iScience, Volume *23*

Supplemental Information

Environmental and Intestinal Phylum *Firmicutes* Bacteria Metabolize the Plant Sugar Sulfoquinovose via a 6-Deoxy-6-sulfofructose Transaldolase Pathway

Benjamin Frommeyer, Alexander W. Fiedler, Sebastian R. Oehler, Buck T. Hanson, Alexander Loy, Paolo Franchini, Dieter Spiteller, and David Schleheck

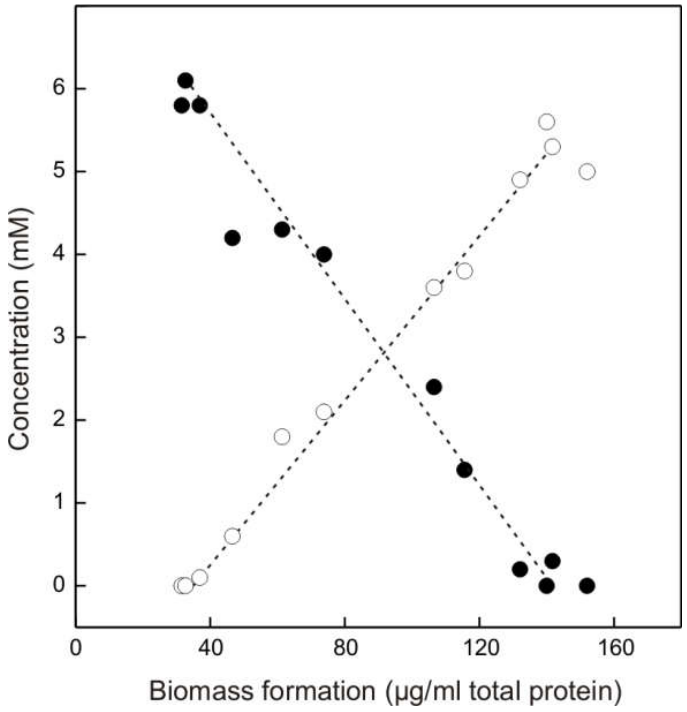

Figure S1. Linearized growth plot demonstrating complete substrate disappearance and stoichiometrical formation of 3-sulfolactate (SL) during growth of *B. aryabhattai* SOS1 with SQ, Related to Figure 1 and 6A. Concentrations of SQ (solid circles) and SL (open circles) were determined by HPLC at intervals during growth and the values plotted against biomass formation (total cellular protein). Data of one growth experiment (n=1) is shown.

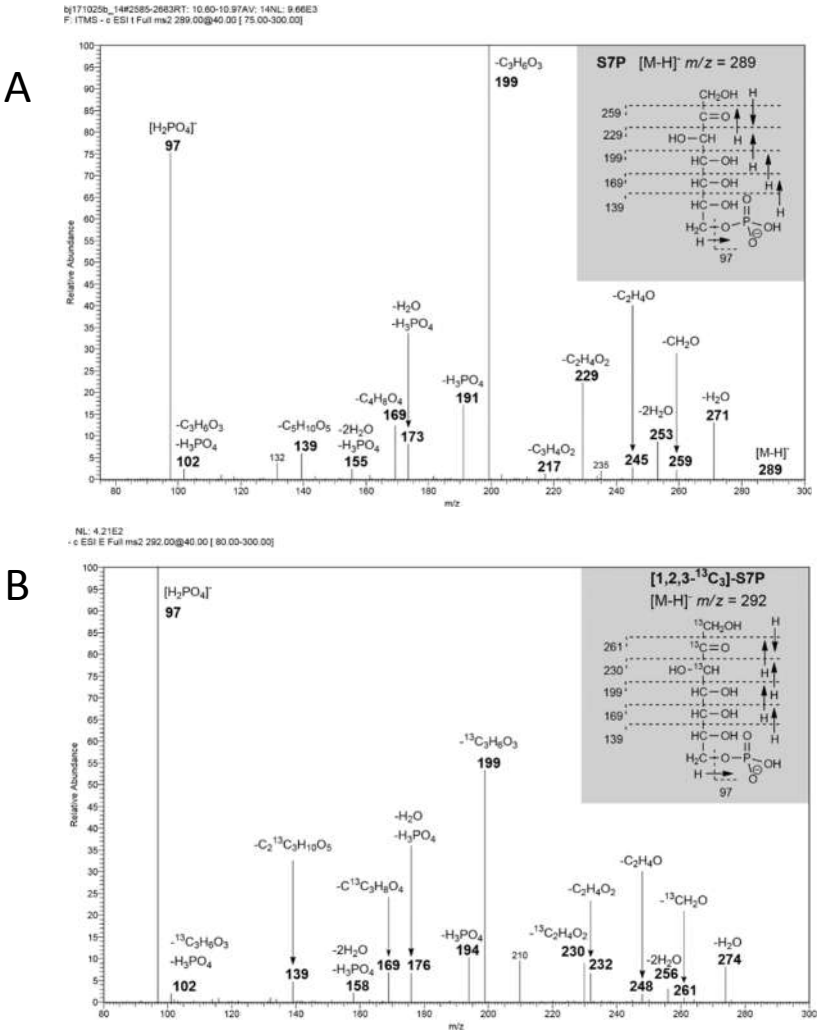

Figure S2. MS/MS fragmentation of unlabeled S7P analytical standard (A) and of [1,2,3-$^{13}C_3$]-S7P (B) as generated from $^{13}C_6$-SQ as substrate and unlabeled E4P as acceptor in cell extract of SQ-grown *B. aryabhattai* SOS1, Related to Figure 3. (**A**) Fragmentation pattern of the $[M-H]^-$ ions of unlabeled S7P standard. Fragmentation led to loss of one water (-18), formaldehyde (-30), two water (-36), ethanal (-44), $C_2H_4O_2$ (-60), $C_3H_4O_2$ (-72), dihydroxyacetone (-90), phosphoric acid (-98), phosphoric acid and water (-116), tetrose (-120), phosphoric acid and two water (-134), pentose (-150), and dihydroxyacetone and phosphate (-187) and the formation of phosphate (97). (**B**) Fragmentation pattern of the $[M-H]^-$ ions of [1,2,3-$^{13}C_3$]-S7P, as generated in enzyme reactions (see main text, **Figures 3, 4**), including corresponding mass shifts relative to that of the analytical standard (**A**) (271 → 274, 259 → 261, 253 → 256, 245 → 248, and 229 → 232/230), confirming a transfer of a $^{13}C_3$-glycerone moiety from $^{13}C_6$-SQ to unlabeled E4P concomitantly with formation of $^{13}C_3$-SLA (see **Figures 3** and **S6**).

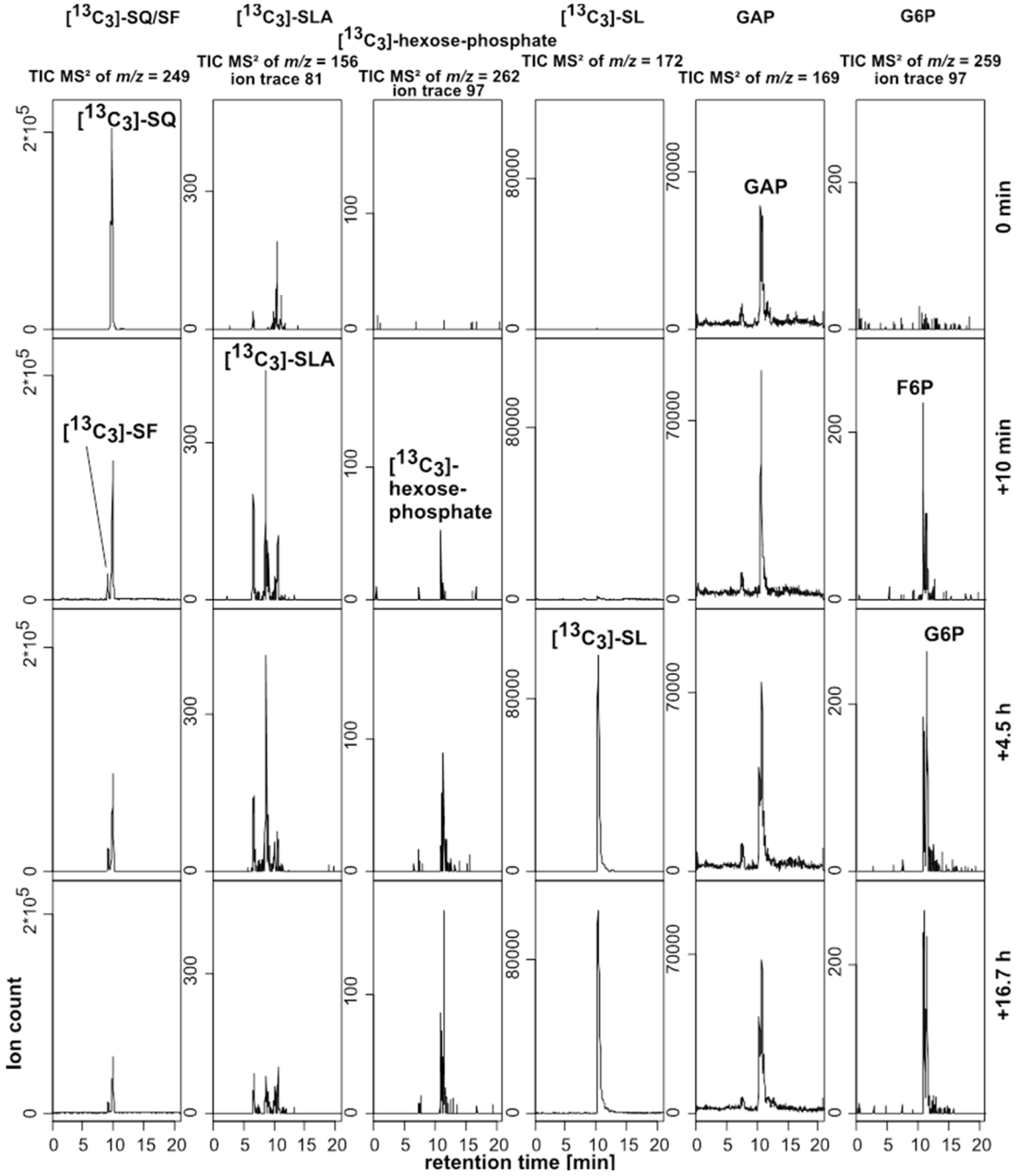

Figure S3. HPLC mass spectrometry confirming a transaldolase reaction in cell extracts of *B. aryabhattai* SOS1 using fully $^{13}C_6$-SQ as substrate and GAP as acceptor molecule, Related to Figure 3. The chromatograms illustrate conversion of $^{13}C_6$-SQ in the presence of GAP by cell extract of SQ-grown *Bacillus aryabhattai* SOS1, to $^{13}C_3$-SLA, which was further oxidized to $^{13}C_3$-SL, and to [$^{13}C_3$]-hexose phosphate. The reaction contained 50 mM $(NH_4)_2CO_3$ (pH 9.0), 1 mM DTT, 1 mM $MnCl_2$, 0.5 mM $MgCl_2$, 2 mM $^{13}C_6$-SQ, 6 mM NAD^+ and 6 mM GAP. Samples were taken and analyzed directly before addition of extract and after 10 min, 4.7

h and 21.4 h. For the organosulfonates, the total-ion chromatograms (TICs) recorded in the negative ion mode from the MS/MS fragmentation of the $[M-H]^-$ ions are shown. For SLA, the MS/MS-ion trace of the fragment from loss of the sulfonate group (*m/z* = 81) is shown. The $[^{13}C_3]$-hexose-phosphates were detected *via* fragmentation of $[M-H]^-$ ion (*m/z* = 262) and recording the phosphate-specific fragment *m/z* = 97. Note that F6P and G6P eluted in two peaks with very similar same retention time (**Figure S12**). Triose phosphates were detected by the $[M-H]^-$ ion trace (*m/z* = 169). Formation of unlabeled F6P and G6P can be explained by enzymes active in the cell extract for gluconeogenesis, e.g. through conversion of the added GAP.

