
Article:
Snow, Alexander J.D., Burchill, Laura, Sharma, Mahima orcid.org/0000-0003-3960-2212 et al. (2 more authors) (2021) Sulfoglycolysis:catabolic pathways for metabolism of sulfoquinovose. Chemical Society Reviews. ISSN 0306-0012

https://doi.org/10.1039/D1CS00846C

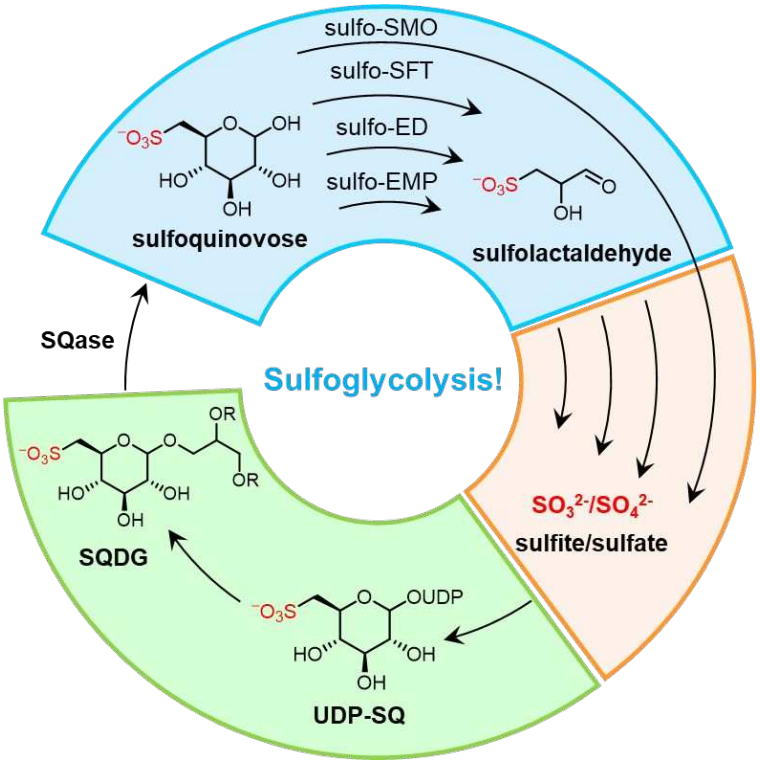
Chem Soc Rev

Sulfoglycolysis: catabolic pathways for the breakdown of sulfoquinovose

Journal:	*Chemical Society Reviews*
Manuscript ID	CS-SYN-09-2021-000846.R2
Article Type:	Review Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Snow, Alexander; University of York, Chemistry Burchill, Laura; University of Melbourne, School of Chemistry Sharma, Mahima; University of York, Department of Chemistry Davies, Gideon; University of York, Gideon Davies; University of York, Gideon Davies Williams, Spencer; University of Melbourne, School of Chemistry

Graphical abstract

A biochemical, structural and mechanistic perspective on the pathways of sulfoglycolysis for catabolism of the sulfosugar sulfoquinovose.



Sulfoglycolysis: catabolic pathways for metabolism of sulfoquinovose

Alexander J.D. Snow,[1] Laura Burchill,[2,3] Mahima Sharma,[1]* Gideon J. Davies,[1]* Spencer J. Williams[2,3]*

[1] York Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, YO10 5DD, U.K.
[2] School of Chemistry, University of Melbourne, Parkville, Victoria 3010, Australia.
[3] Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Victoria 3010, Australia

Email: sjwill@unimelb.edu.au, gideon.davies@york.ac.uk, mahima.sharma@york.ac.uk

Keywords: sulfur cycle; X-ray crystallography; enzyme mechanism; sulfonate; environmental microorganisms

Running title: Sulfoglycolysis pathways

Abstract

Sulfoquinovose (SQ), a derivative of glucose with a C6-sulfonate, is produced by photosynthetic organisms and is the headgroup of the sulfolipid sulfoquinovosyl diacylglycerol. The degradation of SQ allows recycling of its elemental constituents and is important in the global sulfur and carbon biogeochemical cycles. Degradation of SQ by bacteria is achieved through a range of pathways that fall into two main groups. One group involves scission of the 6-carbon skeleton of SQ into two fragments with metabolic utilization of carbons 1-3 and excretion of carbons 4-6 as dihydroxypropanesulfonate or sulfolactate that is biomineralized to sulfite/sulfate by other members of the microbial community. The other involves the complete metabolism of SQ by desulfonylation involving cleavage of the C-S bond to release sulfite and glucose, the latter of which can enter glycolysis. The discovery of sulfoglycolytic pathways has revealed a wide range of novel enzymes and SQ binding proteins. Biochemical and structural characterization of the proteins and enzymes in these pathways have illuminated how the sulfonate group is recognized by Nature's catalysts, supporting bioinformatic annotation of sulfoglycolytic enzymes, and has identified functional and structural relationships with the pathways of glycolysis.

