysis (YihS, YihT, YihU, YihV) and constitutes the first examples of sulfo-EMP bacteria isolated from soil.

Availability of data and materials The datasets generated and/or analyzed during the current study are available in the GenBank database with accession codes SAMN23041292 (strain AK01) and SAMN23041293 (strain AK04).

Code availability Scripts used for bioinformatic search pipeline are available on the GitHub repository, https://github.com/jmui-unimelb/Gene-Cluster-Search-Pipeline.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

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

Abayakoon P et al (2018b) Discovery and characterization of a sulfoquinovose mutarotase using kinetic analysis at equilibrium by exchange spectroscopy. Biochem J 475:1371–1383. https://doi.org/10.1042/bcj20170947

Benson AA, Shibuya I (1961) Sulfocarbohydrate metabolism. Fed Proc 20:79

Benson AA, Daniel H, Wiser R (1959) A sulfolipid in plants. Proc Natl Acad Sci USA 45:1582–1587. https://doi.org/10.1073/pnas.45.11.1582

Burrichter A et al (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. https://doi.org/10.3389/fmicb.2018.02792

Cook AM, Denger K, Smits TH (2006) Dissimilation of C3-sulfonates. Arch Microbiol 185:83–90. https://doi.org/10.1007/s00203-005-0069-1

Denger K, Cook AM (2010) Racemase activity effected by two dehydrogenases in sulfolactate degradation by *Chromohalobacter salexigens*: purification of (S)-sulfolactate dehydrogenase.

Microbiology 156:967–974. https://doi.org/10.1099/mic.0.034736-0

Denger K, Mayer J, Buhmann M, Weinitschke S, Smits TH, Cook AM (2009) Bifurcated degradative pathway of 3-sulfolactate in *Roseovarius nubinhibens* ISM via sulfoacetaldehyde acetyltransferase and (*S*)-cysteate sulfolyase. J Bacteriol 191:5648–5656. https://doi.org/10.1128/jb.00569-09

Denger K, Huhn T, Hollemeyer K, Schleheck D, Cook AM (2012) Sulfoquinovose degraded by pure cultures of bacteria with release of C₃-organosulfonates: complete degradation in two-member communities. FEMS Microbiol Lett 328:39–45. https://doi.org/10.1111/j.1574-6968.2011.02477.x

Denger K et al (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

Eschbach M, Möbitz H, Rompf A, Jahn D (2003) Members of the genus *Arthrobacter* grow anaerobically using nitrate ammonification and fermentative processes: anaerobic adaptation of aerobic bacteria abundant in soil. FEMS Microbiol Lett 223:227–230. https://doi.org/10.1016/S0378-1097(03)00383-5%JFEMSMicrobiologyLetters

Felux AK, Spiteller D, Klebensberger J, Schleheck D (2015) Entner-Doudoroff pathway for sulfoquinovose degradation in *Pseudomonas putida* SQ1. Proc Natl Acad Sci USA 112:E4298-4305. https://doi.org/10.1073/pnas.1507049112

Frommeyer B et al (2020) Environmental and intestinal phylum firmicutes bacteria metabolize the plant sugar sulfoquinovose via a 6-deoxy-6-sulfofructose transaldolase pathway. Science 23:101510. https://doi.org/10.1016/j.isci.2020.101510

Goddard-Borger ED, Williams SJ (2017) Sulfoquinovose in the biosphere: occurrence, metabolism and functions. Biochem J 474:827–849. https://doi.org/10.1042/BCJ20160508

Harwood JL, Nicholls RG (1979) The plant sulpholipid—a major component of the sulphur cycle. Biochem Soc Trans 7:440–447

Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S (2018) High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:5114. https://doi.org/10.1038/s41467-018-07641-9

Jones D, Keddie RM (2006) The Genus Arthrobacter. In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E (eds) The Prokaryotes: volume 3: archaea. Bacteria: firmicutes, actinomycete. Springer, New York, NY, pp 945–960

Kim M, Oh HS, Park SC, Chun J (2014) Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 64:346–351. https://doi.org/10.1099/ijs.0.059774-0

Kuritani Y et al (2020) Conversion of levoglucosan into glucose by the coordination of four enzymes through oxidation, elimination, hydration, and reduction. Sci Rep 10:20066. https://doi.org/10.1038/s41598-020-77133-8

Li J et al (2020) A sulfoglycolytic Entner-Doudoroff pathway in *Rhizobium leguminosarum* bv. *trifolii* SRDI565. Appl Environ Microbiol 86:e00750–e00720. https://doi.org/10.1128/aem.00750-20

Liu Y, Wei Y, Zhou Y, Ang EL, Zhao H, 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

Liu J et al (2021) Mechanistically diverse pathways for sulfoquinovose degradation in bacteria. ACS Catal 11:14740–14750. https://doi.org/10.1021/acscatal.1c04321