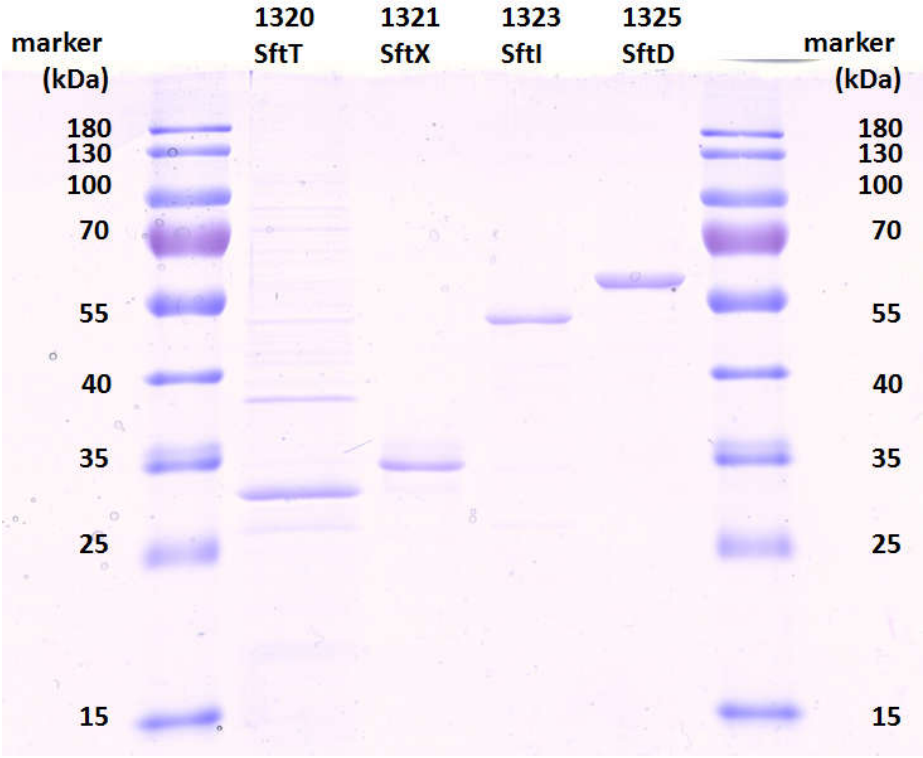

Figure S4. SDS-PAGE of purified recombinant proteins, Related to Figure 4. Protein standards, lane 1 and 6; *B. aryabhattai* SftT (25.3 + 4.1 [His-tag] kDa), lane 2; SftX (25,6 + 4.1 kDa), lane 3; SftI (49.7 + 4.1 kDa), lane 4; and SftD (52.1 + 4.1 kDa), lane 5. Note that we expressed and purified also domain-of-unknown-function (DUF4867) protein SftX, but could not attribute any activity to this protein, e.g., when tested individually or combination with SftITD.

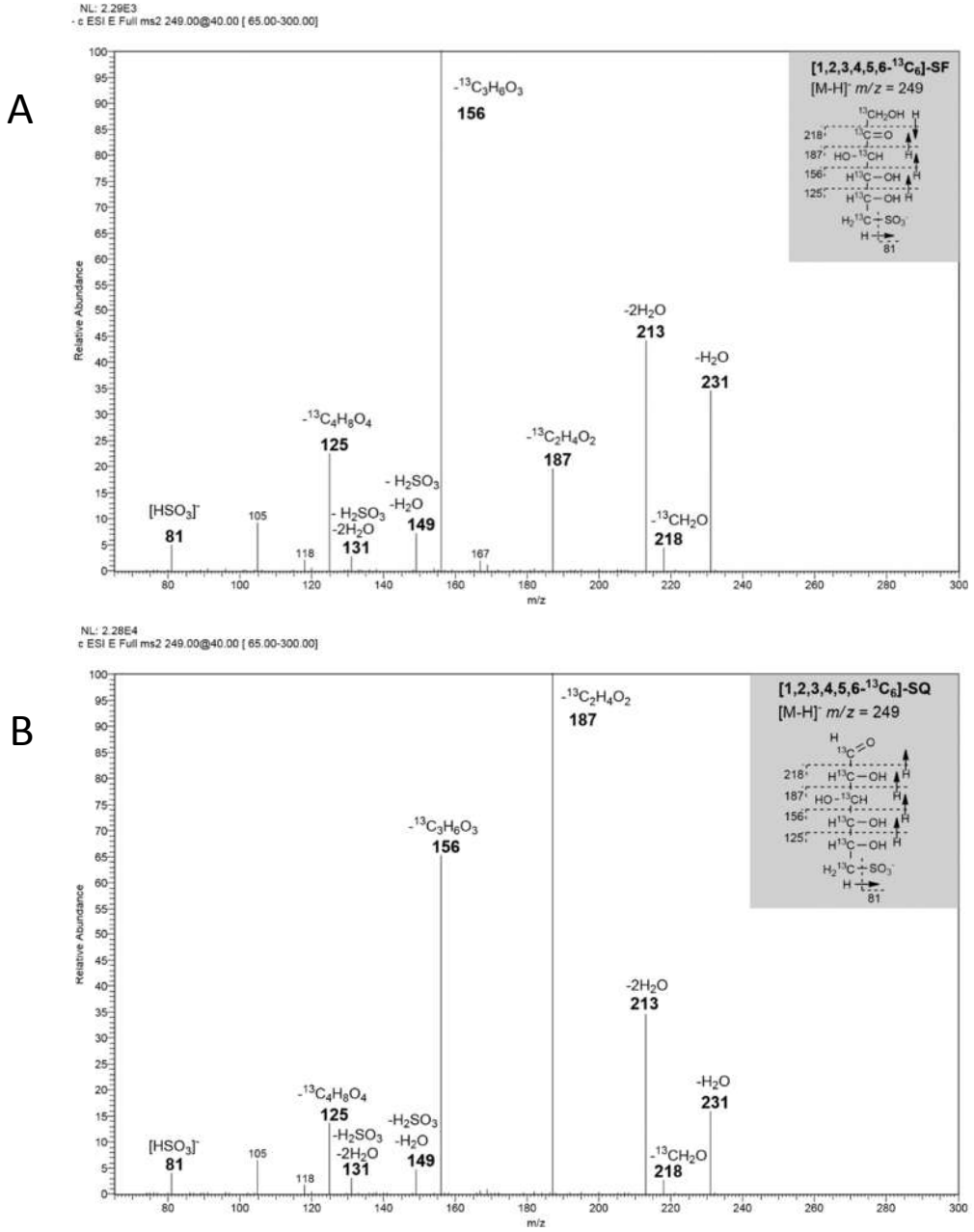

Figure S5. MS/MS fragmentation of $^{13}C_6$-SF (A) and $^{13}C_6$-SQ (B), Related to Figures 3 and 4. (A) Fragmentation of the $[M-H]^-$ ions of the $^{13}C_6$-SF (**Figures 3, 4**) led to a pattern as observed for unlabeled SF (**Denger et al., 2014**) including corresponding mass shifts (225 → 231, 213 → 218, 207 → 213, 183 → 187, 153 → 156, 143 → 149, 125 → 131, and 123 →125). (B) Fragmentation of the $[M-H]^-$ ions of $^{13}C_6$-SQ led to a pattern as for unlabeled SQ (**Denger et al., 2014**) including corresponding mass shifts (225 → 231, 213 → 218, 207 → 213, 183 → 187, 153 → 156, 143 → 149, 125 → 131, and 123 →125).

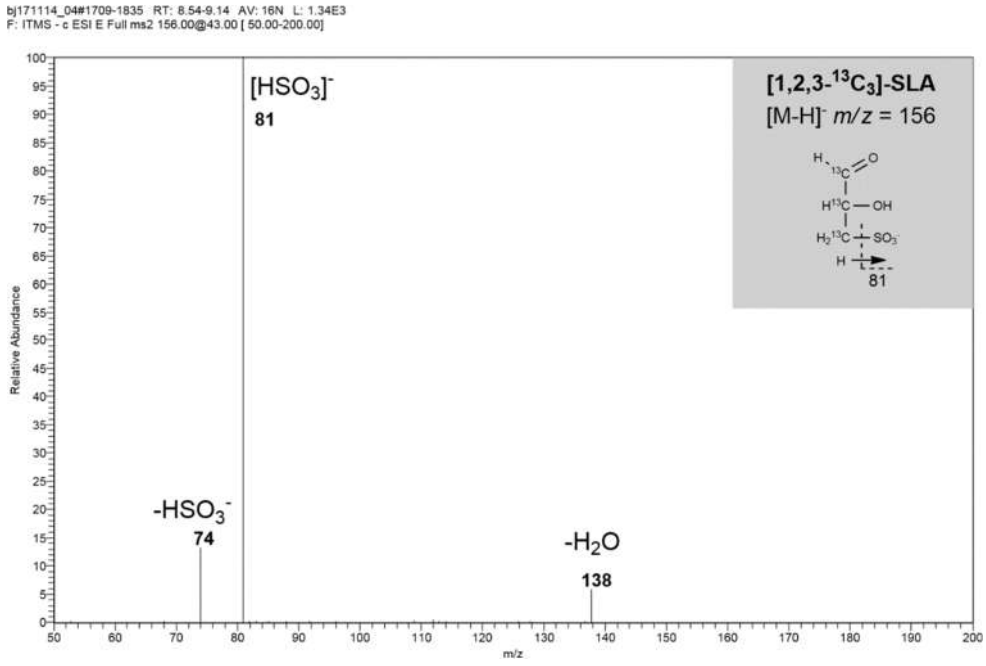

Figure S6. MS/MS fragmentation of $^{13}C_3$-SLA, Related to Figures 3 and 4. Fragmentation of the [M-H]$^-$ ions of the [1,2,3-$^{13}C_3$]-SLA (**Figures 3, 4**) led to a pattern as published previously (**Felux et al., 2015**) but with corresponding mass shifts (135 → 138, 71 → 74).