1. Introduction

Sulfoquinovose (SQ) is a sulfonated hexose analogous to D-glucose, but which contains a sulfur-carbon bond **(Fig. 1)**.[1-3] SQ is produced by photosynthetic organisms and is the anionic headgroup of the sulfolipid sulfoquinovosyl diacylglycerol (SQDG) found in plants, alga, and cyanobacteria. SQ occurs within the related metabolites sulfoquinovosyl monoglyceride (SQMG, *lyso*-sulfolipid),[4] sulfoquinovosyl glycerol (SQGro),[5] sulfoquinovosyl glyceryl ether,[6] and 2'-*O*-acyl-SQDG.[7] SQ is also present within the N-linked glycan of the archaeon *Sulfolobus acidocaldarius*.[8] SQDG is localized to the thylakoid membrane, is found in intimate association with photosynthetic protein complexes,[9, 10] and assists in the function of photosystem II.[11] The biosynthesis of SQ occurs at the level of the sugar nucleotide diphosphate, by reaction of sulfite with UDP-glucose to give UDP-SQ (catalyzed by UDP-sulfoquinovose synthase), which is then transferred to diacylglycerol to give SQDG (catalyzed by the glycosyltransferase SQDG synthase).[12, 13] Owing to its production within photosynthetic tissues,[1, 12, 14] SQ constitutes a major reservoir of organosulfur, with one estimate of its production standing at 10 billion tonnes annually.[15] Regardless of the precise amount, this places SQ among other important organosulfur compounds in the biosphere including the amino acids cysteine and methionine, the simple sulfonate 2,3-dihydroxypropanesulfonate (DHPS)[16] and the osmolytes dimethylsulfoniopropionate[17] and dimethylsulfoxonium propionate.[18]

SQDG is a sulfur reservoir in photosynthetic organisms, and a carbon and sulfur source for diverse microbial communities. Shortly after the discovery of SQ, and amid growing appreciation for its ubiquity in photosynthetic organisms, it was realised that breakdown pathways must exist. Early hypotheses suggested that SQ deconstruction could be analogous to glycolysis, and thus was described as sulfoglycolysis.[19] Subsequently, various organisms were isolated that could grow on SQ as sole carbon source and it was found that they produced a range of end-products including DHPS, sulfolactate (SL) and sulfate.[20-22] While early studies identified possible breakdown products of SQ, until relatively recently it proved difficult to elucidate the individual steps.

In recent years, a series of breakthroughs have led to the identification of four pathways of sulfoglycolysis and has ushered in a new era of genetic, bioinformatic, structural, biochemical and microbiological studies of SQ **(Fig. 1)**. These four pathways fall into two main groups depending on whether they allow microorganisms to use 3 or 6 of the carbons within the SQ skeleton. The first group is comprised of bacteria that cleave the 6-carbon chain of SQ into two C3 chains, leading to production of DHPS or pyruvate that is utilized by the organism, and sulfolactaldehyde (SLA), which is oxidized to SL or reduced to DHPS then excreted. These organisms use one of three pathways: the sulfoglycolytic Embden-Meyerof-Parnas

(sulfo-EMP) **(Fig. 2)**,[23] sulfoglycolytic Entner-Doudoroff (sulfo-ED) **(Fig. 3)**[24] and sulfoglycolytic sulfofructose transaldolase (sulfo-SFT; also termed the sulfoglycolytic transaldolase, sulfo-TAL) **(Fig. 4)**[25, 26] pathways. The excreted SL or DHPS becomes available to organisms that utilize its carbon and effect the biomineralization of the sulfonate group to sulfite or sulfate.[27] The second group contains bacteria that use a fourth pathway, termed the sulfoglycolytic sulfoquinovose monooxygenase (sulfo-SMO) pathway **(Fig. 5)**, which produces glucose and sulfite and is the only pathway in a single organism that results in cleavage of the sulfur-carbon bond of SQ to allow complete utilization of all six of its carbons.[28]

2. Early observations on the catabolism of sulfoquinovose and sulfoquinovosyl glycerol

Early studies of the metabolism of SQ and SQGro focused on algae and plants. The formation and breakdown of ^{35}S-SQDG was studied in the alga *Chlorella ellipsoidea* after labelling with ^{35}S-sulfate. Using 2D-thin layer chromatography a suite of radiolabelled products were observed that included SQGro, SQ, DHPS, SL and SLA.[29] SLA condensed with dihydroxyacetone phosphate in the presence of rabbit muscle aldolase to give a sulfoketose phosphate.[29] This data were used to suggest the existence of a sulfoglycolytic pathway in *Chlorella.*[19] In alfalfa leaves, SQGro is broken down to SQ, SLA and SL, showing the existence of a sulfoglycolytic pathway in plants.[30] In coral tree (*Erythrina crista-galli*), the only product of SQGro breakdown observed was sulfoacetic acid, which was proposed to result from decarboxylation of SL and then oxidation.[30] A recent survey of algae identified DHPS in many but not all species investigated.[31] Bioinformatic analysis of macro- and microalgae revealed only limited occurrence of SQase and SLA reductase homologues, but with no evidence of an intact sulfo-EMP pathway.[31] Likewise, while homologues of SLA dehydrogenase, sulfogluconate (SG) dehydratase and SQ dehydrogenase were observed in the majority of algal species, homologues of SQase and 2-keto-3-deoxysulfogluconate (KDSG) aldolase were found only in a smaller subset and no complete sulfo-ED pathway could be identified. These data suggest that either algae use other pathways for SQ catabolism, or do not accomplish complete sulfoglycolysis.