Nakamura K, Zhu S, Komatsu K, Hattori M, Iwashima M (2020) Deglycosylation of the isoflavone C-glucoside puerarin by a combination of two recombinant bacterial enzymes and 3-oxo-glucose. Appl Environ Microbiol. https://doi.org/10.1128/aem.00607-20

Ramaley R, Fujita Y, Freese E (1979) Purification and properties of *Bacillus subtilis* inositol dehydrogenase. J Biol Chem 254:7684–7690

Rein U, Gueta R, Denger K, Ruff J, Hollemeyer K, Cook AM (2005) Dissimilation of cysteate via 3-sulfolactate sulfo-lyase and a sulfate exporter in *Paracoccus pantotrophus* NKNCYSA. Microbiology 151:737–747. https://doi.org/10.1099/mic.0.27548-0

Rowland P, Basak AK, Gover S, Levy HR, Adams MJ (1994) The three-dimensional structure of glucose 6-phosphate dehydrogenase from *Leuconostoc mesenteroides* refined at 2.0 A resolution. Structure 2:1073–1087. https://doi.org/10.1016/S0969-2126(94)00110-3

Seemann T (2014) Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Sharma M et al (2020) Dynamic structural changes accompany the production of dihydroxypropanesulfonate by sulfolactaldehyde reductase. ACS Catal 10:2826–2836. https://doi.org/10.1021/acscatal.9b04427

Sharma M et al (2021) Molecular basis of sulfosugar selectivity in sulfoglycolysis. ACS Cent Sci 7:476–487. https://doi.org/10.1021/acscentsci.0c01285

Sharma M, Lingford JP, Petricevic M, Snow AJP, Zhang Y, Jarva M, Mui JW-Y, Scott NE, Saunders EC, Mao R, Epa R, da Silva BM, Pires DEV, Ascher DB, McConville MJ, Davies GJ, Williams SJ, Goddard-Borger ED (2022) Oxidative desulfurization pathway for complete catabolism of sulfoquinovose by bacteria. Proc Natl Acad Sci USA 119:e2116022119. https://doi.org/10.1073/pnas.2116022119

Shimada T, Yamamoto K, Nakano M, Watanabe H, Schleheck D, Ishihama A (2019) Regulatory role of CsqR (YihW) in transcription of the genes for catabolism of the anionic sugar sulfoquinovose (SQ) in *Escherichia coli* K-12. Microbiology 165:78–89. https://doi.org/10.1099/mic.0.000740

Shlykov MA, Zheng WH, Chen JS, Saier MH Jr (2012) Bioinformatic characterization of the 4-toluene sulfonate uptake permease (TSUP) family of transmembrane proteins. Biochim Biophys Acta 1818:703–717. https://doi.org/10.1016/j.bbamem.2011.12.005

Snow AJD, Burchill L, Sharma M, Davies GJ, Williams SJ (2021) Sulfoglycolysis: catabolic pathways for metabolism of sulfoquinovose. Chem Soc Rev 50:13628–13645. https://doi.org/10.1039/D1CS00846C

Speciale G, Jin Y, Davies GJ, Williams SJ, Goddard-Borger ED (2016) YihQ is a sulfoquinovosidase that cleaves sulfoquinovosyl diacylglyceride sulfolipids. Nat Chem Biol 12:215–217. https://doi.org/10.1038/nchembio.2023

Sugiura M, Nakahara M, Yamada C, Arakawa T, Kitaoka M, Fushinobu S (2018) Identification, functional characterization, and crystal structure determination of bacterial levoglucosan dehydrogenase. J Biol Chem 293:17375–17386. https://doi.org/10.1074/jbc.RA118.004963

Taberman H, Parkkinen T, Rouvinen J (2016) Structural and functional features of the NAD(P) dependent Gfo/Idh/MocA protein family oxidoreductases. Protein Sci 25:778–786. https://doi.org/10.1002/pro.2877

Weinitschke S, Denger K, Cook AM, Smits THM (2007) The DUF81 protein TauE in *Cupriavidus necator* H16, a sulfite exporter in the metabolism of C2 sulfonates. Microbiology 153:3055–3060. https://doi.org/10.1099/mic.0.2007/009845-0

Yoon SH, Ha SM, Lim J, Kwon S, Chun J (2017) A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie Van Leeuwenhoek 110:1281–1286. https://doi.org/10.1007/s10482-017-0844-4

Yoshida K, Yamaguchi M, Morinaga T, Ikeuchi M, Kinehara M, Ashida H (2006) Genetic modification of *Bacillus subtilis* for production of D-chiro-inositol, an investigational drug candidate for treatment of type 2 diabetes and polycystic ovary syndrome. Appl Environ Microbiol 72:1310–1315. https://doi.org/10.1128/aem.72.2.1310-1315.2006

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.