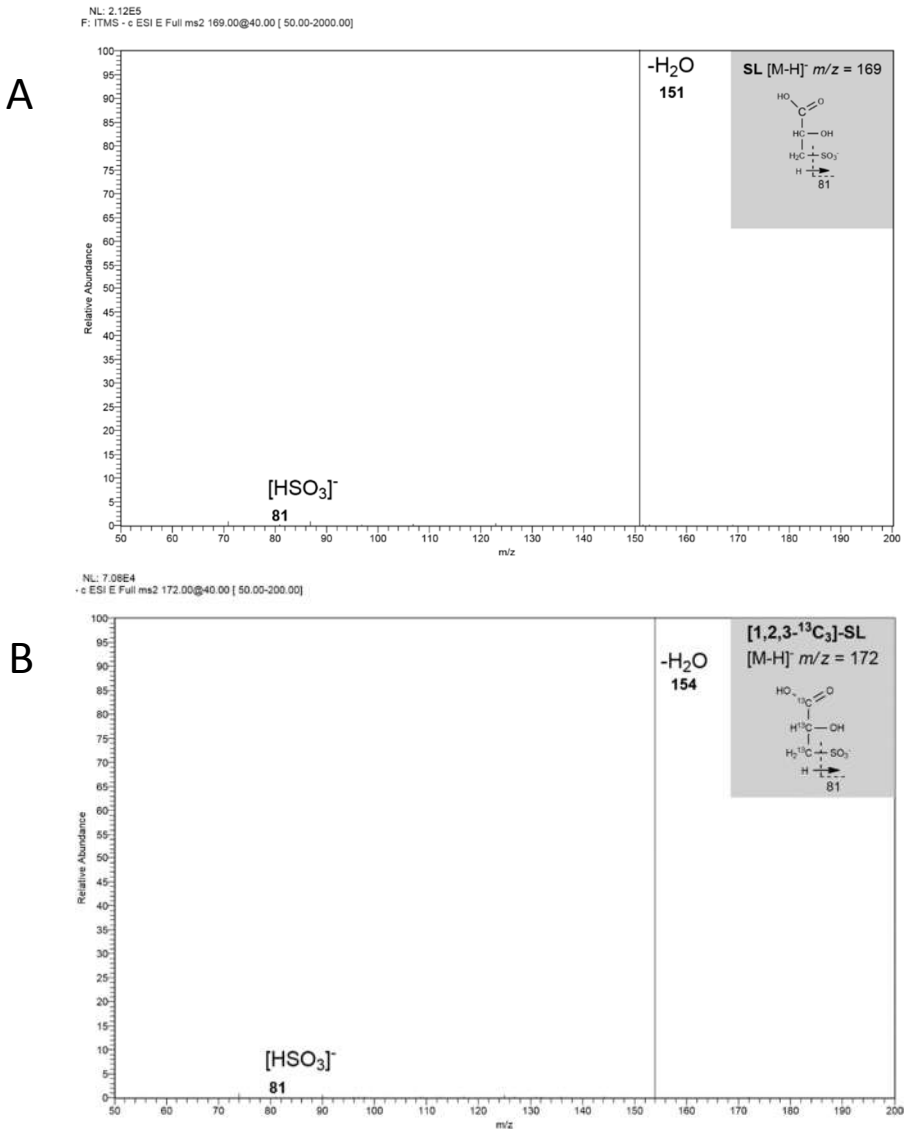

Figure S7. MS/MS fragmentation of $^{13}C_3$-SL, Related to Figures 3 and 4. (**A**) Fragmentation of the [M-H]$^-$ ions of authentic SL standard. Ion intensities were different than previously published (**Felux et al., 2015**) because of a stronger collision-induced dissociation energy was used. (**B**) Fragmentation of the [M-H]$^-$ ions of the [1,2,3-$^{13}C_3$]-SL (**Figures 3, 4**) including corresponding mass shift (151 → 154) of the major fragment.

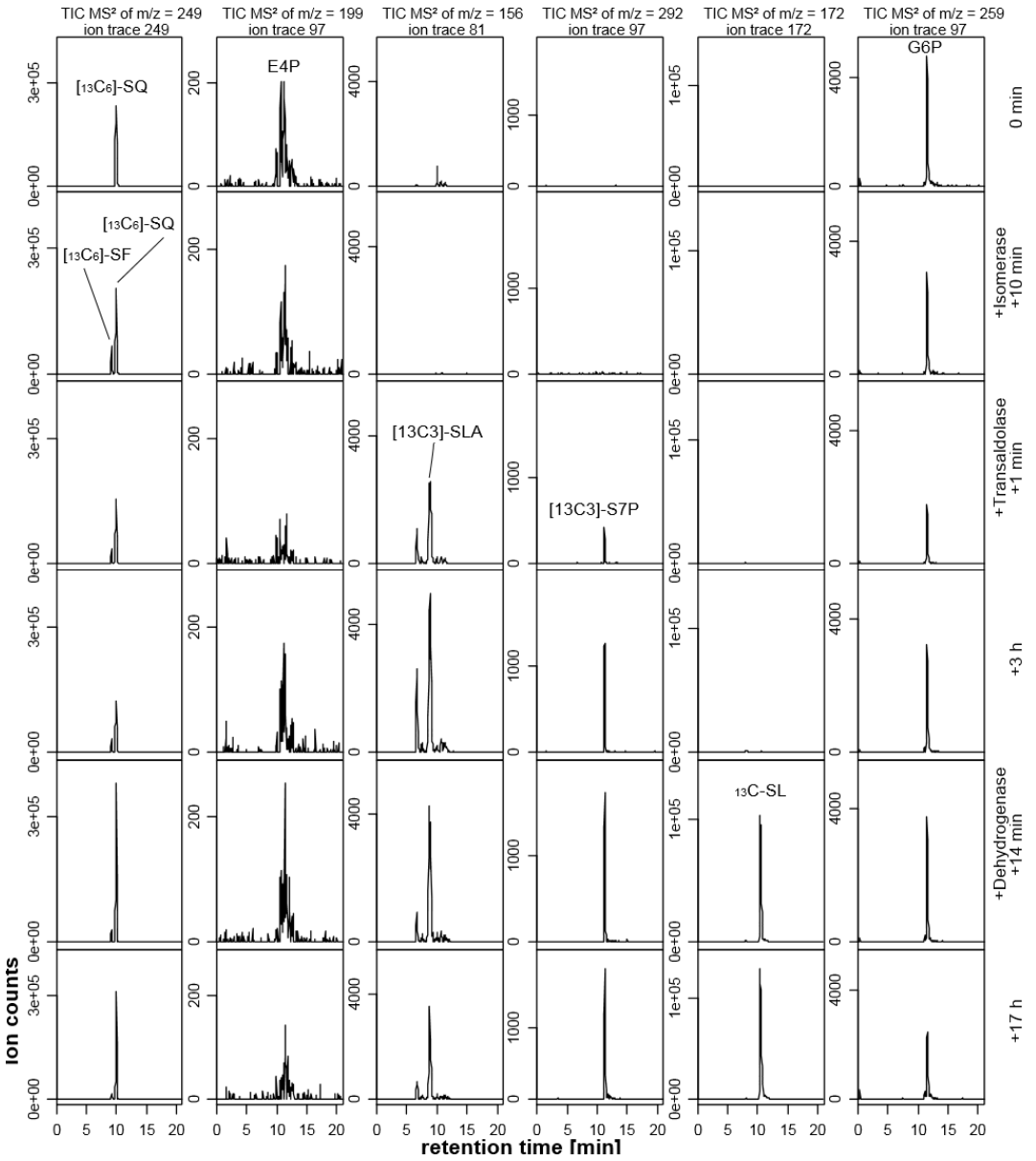

Figure S8. *In-vitro* reconstitution of the SFT pathway by recombinant enzymes using E4P as acceptor instead of GAP, Related to Figure 4. Cleavage of $^{13}C_6$-SQ *via* $^{13}C_6$-SF to [1,2,3-$^{13}C_3$]-S7P and $^{13}C_3$-SLA in the presence of unlabeled E4P, and oxidation of the $^{13}C_3$-SLA to $^{13}C_3$-SL in the presence of NAD^+. The reaction mixture initially contained 2 mM $^{13}C_6$-SQ and 6 mM E4P (0 min); note that purchased E4P contained impurities of G6P. SQ isomerase SftI (200 µg/ml) was added and the reaction sampled after 10 min. Then, SF transaldolase SftT (100 µg/ml) was added and the reaction sampled after 3 h. Finally, SLA-dehydrogenase SftD was added (200 µg/ml) and NAD^+ (6 mM) and the reaction sampled after 14 min and 17 h. Note that even after 17 h reaction time, substrate conversion was incomplete because of the promiscuous activity of SftD oxidizing E4P in the presence of NAD^+.

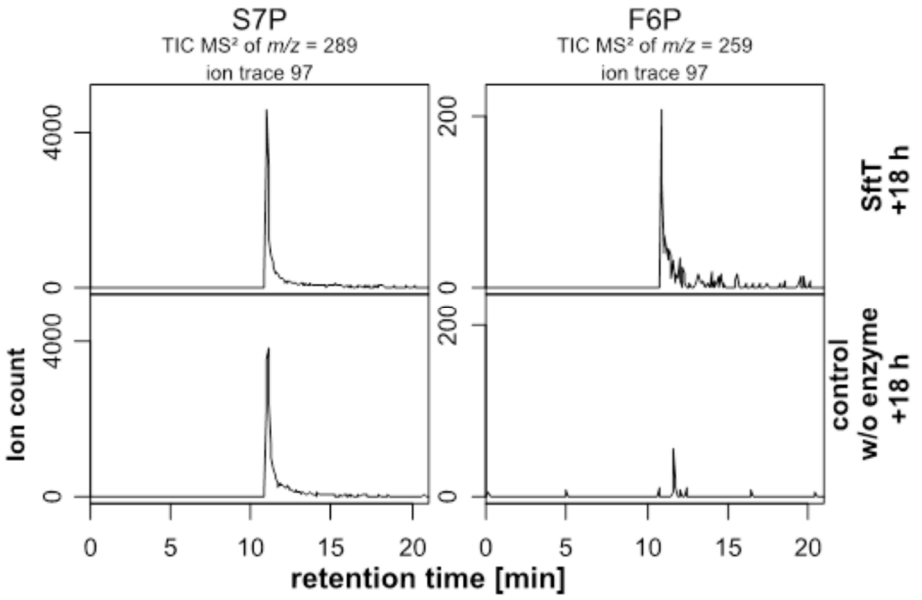

Figure S9. Promiscuous activity of transaldolase SftT as S7P transaldolase in the presence of GAP as acceptor, Related to Figure 4. HPLC MS/MS-ion traces demonstrating conversion of S7P (left panel) to F6P (right panel) in the presence of GAP and transaldolase SftT (100 µg/ml), in comparison to a control reaction containing GAP but not SftT. The reaction mixture initially contained 1 mM S7P, 1 mM GAP, 50 mM $(NH_4)_2CO_3$ (pH 9.0), 1 mM DTT, 1 mM $MnCl_2$ and 1 mM $MgCl_2$. Note that the purchased GAP contained impurities of F6P (as visible for the control reaction).