Other early studies examined the breakdown of SQ by saprophytic soil microorganisms (ie those that digest decaying organic matter). Field studies involving the incubation of SQ in assorted forest soils revealed its rapid mineralization to sulfate, suggesting that sulfoglycolytic microorganisms are widespread.[32] Martelli and Benson isolated a soil *Flavobacterium* sp. that converted methyl α-sulfoquinovoside to sulfoacetate in a phosphate-dependent manner.[33] Roy and co-workers isolated several SQ-metabolizing bacteria from soil and sewage sludge, which produced a range of metabolic end-products: SL, DHPS or

sulfate.[20, 21] *Klebsiella* sp. ABR11 initially produced DHPS and SL, and eventually sulfate.[21] Crude cell-free extracts from this bacterium exhibited 'phosphofructokinase' activity but using SQ as substrate, and NAD^+–dependent SQ dehydrogenase activity. Based on these results it was proposed that ABR11 performed sulfoglycolysis using pathways that were variants of the classical Embden-Meyerhof-Parnas (EMP) or Entner-Doudoroff (ED) glycolysis pathways.[21] By contrast, *Agrobacterium* sp. ABR2 produced only sulfate and bicarbonate as metabolic products, with a significant delay between SQ consumption and sulfate production.[21]

Collectively, this work provided strong evidence for the occurrence of sulfoglycolytic pathways in plants, algae and bacteria, and a framework for understanding the molecular details of sulfoglycolysis. Recent breakthroughs have placed these early observations on firmer foundations and identified dedicated sulfoglycolytic pathways in bacteria. The pathways used in plants and algae remain unclear.

3. Pre-sulfoglycolytic processing of SQDG: Delipidation, SQ glycoside hydrolysis, and mutarotation

3.1 Delipidation of SQDG: formation of SQMG and SQGro

While specific sulfolipid lipases (sulfolipases) have not yet been discovered, there is evidence for enzymes that can effect the partial and complete delipidation of SQDG **(Fig. 6a)**. Benson reported that incubation of ^{35}S-SQDG with a crude extract of the algae *Scenedesmus obliquus* resulted in the sequential formation of SQMG then SQGro.[4] Similar lipase activity on SQDG was observed in *Chlorella ellipsoidea*,[4] alfalfa (*Medicago sativa*),[4] corn-roots extracts[4] and *Chlorella pyrenoidosa.*[34] On the other hand a lipid hydrolase isolated from leaves of the scarlet runner bean, *Phaseolus multiflorus,* converted SQDG to SQGro with no SQMG observed.[35] Hazelwood and Dawson reported the isolation of a lipolytic fatty-acid requiring *Butyrivibrio* sp. S2 from sheep rumen that could catabolize SQDG but its further metabolism was not studied.[36] When *Chlorella protothecoides* was grown in light, levels of SQDG increased while levels of SQGro was stable.[37] By contrast, levels of SQDG declined in 'glucose bleached' *C. protothecoides* cells, while SQGro increased, suggesting that lipase action is connected with the bleaching action and disintegration of the photosynthetic architecture.

The hydrolysis of SQDG has been studied in mammalian systems. Pancreatin (an enzymatic extract from porcine pancreas) can effect the deacylation of SQDG.[4] Gupta and Sastry studied saline extracts of acetone-precipitated pancreas and intestinal mucosa from guinea pig, sheep and rat and observed stepwise deacylation of SQDG to SQMG (ascribed to 'sulfolipase A') and to SQGro (ascribed to 'sulfolipase B').[38] Administration of ^{35}S-SQDG to

guinea pigs resulted in rapid conversion to water soluble forms: SQGro and SO_4^{2-} supporting the conclusion that C—S bond cleavage is brought about by intestinal microflora.[38]

3.2 Hydrolysis of sulfolipid and SQ glycosides

Liberation of SQ from its glycosides is achieved by specialized glycosidases termed sulfoquinovosidases (SQases) **(Fig. 6a)**. While early reports suggested that *E. coli* β-galactosidase can act as an SQase,[39] this observation was subsequently reconized to be erroneous,[40] and in retrospect the most reasonable explanation is contamination, perhaps by the *E. coli* YihQ protein. All characterized sulfoglycolysis gene clusters characterized to date contain a gene encoding an SQase belonging to glycoside hydrolase family 31 of the CAZy sequence-based classification (www.cazy.org;[41] www.cazypedia.org[42]): YihQ from *E. coli*,[43] SftG from *Bacillus aryabhattai*,[25] SqvC from *Bacillus megaterium*,[26] PpSQ1_00425 from *Pseudomonas putida* SQ1,[24] *Rl*SQase from *Rhizobium leguminosarum* SRDI565,[44] and SmoI (AtSQase) from *A. tumefaciens*.