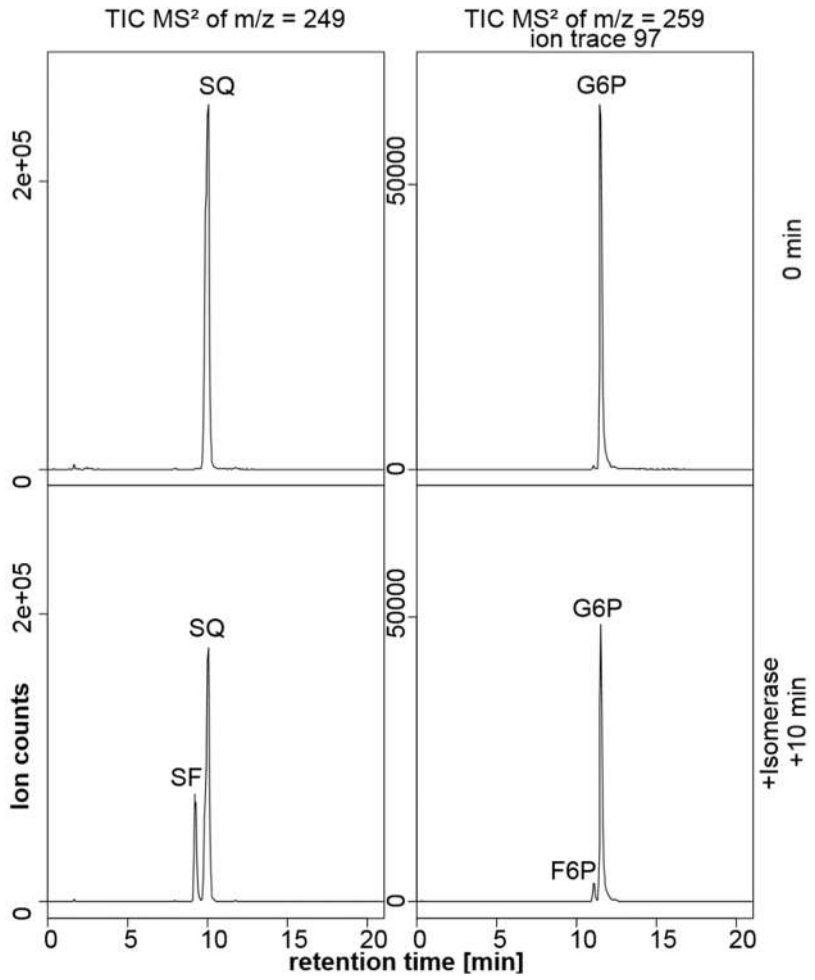

Figure S10. Promiscuous activity of isomerase SftI as G6P isomerase, Related to Figure 4. HPLC-MS chromatograms demonstrating conversion of SQ to SF (left panel) and of G6P to F6P (right panel) by isomerase SftI. The reaction mixture initially contained 2 mM SQ or 6 mM G6P, and 50 mM $(NH_4)_2CO_3$ (pH 9.0), 1 mM DTT, 1 mM $MnCl_2$ and 1 mM $MgCl_2$. Samples were taken before and 10 min after addition of enzyme (100 μg/ml).

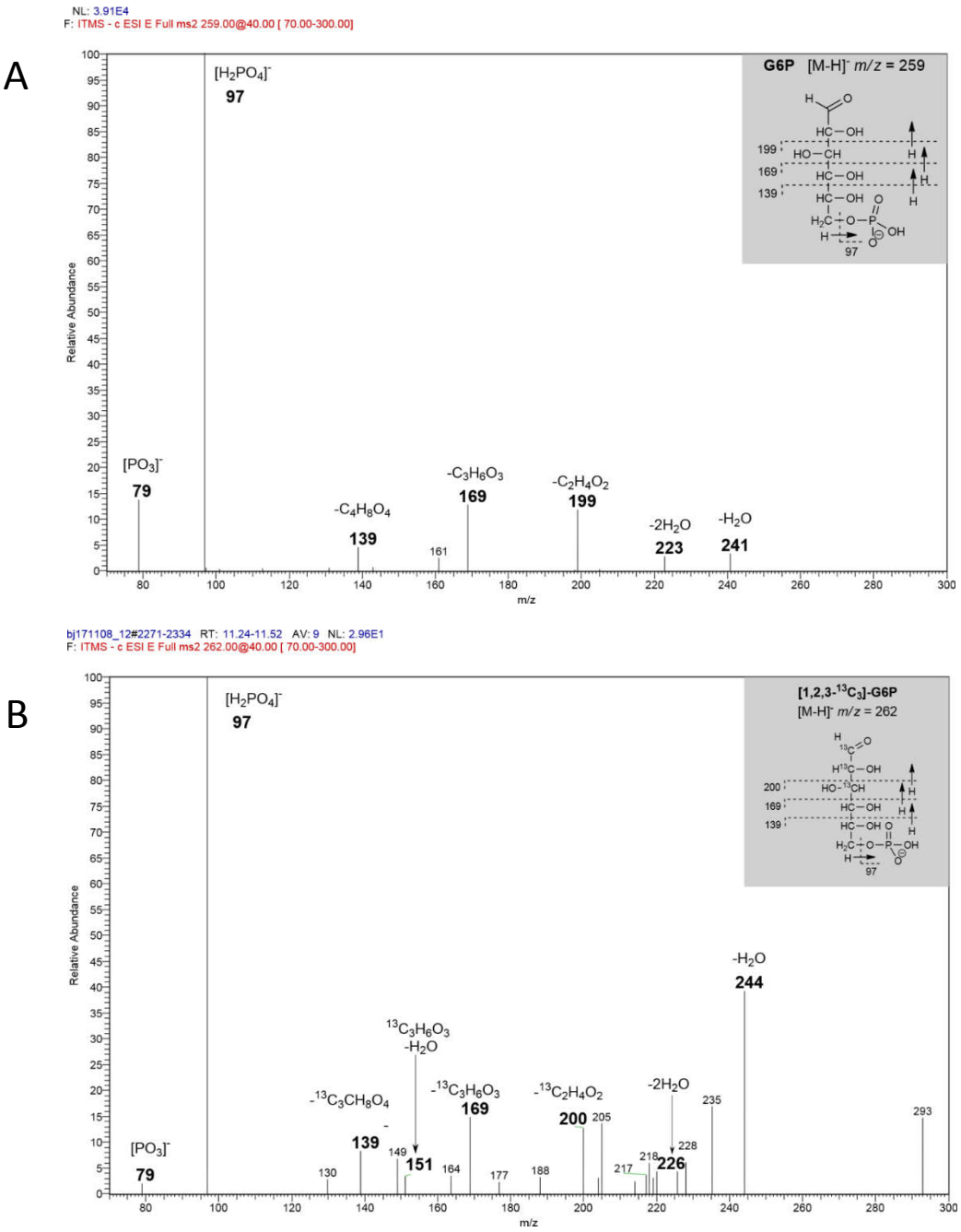

Figure S11. MS/MS fragmentation of G6P (A) and [1,2,3-$^{13}C_3$]-G6P (B), Related to Figure 4. (**A**) Fragmentation of the $[M-H]^-$ ions of G6P led to a loss of water (-18), two water (-36), $C_2H_4O_2$ (-60), triose (-90), tetrose (-120), and the formation of phosphate characteristic peaks 97 and 79 (**de Souza et al., 2009**). (**B**) Fragmentation pattern of the $[M-H]^-$ ions of [1,2,3-$^{13}C_3$]-G6P inclusive corresponding mass shifts (241 → 244, 223 → 226, and 199 → 200). Note that the fragmentation patterns for F6P and G6P are very similar, but that F6P eluted earlier than G6P in HPLC (**Figure S10**).

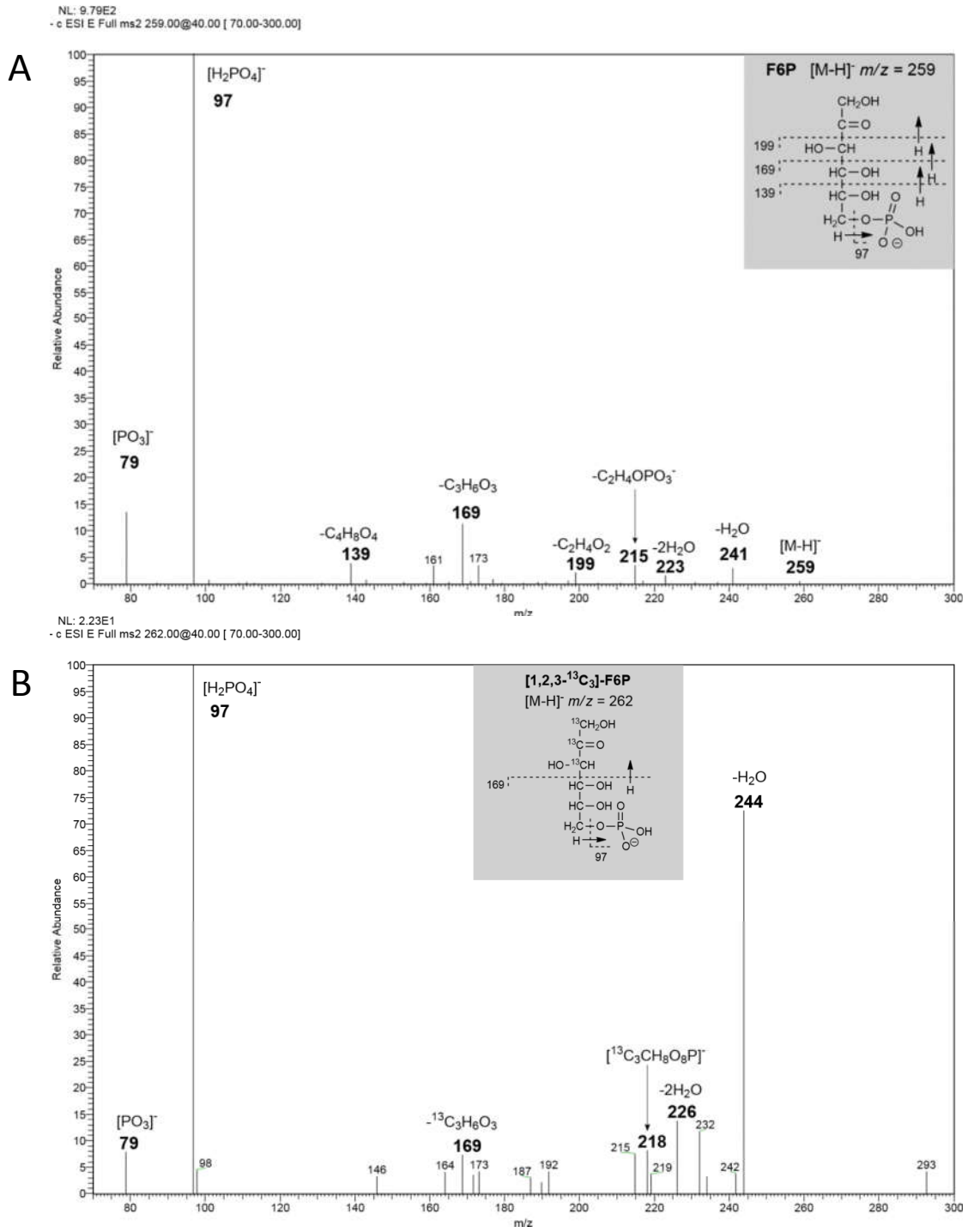

Figure S12. MS/MS fragmentation of F6P (A) and [1,2,3-$^{13}C_3$]-F6P (B), Related to Figure 4. (**A**) Fragmentation of the $[M-H]^-$ ions of F6P led to a loss of water (-18), two water (-36), $C_2H_4O_2$ (-60), triose (-90), tetrose (-120), and the formation of phosphate characteristic peaks 97 and 79 (**de Souza et al., 2009**). (**B**) Fragmentation of the $[M-H]^-$ ions of [1,2,3-$^{13}C_3$]-F6P inclusive corresponding mass shifts (241 → 244; 223 → 226 and 199 → 200). Note that the fragmentation for F6P and G6P is very similar, but that F6P eluted earlier than G6P in HPLC (**Figure S10**).

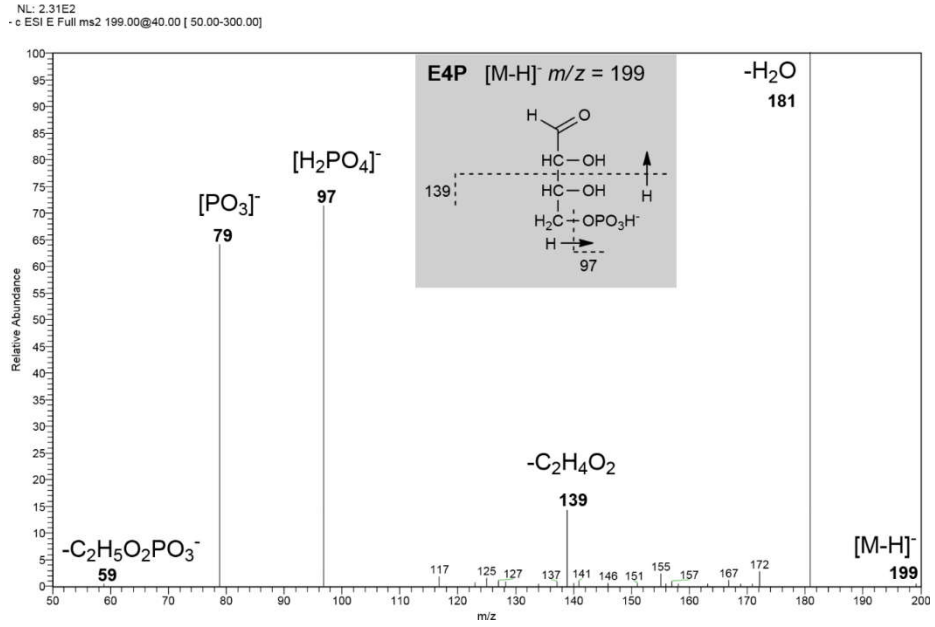

Figure S13. MS/MS fragmentation of E4P parental ion [M-H]$^-$, Related to Figure 3. Fragmentation led to a loss of water (-18), $C_2H_4O_2$ (-30), and the formation of characteristic phosphate peaks 97 and 79 (**de Souza et al., 2009**).

Table S1. Related to Figures 2 and 4. Primer sequences (5'-3') for directional cloning of *Bacillus aryabhattai* SOS1 pathway genes, plasmid-insert sequencing and 16S-rRNA gene sequencing.

1320_for3*	CACCATGAAGTATTTTTTAGATAGTGCCATTTTAGAG
1320_rev2	AAGCCCCTCAAAGAAGATAATAAGATTC
1321_for3*	CACCATGGGCAGCTTTCAATACATGAAAGACTT
1321_rev1	AGTAGAACTACGCCGTTAAGAGCAACTT
1322_for1*	CACCATGTCAGGTACCTTAAACGTAACTAAAGGG
1322_rev2	ATACTAATTTTCCGCTTGACTACACTTCTTAT
1323_for2*	CACCATGCAAAATACGACAGTTTTATATGTGC
1323_rev1	TTGTATCTTTCATGTTCTCAAACTCCTT
1325_for4*	CACCATGACGAGTTTAACTCAAGTCAAACAATATG
1325_rev1	AAAAATCCCGAAAATAGGAAGAAGGA
T7 forward	TAATACGACTCACTATAGGG
T7 reverse	TAGTTATTGCTCAGCGGTGG
8F	AGAGTTTGATYMTGGCTC
1492R	GGYTACCTTGTTACGACTT

*Adaptor sequence for directional cloning underlined

TRANSPARENT METHODS

Enzyme nomenclature. We suggest that *B. aryabhattai* SQ isomerase SftI would belong to NC-IUBMB (Nomenclature Commission of the International Union of Biochemistry and Molecular Biology) subgroup EC 5.3.1., with the name sulfoquinovose isomerase (systematic name 6-deoxy-6-sulfoglucose aldose-ketose-isomerase); it showed activity also as D-glucose 6-phosphate isomerase (EC 5.3.1; D-glucose 6-phosphate aldose-ketose-isomerase). SF transaldolase SftT would belong to EC 2.2.1.2 with the name sulfofructose transaldolase (systematic name 6-deoxy-6-sulfofructose:D-glyceraldehyde-3-phosphate glycerone-transferase); it showed activity also as sedoheptulose-7-phosphate transaldolase (EC 2.2.1.2; sedoheptulose-7-phosphate:D-glyceraldehyde-3-phosphate glyceronetransferase). SLA dehydrogenase SftD belongs to the EC 1.2.1 with the name sulfolactaldehyde dehydrogenase (systematic name 3-sulfolactaldehyde:NAD^+ oxidoreductase); it showed activity also as erythrose-4-phosphate and glyceraldehyde-3-phosphate:NAD^+ oxidoreductase.

Chemicals. NAD^+, fructose-6-phosphate disodium salt, D/L-gyceraldehyde-3-phosphate (45-55 mg/ml in H_2O solution), glucose-6-phosphate sodium salt and chloramphenicol were supplied by Sigma (now Merck KGaA, Darmstadt Germany), and D-erythrose-4-phosphate sodium salt by Sigma, Carbosynth (Compton, UK) and Santa Cruz Biotechnology (Dallas, Texas). D-sedoheptulose-7-phosphate barium salt was from Carbosynth. SQ, $^{13}C_6$-SQ and SL were synthesized by MCAT GmbH (Donaueschingen, Germany). DHPS was synthesized and validated by NMR as reported previously (**Mayer et al., 2010; Denger et al., 2012**). 1,4-Dithiothreitol was supplied by Carl Roth (Karlsruhe, Germany) (p.a. grade) and VWR International GmbH (Electran molecular biology grade). Adenosin-5'-triphosphate disodium salt was from Serva (Heidelberg, Germany), manganese(II) chloride tetrahydrate (p. a. grade) from Riedel-de Haën (now Honeywell, Seelze, Germany). Ampicillin was purchased from Carl

Roth (Karlsruhe, Germany), isopropyl-β-D-thiogalactopyranosid (IPTG) from carbolution chemicals GmbH (St. Ingbert, Germany), and Imidazol from Merck (Kenilworth, New Jersey, US). Gases for anaerobic cultivation were purchased from Messer-Griesheim (Darmstadt, Germany) and Sauerstoffwerke Friedrichshafen (Friedrichshafen, Germany).

Enrichment, isolation and identification of *Bacillus aryabhattai* SOS1. *B. aryabhattai* SOS1 (DSM 104036) was isolated from an enrichment cultures using a maple leaf as inoculum (collected in the forest of the campus of University of Konstanz), and by using the following purification strategy. Sterile phosphate-buffered (pH 7.2) mineral salts medium (**Thurnheer et al., 1986**) (5 ml) with 6 mM SQ as sole carbon and energy source in 30 ml glass tubes (Corning, USA) was used to wash microorganisms off the leaf by vortexing. The leaf was removed and the enrichment culture incubation at 30°C shaking at 165 rpm. For enrichment and isolation of the other SQ-degrading strains, as mentioned in the results section, the inocula were prepared as follows. Samples of different soils (forest or agricultural soils, collected at around University of Konstanz; approx. 1 g) were suspended in sterile mineral-salts medium (10 ml) and sonified (30 s) to detach microorganisms; supernatant of these soil suspensions were used as inoculum (100 µl). Heat-treated (pasteurized) soil inocula were prepared when the soil samples were dried for three days at 40°C. Appr. 50 mg were added to SQ minerals-salts medium, and these inoculated cultures were heated at 65°C or 80°C for 1 h prior to incubation at 30°C, 165 rpm shaking. A water sample from a pond (collected at around University of Konstanz) was used as inoculum (100 µl) when added directly to the culture tube without heat treatment. When growth was visible as turbidity and by presence of bacteria (microscopy), and after SQ disappearance was confirmed by HPLC analysis, 5 - 15 µl of the outgrown cultures were transferred into fresh, sterile 3 ml SQ mineral-salts medium in 30-ml glass tubes. Sub-cultivations were pursued until to the 5th transfer, after which samples of culture fluid were streaked on LB5 plates (10 g/l tryptone, 5 g/l yeast extract, 5 g/l NaCl, 15 g/l agar); the plates were incubated at 30°C in the dark. From these plates, each colony type was picked individually and transferred back into

liquid cultures with SQ-salts medium. This procedure was repeated until a homogeneous colony picture on LB plates was obtained. The isolates were identified by sequencing of the 16S rRNA gene (primers 8F and 1492R; **Table S1**) through colony PCR. Sanger sequencing was performed on purified amplicons (Eurofins, Germany) and characterized against the NCBI and RDP databases.