Phylogenies of family GH31 proteins using hidden Markov models (a statistical method used for pattern recognition)[45] revealed that SQases occupy a small sub-group in this large family that is dominated by α-glucosidases and includes α-xylosidase, α-galactosidases, α-*N*-acetylgalactosaminidases and α-glucan lyases.[46, 47] SQases operate through a retaining mechanism involving a two-step double displacement via a glycosyl enzyme intermediate.[43] In the first step a general acid residue (Asp472 in *E. coli* YihQ) assists the departure of the leaving group while a second carboxylate (Asp405) acts as a nucleophile to form the glycosyl enzyme intermediate with inversion of anomeric stereochemistry. In the second step Asp472 acts as a general base to deprotonate a water molecule and assist the hydrolysis of the glycosyl enzyme with a second inversion of anomeric stereochemistry, leading to release of α-SQ **(Fig. 6c)**. YihQ SQase acts on both SQDG and SQGro and displays a 6-fold preference for the natural 2'*R*-SQGro stereoisomer, which is also preferred by SmoI.[48] 4-Nitrophenyl α-D-sulfoquinovoside (PNPSQ) is an effective chromogenic substrate for *E. coli* YihQ,[43] *A. tumefaciens* SQase[48] and *Rhizobium leguminosarum* SQase,[44] enabling their characterization using real-time assays with a UV/Vis spectrophotometer. On the other hand the fluorogenic substrate 4-methylumbelliferyl α-D-sulfoquinovoside (MUSQ) is a poor substrate for YihQ and SmoI with k_{cat}/K_M values some 10^4-10^5-fold lower, indicating that the SQase active site has difficulty accommodating the bulky methylumbelliferone group.[49]

X-ray structures of two SQases have been reported that reveal structurally homologous $(\alpha/\beta)_8$ barrel folds.[43, 48] Complexes of wildtype *E. coli* YihQ and *A. tumefaciens* SmoI enzymes with the aza-sugar competitive inhibitor sulfoquinovose-isofagomine (SQ-IFG), and of the catalytically disabled acid-base mutant with the artificial substrate PNPSQ identified

a conserved sulfonate binding pocket comprised of Arg301, Trp304 and Tyr508 via a bridging water molecule (*E. coli* YihQ numbering). The RWY sulfonate binding triad is present within predicted GH31 enzymes from plants, bacteria, fungi, animals and protists. Second sphere residues around the active site are largely conserved, but with two subgroups containing either QQ (*E. coli* Q262-Q288) or KE (*A. tumefaciens* K245-E270) residues.[48] Mutant enzymes with partially switched second-shell residues were around 1000-fold less active than wildtype. Mutants with fully swapped neutral pairs were less active than wildtype but were 10-fold more active than the partially swapped mutants, showing that charge neutrality is required for optimum SQase activity **(Fig. 6b)**.

SQase activity serves to explain the ability of *E. coli* K-12, *R. leguminosarum* SRDI585, *P. putida* SQ1 and *B. aryabhattai* SOS1 to grow on SQGro as sole carbon source.[25, 44, 48] In the case of *E. coli*, growth on SQGro led to cell densities similar to growth on glucose, while growth on SQ led to similar density to Gro, consistent with the utilization of 6C and 3C in the respective substrate pairs.[48]

3.3 Sulfoquinovose mutarotase

Sulfoglycolysis gene clusters from sulfo-EMP and sulfo-ED pathways typically contain genes annotated as aldose-1-epimerase (also termed mutarotases) that are upregulated upon growth on SQ **(Fig. 2 and 3)**.[23, 24] Mutarotases catalyse the interconversion of anomers of sugars, and in the case of SQ, mutarotation is a relatively slow process, with a half-life of 300 min under phosphate free conditions at 26 °C.[50] A sole SQ mutarotase (*Hs*SQM) from the putative sulfo-ED gene cluster from *Herbaspirillum seropedicaea* AU14040 has been experimentally studied using the nuclear magnetic resonance spectroscopy technique of chemical exchange spectroscopy.[50] *Hs*SQM catalyzes the mutarotation of SQ and glucose-6-phosphate (G6P), but not glucuronic acid, mannose, glucose or galactose **(Fig. 6d)**. Unidirectional rate constants at equilibrium were measured, and revealed that the k_{cat}/K_M value for SQ mutarotation is 5-fold higher than for G6P. Combining these results with the reported rate for spontaneous mutarotation of G6P ($t_{1/2}$ = 6 sec),[51, 52] allowed calculation of the proficiency ratio $(k_{cat}/K_M)/k_{uncat}$, which revealed that *Hs*SQM is 17,000-fold more proficient as a catalyst for enhancing the rate of SQ mutarotation compared to G6P.[50] HsSQM shares highly conserved residues with other hexose mutarotases: His92, His162 and Glu254.[50] A histidine and Glu254 are proposed to act in roles of general acid and general base,[53, 54] respectively, in the first half of the reaction leading to the acyclic aldehyde.[50]

4. The sulfoglycolytic Embden-Meyerhof-Parnas (sulfo-EMP) pathway

Denger *et al.* and co-workers reported that *E. coli* K-12 (strains BW25113, DH1, MG1655 and W3100) grow on SQ under aerobic conditions.[23] Working with strain MG1655 they observed that consumption of SQ coincided with release of DHPS into the culture media. Addition of a DHPS-degrading bacterium, *Cupriavidus pinatubonensis* JMP134, to the spent culture medium resulted in consumption of DHPS and production of sulfate. By comparative two-dimensional polyacrylamide gel electrophoresis (2D SDS-PAGE) of glucose and SQ-grown *E. coli*, and peptide fingerprinting–mass spectrometry sequencing of differentially upregulated proteins excised from the gel, the responsible gene cluster was identified as *yihO-yihW* **(Fig. 2)**.[23] This operon had previously been assigned as encoding O-antigen biosynthesis in *Salmonella enterica*.[55] Single-gene knockouts of *yihO*, *yihS*, *yihT* and *yihV* in strain BW25113 did not grow on SQ, supporting the contribution of these genes to SQ catabolism. This same gene cluster contributes to growth of *E. coli* K-12 under anaerobic conditions, where a mixed type of fermentation produces DHPS, succinate, acetate and formate.[56]