***B. aryabhattai* SOS1 cultivation and preparation of cell extracts.** Strain SOS1 was grown at 30°C, 150-180 rpm shaking, in the range from 5 ml culture medium in 30-ml glass tubes up to 2 liters in Erlenmeyer flasks, using the mineral salts medium described above with either SQ (6 or 12 mM) or glucose (3 or 6 mM) as substrate; concentrations were adjusted to compensate for half the molar growth yield with SQ *vs.* glucose. Cultures were inoculated with a colony from LB5-agar plates (5-ml scale) or with 0.5-5% (v/v) of outgrown pre-culture. Growth was monitored as optical density at 580 nm (OD580) either directly in the glass tubes or through side-arm flasks in a tube photometers (model M 107; Campspec), or in samples of 1 ml culture fluid in cuvettes (Novaspec Plus; Amersham Biosciences). For determination of substrate disappearance and product formation by HPLC (see below), samples from the supernatant obtained after centrifugation (10 minutes, 16,100 x g, 20°C) of 1 ml of culture fluid were taken; the cell pellet was retained for total protein determination by Lowry assay (**Kennedy and Fewson 1968**) against bovine serum albumin (BSA) as a standard. For proteomic analysis and preparation of cell extracts, cells were grown in larger scales in Erlenmeyer flasks and harvested by centrifugation (15 min, 15,000 x g, 10°C). The cell pellets were stored frozen (-20°C).

For two-dimensional (2D) protein gel electrophoresis and total proteomics, cell pellets were re-suspended in Tris-HCl buffer (pH 7.5) containing 25 µg/ml DNAse and 2 mM $MgCl_2$. The cells were disrupted by five passages through a chilled French Pressure Cell (SLM Instruments, USA) at 140 MPa and unbroken cells and debris was removed by centrifugation (10 min, 15,000 x g, 4°C) to obtain cell extracts. Samples of cell extract were submitted to total proteomic

analysis (see below). For 2D-protein gel electrophoresis, membrane fragments were removed by ultracentrifugation (1 h, 70,000 x g, 4°C; Beckman Optima TL) to obtain soluble protein fraction, which was desalted using PD-10 columns (GE Healthcare) prior to 2D-protein gel electrophoresis (see below). Protein concentrations in the extracts were determined through Bradford assay (**Bradford 1976**) against BSA as standard.

For enzyme assays, the cell pellets were resuspended and washed twice with $(NH_4)_2CO_3$ buffer (50 mM, pH 9.0) by centrifugation (15 min, 5000 x g, 4°C); after the washing, 1 U/ml DNase was added. The lysis was performed by five passages through a chilled French pressure cell at 140 MPa. Cell debris was removed by centrifugation (15 min, 21.380 x g, 4°C) and small molecules were removed by gel filtration if appropriate (exclusion size 1-5 kDa; illustra NAP-10 gel filtration columns, GE Healthcare Life Sciences).

Draft-genome sequencing and proteomic analysis of *B. aryabhattai* SOS1. Genomic DNA of strain SOS1 was prepared by the CTAB-protocol (**William et al., 2012**) including RNase treatment. Whole genome shotgun sequencing of strain SOS1 was performed by GATC Biotech (Konstanz, Germany; now Eurofins) using an Illumina HiSeq 2500 platform and a 2x125bp paired-end library, which generated 11,92 million reads. The program Trimmomatic v0.33 (**Bolger et al., 2014**) was used to remove adapters and filter the reads by quality with default settings, as well as to discard sequences shorter than 50 nucleotides. The filtered reads were assembled *de novo* using the program SOAPdenovo v2.04 (**Luo et al., 2012**) with a k-mer size of 67 and setting the minimum contig length at 200 bp. The *de novo* assembly procedure resulted in 383 sequences (N50: 0.808 Mb) that were further scaffolded by the reference-guided algorithm implemented in Ragout 2.0b (**Kolmogorov et al., 2014**) using the genome of *Bacillus aryabhattai* T61 as reference. The final assembly of *B. aryabhattai* SOS1 included 359 sequences (N50: 4.638 Mb).

2D-protein gel electrophoresis (isoelectric focusing and SDS gel electrophoresis) of the soluble protein fraction of strain SOS1 was conducted as described previously (**Felux et al., 2015**). Total proteomics of cell extracts, and identification of proteins in spots excised from the SDS gels, was done as described previously (**Denger et al., 2014; Felux et al., 2015**, **Schmidt et al., 2013**) at the Proteomics Centre of the University of Konstanz (https://www.biologie.uni-konstanz.de/proteomics-centre/) with the exception that each sample was analyzed twice on a Orbitrap Fusion with EASY-nLC 1200 (Thermo Fisher Scientific), and tandem mass spectra were searched against the protein database using Mascot (Matrix Science) and Proteom Discoverer V1.3 (Thermo Fisher Scientific) with "Trypsin" enzyme cleavage, static cysteine alkylation by chloroacetamide, and variable methionine oxidation (**Peck et al., 2019**).

HPLC for determination of substrate and product concentrations in cultures. SQ, DHPS and SL in samples of culture supernatant (see above) were separated by hydrophilic interaction chromatography (HILIC) using an HPLC apparatus (Prominence LC-20A System; Shimadzu), and were detected by an Evaporative Light Scattering Detector (ELSD) (model ZAM 3000; Schambeck SFD GmbH, Germany). A SeQuant ZIC-HILIC HPLC Column (Merck, Germany) as described previously (**Denger et al., 2014**) was used for the separation. Samples of culture supernatants (0.6 ml) were mixed with acetonitrile (0.3 ml) in HPLC-vials. The injection volume was 10 µl. The column temperature was set at 30°C. The eluents were (A) 0.1 M ammonium acetate in water supplemented with 10% acetonitrile, and (B) 100% acetonitrile. The flow rate was 0.75 ml/min. The HPLC gradient was set from 90% B to 65% B in 25 min; hold at 65% B for 10 min; gradient to 90% B in 0.5 min; hold (reequilibration) at 90% B for 9.5 min. SQ, DHPS and SL were identified by their specific retention time against authentic standards (SQ, 25.9 - 26.4 min; DHPS, 17.6 - 17.7 min; SL, 27.2 - 27.6 min) and quantified by peak area integration against authentic standards (see above). SL was quantified also by ion chromatography with suppression under conditions described previously (**Styp von Rekowski et al., 2005**); SL eluted after 6.1 min.

HPLC-MS of substrate turnover in *B. aryabhattai* SOS1 cell extracts and in reactions with recombinant enzymes. HPLC-MS/MS measurements were conducted with a Finnigan Surveyor Autosampler Plus and MS Pump Plus coupled with a Finnigan LTQ (Thermo Fisher Scentific). The SeQuant ZIC-HILIC column (see above) was used for the separation fitted with an upstream prefilter (HICHROM HI-704 with 2 µm frit) and precolumn (ZIC HILIC Guard SeQuant, Merck). Solvent A was 0.1 M ammonium acetate in milliQ H_2O or double distilled H_2O with 10 % acetonitrile. Solvent B was acetonitrile with 0.1% glacial acetic acid. The flow rate was 0.3 ml/min. The HPLC gradient program started at 90% solvent B, decreasing to 10% B over 20 min, increasing back to 90% B within 1 min and equilibrating on 90% B for 5 min. The compounds were ionized by negative electrospray ionization (ESI). MS chromatograms and fragmentation patterns were analyzed with Xcalibur 2.0 (Thermo Fisher Scientific). Multiple plot figures were created by using the raw data exported from Qual browser 2.0 *via* .csv-files into Excel 2010 (Microsoft Corporation). These files were used by a custom written r-studio (Version 1.0.143; RStudio, Inc.) script to generate the figures.

The retention times of the substances separated by the HPLC gradient program and the ESI-MS/MS fragmentation patterns of the observed analytes were as follows: [1,2,3,4,5,6-$^{13}C_6$]-SQ retention time 9.8 min; [1,2,3,4,5,6-$^{13}C_6$]-SQ ESI-MS *m/z* (% base-peak): $[M-H]^-$ 249 (100); [1,2,3,4,5,6-$^{13}C_6$]-SQ ESI-MS/MS of $[M-H]^-$ 249: 231 (16), 218 (3), 213 (35), 187 (100), 156 (65), 149 (5), 131 (3), 125 (14), 81 (4). [1,2,3,4,5,6-$^{13}C_6$]-SF retention time 9.3 min; [1,2,3,4,5,6-$^{13}C_6$]-SF ESI-MS *m/z* (% base-peak) $[M-H]^-$ 249 (100); [1,2,3,4,5,6-$^{13}C_6$]-SF ESI-MS/MS of $[M-H]^-$ 249: 231 (33), 218 (4), 213 (43), 187 (20), 156 (100), 149 (7), 131 (3), 125 (23), 81 (5). [1,2,3-$^{13}C_3$]-SLA retention time 8.8 min. [1,2,3-$^{13}C_3$]-SLA ESI-MS *m/z* (% base-peak) $[M-H]^-$ 249 (100); [1,2,3-$^{13}C_3$]-SLA ESI-MS/MS of $[M-H]^-$ 156: 138 (6), 81 (100), 74 (13). SL retention time 10.5; SL ESI-MS *m/z* (% base peak) $[M-H]^-$ 169 (100); SL ESI-MS/MS of $[M-H]^-$ 169: 151 (100), 81 (<1). [1,2,3-$^{13}C_3$]-SL retention time 10.5; [1,2,3-$^{13}C_3$]-SL ESI-MS *m/z* (% base peak) $[M-H]^-$ 169 (100); [1,2,3-$^{13}C_3$]-SL ESI-MS/MS of $[M-H]^-$ 172: 154 (100),

81 (>1). GAP retention time 12.5; GAP ESI-MS *m/z* (% base peak) $[M-H]^-$ 169 (100); GAP ESI-MS/MS of $[M-H]^-$ 169: 151 (21), 141 (7), 97 (100), 79 (32). E4P retention time 10.6; E4P ESI-MS *m/z* (% base peak) $[M-H]^-$ 199 (100); E4P ESI-MS/MS of $[M-H]^-$ 199: 199 (<1), 181 (100), 139 (14), 97 (71), 79 (64). S7P retention time 11.2; S7P ESI-MS *m/z* (% base peak) $[M-H]^-$ 289 (100); S7P ESI-MS/MS $[M-H]^-$ 289: 289 (<1), 271 (13), 259 (2), 253 (8), 245 (2), 229 (22), 217 (1), 199 (100), 191 (17), 173 (8), 169 (13), 155 (2), 139 (6), 102 (3), 97 (75). [1,2,3-$^{13}C_3$]-S7P retention time 11.2; [1,2,3-$^{13}C_3$]-S7P ESI-MS *m/z* (% base peak) $[M-H]^-$ 292 (100); [1,2,3-$^{13}C_3$]-S7P ESI-MS/MS $[M-H]^-$ 292: 274 (8), 261 (1), 256 (3), 248 (2), 232 (7), 230 (9), 199 (53), 194 (10), 176 (7), 169 (7), 158 (2), 139 (5), 102 (1), 97 (100). G6P retention time 11.5; G6P ESI-MS *m/z* (% base peak) $[M-H]^-$ 259 (100); G6P ESI-MS/MS $[M-H]^-$ 259: 241 (3), 223 (3), 199 (12), 169 (13), 139 (5), 97 (100), 79 (14). [1,2,3-$^{13}C_3$]-G6P retention time 11.5; [1,2,3-$^{13}C_3$]-G6P ESI-MS *m/z* (% base peak) $[M-H]^-$ 262 (100); [1,2,3-$^{13}C_3$]-G6P ESI-MS/MS $[M-H]^-$ 262: 244 (2), 226 (1), 200 (4), 169 (12), 139 (4), 97 (100), 79 (13). F6P retention time 11.0; F6P ESI-MS *m/z* (% base peak) $[M-H]^-$ 259 (100); F6P ESI-MS/MS $[M-H]^-$ 259: 241 (3), 223 (1), 199 (2), 169 (11), 139 (4), 97 (100), 79 (14). [1,2,3-$^{13}C_3$]-F6P retention time 11.0; [1,2,3-$^{13}C_3$]-F6P ESI-MS *m/z* (% base peak) $[M-H]^-$ 262 (100); [1,2,3-$^{13}C_3$]-F6P ESI-MS/MS $[M-H]^-$ 262: 244 (8), 226 (4), 200 (4), 169 (11), 139 (3), 97 (100), 79 (8).

Enzyme assays in cell extracts. For screening of formation of sulfo-EMP or sulfo-ED pathway intermediates in cell extracts of strain SOS1, Tris-HCl buffer (50 mM, pH 7.8) or potassium phosphate buffer (50 mM, pH 6.5) containing 200 µg/ml soluble protein fraction, 2 mM SQ, 4 mM NAD^+, 4 mM $NADP^+$, 8 mM ATP and 0.5 mM Mg^{2+} was incubated at room temperature. At intervals, samples (150 µl) of the enzyme assay were transferred to dichloromethane (50 µl) and vortexed for 3 s, in order to stop enzymatic activity, and centrifuged (15 min, 21.380 x g, 4°C); 100 µl of the upper aqueous phase was transferred into HPLC vials. Substrate disappearance and product formation was analyzed *via* HPLC-MS (see above) but with additional screening for the ions of sulfo-EMP or sulfo-ED pathway intermediates (**Denger et**

al., 2014): *m/z* $[M-H]^-$ of 241 for sulfogluconolactone and 2-keto-3,6-dideoxy-6-sulfogluconate, 256 for 6-deoxy-6-sulfogluconate and 323 for 6-deoxy-6-sulfofructose-1-phosphate.

For enzymatic assays with the SFT pathway enzymes in cell extracts of strain SOS1, the $(NH_4)_2CO_3$ buffered cell extract or gel-filtered cell extract (see above) was used (0.1 mg/ml protein) and incubated in 50 mM $(NH_4)_2CO_3$ buffer (pH 9.0) containing 1 mM 1,4-dithiothreitol (DTT), 1 mM $MnCl_2$, 0.5 mM $MgCl_2$, 6 mM SQ or $^{13}C_6$-SQ, 4 mM NAD^+ and optionally 6 mM erythrose-4-phosphate or 6 mM glyceraldehyde-3-phosphate as acceptor. Samples (150 µl) for HPLC-MS analysis were taken at intervals and transferred to dichloromethane (50 µl) and vortexed for 3 s, in order to stop enzymatic activity, and centrifuged (15 min, 21.380 x g, 4°C); 100 µl from the aqueous phase was transferred into HPLC vials.

***In-vitro* reconstitution of the *B. aryabhattai* SOS1 SFT pathway with recombinant enzymes.** Cloning of genes, and heterologous expression and purification of recombinant proteins *via* His-tag, was performed following a previously published protocol (**Felux et al., 2013**). The PCR primers used for directional cloning are listed in **Table S1**. PCR reaction mixtures consisted of 0.5-fold HF buffer (Thermo Fisher Scientific) containing 0.2 mM dNTP mix, 1 µM forward primer and 1 µM reverse primer, 5 ng genomic DNA of strain SOS1, and 20 U/ml Phusion High Fidelity DNA-Polymerase (Thermo Fisher Scientific). The PCR program for all genes was: initial denaturation at 98°C for 2 min; 25 cycles of 30 s denaturation at 98°C, 45 s annealing at 58°C, and 1.5 min elongation at 72°C; final elongation for 5 min at 72°C. PCR products were purified using the DNA Clean & Concentrator-5 kit (Zymo Research) and 5 ng of purified PCR product was used in the TOPO cloning reaction (Champion pET101 Directional TOPO Expression, Invitrogen). The cloning reaction was used to transform chemically competent *E. coli* NovaBlue (Novagen) for genes 1320, 1322, 1323 and 1325, and OneShot TOP 10 *E. coli* (Invitogen) for genes 1321 and 1323. The colony PCR reaction mixture

contained each 1 µM T7 forward and reverse primer (**Table S1**). The PCR program was: initial denaturation at 95°C for 10 min; 35 cycles of 45 s denaturation at 95°C, 45 s annealing at 55°C, and 1.5 min elongation at 72; final elongation step for 5 min at 72°C. Plasmids were prepared from overnight LB-liquid cultures (3.5 ml) containing 150 µg/ml ampicillin, using the Zyppy Plasmid Miniprep kit (Zymo Research). The correct inserts were confirmed by Sanger sequencing (Eurofins, Germany) using T7 primers (see above). The purified plasmids were used to transform chemical competent *E. coli* Rosetta 2 (DE3) (Invitrogen) according to Invitrogen's manual. For expression, transformands were grown in LB medium containing 100 µg/ml ampicillin and 35 µg/ml chloramphenicol, at 37°C and 200 rpm shaking, first in the 20-ml scale overnight and then in the 1-liter scale. These cultures were grown to an OD_{580nm} of 0.8-1.0. Then, IPTG (1 mM) was added to induce expression of the recombinant protein, and 3% (v/v) ethanol was added to induce chaperon formation (**Neidhardt et al., 1987; Thomas and Baneyx 1997**). After induction, the cultures were transferred to 15°C, 200 rpm shaking, for 24-25 h. Cells were harvested by centrifugation (15 min, 5000 x g, 4°C) and the cell pellets were stored at -20°C. Cells were resuspended in 2.5 ml of 10 mM imidazole buffer (pH 7.8) containing 1 U/ml DNase. Cell extracts were prepared by five passages through a French pressure cell (see above) and removal of the cell debris by centrifugation (15 min, 21.380 x g, 4°C). The overexpressed proteins were purified *via* His_6-tag through His SpinTrap columns (GE Healthcare Life Sciences); all centrifugation steps were performed at 200 x g for 30 sec at 4°C. For washing, 600 µl of 40 mM imidazole buffer (pH 7.8) and subsequently 600 µl of 60 mM imidazole buffer (pH 7.8) were used. The recombinant protein was eluted by adding three times 200 µl of 400 mM imidazole buffer (pH 7.8). The eluates were pooled and then desalted against 10 mM imidazole buffer (pH 7.8) (illustra NAP-10 gel filtration columns, GE Healthcare Life Sciences). Protein concentration in the recombinant protein preparations was determined by Bradford assay (see above). Enzymatic activities of the recombinantly produced enzymes was monitored by HPLC-MS in reactions containing 50 mM $(NH_4)_2CO_3$ buffer (pH 9.0), 1 mM

DTT, 1 mM $MnCl_2$, 0.5 mM $MgCl_2$ and 2 mM $^{13}C_6$-SQ, at room temperature. Optionally, 6 mM F6P or G6P or 12 mM GAP as acceptor was added. Thereafter, the assay was started by addition of SQ isomerase. Then, SF transaldolase was added, and finally, SLA dehydrogenase and 6 mM NAD^+. Before and after addition of each enzyme, samples were taken for HPLC-MS analysis (see above).

Photometric assays of SLA dehydrogenase activity. Recombinant SLA dehydrogenase activity was assayed photometrically using 10 µg/ml enzyme in 50 mM piperazine-N,N′-bis(2-ethanesulfonic acid) (PIPES) buffer (pH 6.8) containing 1 mM DTT, 1 mM $MnCl_2$, 0.5 mM $MgCl_2$, and NAD^+ or NADH in concentrations as specified below. The reduction of NAD^+ or oxidation of NADH was followed as absorbance at 365 nm in a spectrophotometer (Uvicon 922, Kontron Instruments); additionally, samples for HPLC-MS analysis of formation or disappearance of SL, SLA and/or DHPS were taken before, in between, and after the reactions (see above). Since we had no authentic SLA available as substrate, SLA was generated from DHPS (2 mM) in a reverse reaction of recombinant *E.coli* SLA reductase YihU (**Felux et al., 2015**, **Burrichter et al., 2018**) (1 µg/ml YihU; prepared as described previously, ref. **Denger et al., 2014**) in presence of NAD^+ (4 mM); subsequently, SLA dehydrogenase was added (100 µg/ml). A reverse reaction of SLA dehydrogenase 1325 (100 µg/ml) with SL (5 mM) and NADH (0.1 mM) was tested negative. For determination of promiscuous activities of the SLA dehydrogenase, GAP or E4P (6 mM) were used as substrates with NAD^+ (4 mM).

Isolation and genome sequencing of *Clostridium symbiosum* LT0011. *C. symbiosum* LT0011 was isolated from fecal material suspended in 30 mM bicarbonate buffer followed by serial dilution and plating on modified YCFA agar (sulfur compound- and short-chain fatty acid-free, 1.5% agar) (**Lopez-Siles et al., 2012**) supplemented with 10 mM SQ as the sole carbon and energy source. Colonies were streaked onto fresh plates and characterized by colony PCR and

sequencing of 16S rRNA genes (primers 8F and 1492R; **Table S1**). For genome sequencing, DNA was extracted using the Wizard Genomic DNA Purification KIT (Promega) and diluted to 0.1 ng/µl in 130 µl and sheared on a Covaris S220 Focused-ultrasonicator Instrument (Covaris, USA) to a target length of 350 bp. Library preparation was conducted according to the NEBNext Ultra II DNA Library Prep Kit for Illumina protocol (New England BioLabs). Indexing primers were ligated conforming to NEBNext Multiplex Oligos for Illumina manual (Index Primers Set 1, New England BioLabs). Samples were submitted for 150 bp paired-end sequencing on an Illumina HiSeq 3000/4000 at the Biomedical Sequencing Facility (BSF) of the Research Centre for Molecular Medicine (CeMM, Vienna, Austria). Additionally, a DNA library was prepared for MinION sequencing according to the manufacturer's protocol using a R9.4/FLO-MIN106 Flow Cell (Oxford Nanopore Rapid Sequencing, Oxford Nanopore Technologies, Oxford, UK). Illumina and Nanopore sequences were trimmed and filtered (> 100 bp and > 1 kb, respectively; PRINSEQlite v 0.20.4) (**Schmieder and Edwards 2011**). A hybrid assembly of both Illumina and Nanopore reads was performed using SPAdes (v.3.12.0) (**Bankevich et al., 2012**). Genome annotation was performed using PROKKA v.1.12 (**Seemann 2014**).