The sulfo-EMP gene cluster of *E. coli* encodes: a predicted transporter (YihO) for importing SQ and SQ glycosides; an SQase (YihQ) to hydrolyze SQ glycosides (*vide supra*); SQ mutarotase (YihR) to catalyze conversion of α-SQ to β-SQ (*vide supra*); aldose–ketose isomerase (YihS) to isomerize SQ to sulfofructose (SF); SF kinase (YihV) to phosphorylate SF to SF-1-phosphate (SFP); SFP aldolase (YihT) to convert SFP into dihydroxyacetone phosphate and SLA; SLA reductase (YihU) to convert SLA into DHPS; and a second predicted transporter (YihP) that may export DHPS from the cell. *yihW* was renamed *csqR* and encodes a DeoR-type transcription factor for the operon. Overall, the sulfo-EMP pathway shares remarkable similarity with the EMP pathway of glycolysis.[57] However, consistent with specialization to act on SQ, the sulfo-EMP pathway is induced upon growth on SQ and enzymes within the pathway have undetectable reactivity on the corresponding intermediates in glycolysis.[58, 59] This is likely to be of significance as growth under sulfoglycolytic conditions requires gluconeogenesis to produce intermediates to supply the pentose phosphate pathway (PPP) and cell wall biogenesis. If the sulfoglycolytic enzyme SF kinase acted on G6P, this would result in a futile cycle that would consume ATP. Bioinformatics analysis showed that the sulfo-EMP pathway is present in the majority of commensal and pathogenic *E. coli* strains, and a wide range of other Enterobacteriacaea including *Salmonella enterica*, *Chronobacter sakazakaii*, *Klebsiella oxytoca* and *Pantoea anantis*, suggesting a role in sulfoglycolysis in the gastrointestinal tract of omnivores and herbivores.[23]

4.1 Sulfoquinovose-sulfofructose isomerase (YihS)

SQ-SF isomerase activity was demonstrated for *E. coli* YihS using recombinantly expressed protein and demonstrating conversion by liquid chromatography-mass spectrometry.[23] Sharma *et al.* re-examined the activity of YihS (and the homolog from *S. enterica*, *Se*YihS) and showed that these proteins catalyze the equilibration of SQ to SF and sulforhamnose (SR, the C2-epimer of SQ) in an equilibrium ratio of 30:21:49, respectively.[58] YihS also acts as a glucose:fructose:mannose isomerase, with the activity for mannose as substrate 178-fold lower than for SQ.[58, 60] ^{1}H NMR spectroscopy was used to monitor H/D exchange at C2 catalyzed by YihS, which revealed that YihS acts preferentially on β-SQ.[58] YihS was inactive on the glycolytic intermediate G6P.[58]

YihS and SeYihS form hexamers in solution, with each monomer exhibiting an α_6/α_6 barrel fold.[58, 60] A 3D structure of a complex of *Se*YihS-H248A with β-SF helped define the active site architecture **(Fig. 7a)**.[58] Overlay of the structure of the mutant complex with the wildtype structure showed that His248 and His383 are positioned on the α-face of SF and two loops move to enclose the ligand in the active site. The sulfonate of SF is bound by Arg55-Gln379-Gln362; this active site architecture is shared with YihS. Collectively, the structural and biochemical data is supportive of a mechanism involving deprotonation of C-2 in acyclic SQ to give a 1,2-enediol **(Fig. 8a)**. Protonation at C-1 forms SF, protonation at C-2 from the bottom face forms SR while protonation at C-2 from the top face regenerates SQ.

4.2 Sulfofructose kinase (YihV)

Denger *et al.* demonstrated *E. coli* YihV is an SF kinase using recombinantly expressed protein and demonstrating ATP-dependent conversion of *in situ* generated SF to SFP by liquid chromatography-mass spectrometry,[23] which has been confirmed using chemically synthesized SF[61]. Analysis of the thermal stability of YihV with SF, ADP, or both showed stabilisation upon sugar binding.[58] YihV activity towards SF is highly sensitive to ATP concentration: at [ATP] = 1 mM, K_M = 8 mM, while at [ATP] = 0.1 mM, K_M = 0.3 mM.[58] Sensitivity of activity to modulation by small molecules extends to other cellular metabolites: substrate inhibition by SF, product inhibition by ADP, and activation by SQ, SLA, fructose-6-phosphate (F6P), fructose bisphosphate (FBP), phosphoenolpyruvate (PEP), dihydroxyacetone phosphate (DHAP), and citrate. This data suggest that YihV is an important control point for managing flux through sulfoglycolysis in the event of large fluctuations in substrate concentration and concentrations of cellular metabolites. This modulation of activity is reminiscent of the regulatory role played by phosphofructokinase (PFK), which catalyzes the first committed step in the EMP glycolysis pathway.[62, 63] YihV was inactive on the glycolytic

intermediate F6P, showing no potential for cross-talk between sulfoglycolysis and glycolysis/gluconeogenesis.[58]