Cultivation of SQ-fermenting strains, total proteomics and detection of fermentation products. *Enterococcus gilvus* DSM15689 (**Tyrrell et al., 2002**) and *Eubacterium rectale* (*Agathobacter rectalis*) DSM17629 (**Duncan and Flint 2008**) were purchased from the Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures GmbH (Braunschweig). *Clostridium symbiosum* LT0011 (DSM180250) was isolated in this study (see above). For growth experiments and for generation of cell extracts for total proteomics (see above), the strains were grown anaerobically in a carbonate-buffered (pH 7.1) mineral salts medium reduced with 1 mM titanuim(III)nitriloacetate (Ti(III)-NTA) as reducing agent (basal medium, ref. **Widdel and Pfennig 1981**; trace elements solution, ref. **Widdel et al., 1983**; Ti(III)NTA solution, ref. **Moench and Zeikus 1983**; selenium-tungstate solution, ref.

Tschech and Pfennig 1984; vitamin solution, ref. **Pfennig 1978**) with 10 mM SQ and yeast extract (0.1% *w*/*v*) as supplement. The cultures were incubated in serum bottles enclosed with butyl-rubber stoppers under a N_2/CO_2 (80:20) gas atmosphere. Inoculation and sampling was done through the rubber stoppers with syringe and needle. Samples were taken at intervals for monitoring growth (OD580) (see above) and for HPLC analysis of SQ, DHPS and SL (see above). Short chain fatty acids and alcohols against authentic standards were analyzed on a HPLC system (system 10A, Shimadzu) with an Aminex column (HPX-87H, BioRad) and a refractive index detector (RID-10A, Shimadzu) at 60°C. The eluent was 5 mM H_2SO_4 at an isocratic flow of 0.6 ml/min. Under these conditions, the fermentation product acetate eluted at 11.7 min. This Aminex HPLC-RID method generally separates and detects carbohydrates, carboxylic acids, short-chain fatty acids, alcohols, ketones and other metabolites; no other products (peaks) than acetate were detected in any growth experiment. Hydrogen production was tested with a Peak Perfomer 1 (Peak Laboratories) trace gas analyzing gas chromatography system with a reducing compound photometer (RCP) detector; N_2 was used as carrier gas. For harvesting of the cultures, the serum bottles were opened and the cell suspension centrifuged under air atmosphere; the cell were disrupted, the cell extracts prepared, and the total proteomic analyses were done as, described above in the appropriate sections for *B. aryabhattai* SOS1.

SUPPLEMENTAL REFERENCES

Bankevich, A., S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, S. I. Nikolenko, S. Pham, A. D. Prjibelski, A. V. Pyshkin, A. V. Sirotkin, N. Vyahhi, G. Tesler, M. A. Alekseyev and P. A. Pevzner (2012). "SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing." Journal of Computational Biology **19**(5): 455-477.

Bolger, A. M., M. Lohse and B. Usadel (2014). "Trimmomatic: a flexible trimmer for Illumina sequence data." Bioinformatics **30**(15): 2114-2120.

Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." Analytical Biochemistry **72**(1): 248-254.

Burrichter, A., K. Denger, P. Franchini, T. Huhn, N. Müller, D. Spiteller and D. Schleheck (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." Frontiers in Microbiology **9**: 2792-2792.

de Souza, L. M., M. Müller-Santos, M. Iacomini, P. A. J. Gorin and G. L. Sassaki (2009). "Positive and negative tandem mass spectrometric fingerprints of lipids from the halophilic Archaea *Haloarcula marismortui*." Journal of Lipid Research **50**(7): 1363-1373.

Denger, K., T. Huhn, K. Hollemeyer, D. Schleheck and A. M. Cook (2012). "Sulfoquinovose degraded by pure cultures of bacteria with release of C3-organosulfonates: complete degradation in two-member communities." FEMS Microbiolgy Letters **328**(1): 39-45.

Denger, K., M. Weiss, A. K. Felux, A. Schneider, C. Mayer, D. Spiteller, T. Huhn, A. M. Cook and D. Schleheck (2014). "Sulphoglycolysis in *Escherichia coli* K-12 closes a gap in the biogeochemical sulphur cycle." Nature **507**: 114-117.

Duncan, S. H. and H. J. Flint (2008). "Proposal of a neotype strain (A1-86) for *Eubacterium rectale*. Request for an Opinion." International Journal of Systematic and Evolutionary Microbiology **58**(7): 1735-1736.

Felux, A.-K., K. Denger, M. Weiss, A. M. Cook and D. Schleheck (2013). "*Paracoccus denitrificans* PD1222 utilizes hypotaurine via transamination followed by spontaneous desulfination to yield acetaldehyde and, finally, acetate for growth." Journal of Bacteriology **195**(12): 2921-2930.

Felux, A.-K., D. Spiteller, J. Klebensberger and D. Schleheck (2015). "Entner-Doudoroff pathway for sulfoquinovose degradation in *Pseudomonas putida* SQ1." Proceedings of the National Academy of Sciences **112**(31): E4298-E4305.

Kennedy, S. and C. Fewson (1968). "Enzymes of the mandelate pathway in bacterium NCIB 8250." Biochemical Journal **107**(4): 497.

Kolmogorov, M., B. Raney, B. Paten and S. Pham (2014). "Ragout - a reference-assisted assembly tool for bacterial genomes." Bioinformatics **30**(12): i302-309.

Lopez-Siles, M., T. M. Khan, S. H. Duncan, H. J. Harmsen, L. J. Garcia-Gil and H. J. Flint (2012). "Cultured representatives of two major phylogroups of human colonic *Faecalibacterium prausnitzii* can utilize pectin, uronic acids, and host-derived substrates for growth." Applied and Environmental Microbiology **78**(2): 420-428.

Luo, R., B. Liu, Y. Xie, Z. Li, W. Huang, J. Yuan, G. He, Y. Chen, Q. Pan, Y. Liu, J. Tang, G. Wu, H. Zhang, Y. Shi, C. Yu, B. Wang, Y. Lu, C. Han, D. W. Cheung, S. M. Yiu, S. Peng, Z. Xiaoqian, G. Liu, X. Liao, Y. Li, H. Yang, J. Wang and T. W. Lam (2012). "SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler." Gigascience **1**(1): 18.

Mayer, J., T. Huhn, M. Habeck, K. Denger, K. Hollemeyer and A. M. Cook (2010). "2,3-Dihydroxypropane-1-sulfonate degraded by *Cupriavidus pinatubonensis* JMP134: purification of dihydroxypropanesulfonate 3-dehydrogenase." Microbiology **156**(5): 1556-1564.

Moench, T. T. and J. G. Zeikus (1983). "An improved preparation method for a titanium (III) media reductant." Journal of Microbiological Methods **1**(4): 199-202.

Neidhardt, F. C., J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter and H. Umbarger (1987). *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, American Society for Microbiology Washington, DC.

Peck, S. C., K. Denger, A. Burrichter, S. M. Irwin, E. P. Balskus and D. Schleheck (2019). "A glycyl radical enzyme enables hydrogen sulfide production by the human intestinal bacterium *Bilophila wadsworthia*." Proceedings of the National Academy of Sciences **116**(8): 3171-3176.

Pfennig, N. (1978). "*Rhodocyclus purpureus gen. nov*. and *sp. nov.*, a ring-shaped, vitamin B12-requiring member of the family *Rhodospirillaceae*." Int. J. Syst. Evol. Microbiol. **28**(2): 283-288.

Schmidt, A., N. Müller, B. Schink and D. Schleheck (2013). "A Proteomic view at the biochemistry of syntrophic butyrate oxidation in *Syntrophomonas wolfei*." PLOS ONE **8**(2): e56905.

Schmieder, R. and R. Edwards (2011). "Quality control and preprocessing of metagenomic datasets." Bioinformatics **27**: 863-864.

Seemann, T. (2014). "Prokka: rapid prokaryotic genome annotation." Bioinformatics **30**(14): 2068-2069.

Styp von Rekowski, K. , K. Denger and A.M. Cook (2005). "Isethionate as a product from taurine during nitrogen-limited growth of *Klebsiella oxytoca* TauN1." Archives of Microbiology **183**: 325-330.

Thomas, J. G. and F. Baneyx (1997). "Divergent effects of chaperone overexpression and ethanol supplementation on inclusion body formation in recombinant *Escherichia coli*." Protein Expression and Purification **11**(3): 289-296.

Thurnheer, T., T. Köhler, A. M. Cook and T. Leisinger (1986). "Orthanilic acid and analogues as carbon sources for bacteria: growth physiology and enzymic desulphonation." Microbiology **132**(5): 1215-1220.

Tschech, A. and N. Pfennig (1984). "Growth yield increase linked to caffeate reduction in *Acetobacterium woodii*." Arch. Microbiol. **137**: 163-167.

Tyrrell, G. J., L. Turnbull, L. M. Teixeira, J. Lefebvre, M. d. G. S. Carvalho, R. R. Facklam and M. Lovgren (2002). "*Enterococcus gilvus* sp. nov. and *Enterococcus pallens* sp. nov. isolated from human clinical specimens." Journal of Clinical Microbiology **40**(4): 1140-1145.

Widdel, F., G.-W. Kohring and F. Mayer (1983). "Studies on dissimilatory sulfate-reducing bacteria that decompose fatty acids III. Characterization of the filamentous gliding *Desulfonema limicola gen. nov. sp. nov*., and *Desulfonema magnum sp. nov.*" Arch. Microbiol. **134**(4): 286-294.

Widdel, F. and N. Pfennig (1981). "Studies on dissimilatory sulfate-reducing bacteria that decompose fatty acids. I. Isolation of new sulfate-reducing bacteria enriched with acetate from saline environments. Description of *Desulfobacter postgatei gen. nov., sp. nov*." Arch. Microbiol. **129**(5): 395-400.

William, S., H. Feil and A. Copeland (2012). "Bacterial genomic DNA isolation using CTAB." Sigma **50**: 6876.