A series of 3D structures of apo YihV, and complexes of YihV with SF and ATP analogue adenylyl-imidodiphosphate (AMPPNP) or ADP, and with SFP were determined by X-ray crystallography **(Fig. 7b)**.[58] YihV is a pfkB-family ribokinase[64] with a characteristic nucleotide-binding domain and a β-domain forming a lid to enclose the substrate binding site. YihV is dimeric with a β-clasp forming the interface, with residues from the dimer interface on each subunit protruding into the active site of their counterpart, which indicates enzyme activity is dependent on dimerization. The lid and nucleotide binding domains have open and closed conformations, related by an interdomain rotation and closed upon binding of SF. SF binds in the cleft between the nucleotide binding domain and the lid, with a sulfonate-binding pocket formed by Arg138 and Asn109 from one subunit and Lys27 from the other dimer subunit. Lys27 also forms hydrogen bonds to the C1 hydroxyl and ring oxygen of SF. The YihV•SF•ADP•Mg^{2+} quaternary complex adopts the closed conformation, with the putative catalytic base Asp244 forming a hydrogen bond to the C1 hydroxyl to facilitate the nucleophilic attack on the γ-phosphate of ATP and is consistent with a mechanism involving an in-line transfer of the phosphoryl group **(Fig. 8b)**.

4.3 Sulfofructose-1-phosphate aldolase (YihT)

Denger *et al.* assigned SFP aldolase function to YihT using recombinantly-expressed protein and demonstrating ATP-dependent conversion of *in situ* generated SFP to SLA and DHAP by liquid chromatography-mass spectrometry.[23] Activity has been demonstrated for the homologue *Se*YihT from *S. enterica* using chemoenzymatically synthesized SFP[61], which was also shown to catalyze the reverse reaction, condensation of SLA and DHAP to give SFP,[58] in concordance with early observations by Benson using rabbit muscle aldolase.[19, 29] *Se*YihT was inactive on the glycolytic intermediate fructose-1,6-bisphosphate.[58]

Sequence analysis defines YihT as a class I aldolase, a group of enzymes that catalyze a retro-aldol reaction using an active-site lysine to form a Schiff base intermediate.[65] *Ec*YihT and *Se*YihT have a TIM-barrel fold and share close structural homology with other class I aldolases **(Fig. 7c)**.[58] *Se*YihT is a tetramer in solution and crystallized with 12 independent molecules in the unit cell (three tetrameric assemblies). Upon soaking with SFP, a covalent Schiff base complex of SFP with Lys193 was obtained in two of the four subunits in the tetramers, and a covalent Schiff base complex with DHAP was obtained in the other two subunits.[58] Collectively, these data are consistent with the initial formation of a Schiff base with SFP, cleavage of the C3-C4 bonds, to generate a Schiff base with DHAP and release of SLA, and finally hydrolysis of the DHAP Schiff base to release DHAP **(Fig. 8c)**. The sulfonate

group is recognized by a diad of conserved Arg253-Ala26, the latter via a bridging water molecule.

4.4 Sulfolactaldehyde reductase (YihU)

The final chemical step in the E. coli sulfo-EMP pathway is the reduction of SLA to DHPS. Denger *et al.* assigned YihT as DHPS reductase using recombinantly-expressed protein and demonstrating ATP-dependent conversion of *in situ* generated SLA to DHPS by liquid chromatography-mass spectrometry; NADPH was not a substrate.[23] Activity of recombinant YihU was confirmed using chemically-synthesized SLA.[59] YihU was inactive on the glycolytic intermediate glyceraldehyde-3-phosphate (GAP), ensuring that YihU would not consume NADH in the futile reduction of this substrate.[59] A random sequential Bi Bi mechanism was demonstrated for YihU involving sequential binding of NADH then SLA followed by the reduction to give DHPS.[59] Saito *et al.* demonstrated that YihU catalyzes succinate semialdehyde reduction, but with a catalytic efficiency 42,000-fold worse than measured for reduction of SLA.[59, 66] YihU was inhibited by the reduced NADH analogues: tetrahydro- and hexahydro-NADH.

YihU is a member of the β-hydroxyacid dehydrogenase (β-HAD) family.[67] YihU forms a solution tetramer and crystallised as a pair of intimately associated dimers **(Fig. 7d)**.[59] The protein has a two-domain architecture containing an N-terminal nucleotide binding domain forming a Rossmann fold, and a C-terminal helical bundle, connected by an interdomain helix. The protein dimerises with a domain swap of the C-terminus: helix α8 from one monomer inserts into the bundle of the counterpart. 3D structures of apo YihU and a complexes with NADH, or NADH and DHPS have been obtained by X-ray crystallography.[59] The YihU•NADH structure adopts a closed conformation when compared to the ligand-free protein, with an 8° hinge motion along the interdomain helix, resulting in enclosure of NADH. NADH specificity is ascribed to a hydrogen bond between the 2-hydroxyl on the NADH ribose ring and Asp31, which appears unable to accommodate the 2'-phosphate of NADPH. In the YihU•NADH complex, a 10.4 Å long channel decorated with a pair of positively charged residues at the entrance leads to the active site. These cationic residues may allow recruitment of anionic SLA to the YihU•NADH complex. A ternary complex of YihU•NADH•DHPS was produced by soaking DHPS into YihU•NADH crystals. This showed 2*S*-DHPS bound using through its hydroxyl groups to Lys171, and with the terminal hydroxyl group close to the dihydronicotinamide ring, consistent with a mechanism for reduction involving direct hydride transfer to SLA **(Fig. 8d)**. A dedicated sulfonate binding pocket was identified comprised of Arg123-Asn174-Ser178 that is common to all putative SLA reductases.

In the YihU•NADH•DHPS complex Arg123 within the triplet Gly122-Arg123-Thr124 interacts with one sulfonate oxygen,[59] whereas in the complex of the closely-related tartronate semialdehyde reductase GarR from *Salmonella typhimurium* bound to the substrate analogue L-tartrate the equivalent (and conserved among classical carboxylate-processing β-HADs) residues Ser123-Gly124-Gly125 exhibit a 180° flip in the central glycine allowing the carboxylate oxygens to bind to the triplet of residues Ser123-Gly124-Gly125.[68] Site-directed mutagenesis of YihU was used to probe the effect of individually converting each residue from Gly122-Arg123-Thr-124 to the corresponding residue in GarR.[59] While only minor changes were noted for k_{cat}/K_M for varying [NADH] at constant [SLA], k_{cat}/K_M values varied more significantly when varying [SLA] at constant [NADH], consistent with these residues being important in the binding of SLA.

4.5 DeoR-like transcription factor (CsqR)

The transcription factor for the sulfo-EMP operon in *E. coli* was initially named YihW and was renamed CsqR.[69] Ishihama and co-workers identified binding of CsqR at two sites: one inside the spacer between the yihUTS operon and the yihV gene, and the other upstream of the yihW gene itself **(Fig. 2)**. CsqR is a repressor for all sulfo-EMP transcription units and binding of CsqR was de-repressed by SQ and SQGro,[69] and more weakly by SR.[58] Lactose, glucose and galactose did not affect DNA binding. Atomic force microscopy of CsqR in the presence of DNA containing the CsqR binding sequence revealed large aggregates consistent with a high level of cooperativity in binding to the target DNA that were alleviated upon de-repression by SQ or SR.[58, 69]

5. The sulfoglycolytic Entner-Doudoroff (sulfo-ED) pathway

Pseudomonas putida SQ1 was isolated by enrichment culture using SQ as sole carbon source from nearshore sediment in Lake Constance (Germany).[22] Growth of SQ1 was coincident with production of equimolar SL into the growth media, and led to cell density that was approximately half that of growth on glucose.[22] By genome sequencing[70] followed by comparative 2D SDS-PAGE of glucose and SQ-grown *E. coli*, and peptide fingerprinting–mass spectrometry sequencing of differentially upregulated proteins excised from the gel, as well as total-proteome analysis, a sulfoglycolytic Entner-Doudoroff pathway encoded by *PpSQ1_00088-00100* was identified **(Fig. 3)**.[24] Five steps of the sulfo-ED pathway were reconstituted in vitro using recombinantly produced proteins, with LC-MS used to demonstrate conversion of SQ to SGL, SG, KDSG, SLA and pyruvate, and finally SL.

The sulfo-ED gene cluster of *P. putida* SQ1 encodes: a predicted importer for SQ and SQ glycosides and an exporter for SL (PpSQ1_00435, PpSQ1_00440); an SQase

(PpSQ1_00425) to hydrolyze SQ glycosides; an SQ mutarotase (PpSQ1_00415) to catalyze the equilibration of α-SQ and β-SQ; an NAD^+-dependent SQ dehydrogenase (PpSQ1_00405) to convert SQ to 6-deoxy-6-sulfogluconolactone (SGL); SGL lactonase (PpSQ1_00410) to hydrolyze SGL to 6-deoxy-6-sulfogluconate (SG); SG dehydratase (PpSQ1_00400) to convert SG to 2-keto-3,6-dideoxy-6-sulfogluconate (KDSG); KDSG aldolase (PpSQ1_00455) to cleave KDSG to SLA and pyruvate; and an $NAD(P)^+$-dependent SLA dehydrogenase (PpSQ1_00395, GabD) that converts SLA to SL. The sulfo-ED pathway shares a striking similarity with the ED pathway of glycolysis, yet SQ dehydrogenase, SG dehydratase and KDSG aldolase were inactive on G6P, phosphogluconate and keto-deoxy-phosphogluconate, respectively, showing that these enzymes are highly specific for the sulfonated substrates. Similar gene clusters were identified in a limited range of *Pseudomonas putida* strains, other γ-Proteobacteria, as well as α- and β-Proteobacteria.

Li and co-workers have studied a sulfo-ED pathway in *Rhizobium leguminosarum* SRDI565 (isolated from soil in eastern Australia) that produces SL upon growth on SQ or SQGro.[44] This pathway shares genes encoding putative enzymes that are homologous to the *P. putida* SQ1 proteins, with a key difference being the absence of a putative SQ mutarotase. SRDI565 shared the putative TauE-type SL exporter with SQ1 but it instead possessed an ABC solute importer cassette involving a putative SQ binding protein, and ABC-type permease components, also identified in the sulfo-SMO pathway.[28] This suggests that the SQ importation machinery can be interchanged between different sulfoglycolytic gene clusters. Metabolomics analysis of crude cell extracts from SRDI565 grown on SQ revealed the presence of the canonical sulfo-ED pathway intermediate SG and the endproduct SL, as well as G6P and F6P, providing evidence that gluconeogenesis is used to satisfy metabolic requirements for the PPP and cell wall biogenesis.[44] In addition, low amounts of SF and DHPS were also detected, which were proposed to arise from moonlighting activities of phosphoglucose isomerase and a non-specific reductase.

6. The sulfoglycolytic Sulfofructose Transaldolase (sulfo-SFT) pathway

Sulfoglycolytic pathways that lack the SF kinase and SFP aldolase of the sulfo-EMP pathway, but instead possess a transaldolase active on SF have been identified, termed the sulfo-SFT pathway (or the sulfoglycolytic transaldolase, sulfo-TAL pathway) **(Fig. 4)**. *Bacillus aryabhattai* SOS1, was isolated from a maple leaf (Konstanz, Germany) and grows aerobically on SQ to produce SL.[25] Using the draft genome sequence, comparative proteomics and 2D SDS-PAGE were used to identify upregulated proteins from growth on SQ leading to identification of the *sftATXGIFDE* gene cluster. Similar results were reported in a

contemporaneous study by Zhang and co-workers using *Bacillus megaterium* DSM1804 for equivalent genes named *sqvUABCD-slaB-SqvE*.[26]

The core steps of the sulfo-SFT pathway were reconstituted in vitro with recombinant enzymes.[25, 26] This pathway contains: a putative SQase (SftG, SqvC), SQ-SF isomerase (SftI, SqvD), and SF transaldolase (SftT, SqvA), which uses SF and GAP as substrate to produce F6P, which can enter glycolysis, or SF and erythrose-4-phosphate to produce sedoheptulose-7-phosphate, which can enter the PPP. The pathways also contain an NAD^+-dependent SLA dehydrogenase (SftD, SlaB) that oxidizes SLA to SL, and which is excreted into the growth media. Sulfo-SFT gene clusters are predominantly present within the classes Bacilli and Clostridia of the phylum Firmicutes, and representatives were also found in individual genomes of members of the phyla Fusobacteria, Chloroflexi, Actinobacteria, Spirochaetes, and Thermotogae.[25] Human gut Firmicutes *Enterococcus gilvus*, *Clostridium symbiosum* and *Eubacterium rectale* use the sulfo-SFT pathway for fermentative growth on SQ. However, while *E. gilvus* produced SL during fermentative growth on SQ, *C. symbiosum* and *E. rectale* produced DHPS, consistent with their genomes instead encoding an SLA reductase. One unresolved question is the function of a conserved gene encoding DUF4867 (domain of unknown function) in the sulfo-SFT clusters (SftX in *B. aryabhattai,* SqvB in *B. megaterium*).[25, 26]

Human fecal slurry-derived microcosms grown under anoxic conditions rapidly consume SQ to transiently produce DHPS and culminate in production of H_2S.[47] Levels of *E. rectale* and *Bilophila wadworthia* increased strongly, and correlated with disappearance of SQ and DHPS, respectively. DHPS production by *E. rectale* was associated with a sulfo-SFT pathway and DHPS consumption by *B. wadworthia* with a sulfite-lyase pathway through the HpsG-HpsH system involving HspG, a glycyl radical sulfite-lyase, and HspH as its cognate activator.[71] Sulfo-SFT pathway expression was two orders of magnitude higher than that of proteobacterial sulfo-EMP pathway, suggesting that *Enterobacteriaceae* are only minor contributors to SQ degradation in the human gut.

7. The sulfoglycolytic Sulfoquinovose Monooxygenase (sulfo-SMO) pathway

All pathways described above involve the degradation of SQ in two tiers, with production of the C3-fragments DHPS or SL through sulfoglycolysis. A distinct pathway termed the sulfo-SMO pathway, reported in *Agrobacterium tumefaciens* C58, achieves the complete degradation in a single organism **(Fig. 5)**.[28] *A. tumefaciens* grows on SQ but the only observable carbon metabolite produced in the growth media is bicarbonate, while growth correlates with release of sulfite, suggestive of complete degradation of SQ. Comparative proteomics revealed the upregulated proteins from growth on SQ derive from the gene cluster

smoABCDEFGHI (*Atu3277-3285*). This gene cluster encodes an ATP binding cassette (ABC) transporter system (SmoEGH) with an associated periplasmic SQ-binding protein (SmoF), a previously characterized SQase (SmoI) for hydrolysis of SQ glycosides,[48] an NAD(P)H-dependent flavin mononucleotide (FMN) reductase (SmoA), an FMN-dependent SQ monooxygenase (SmoC) that converts SQ to 6-oxo-glucose (6-OG), and a 6-OG reductase (SmoB) that converts 6-OG to glucose, which enters glycolysis. This sulfo-SMP pathway is distributed across α- and β-proteobacteria and is particularly prevalent among members of the *Agrobacterium* and *Rhizobium* genus within the *Rhizobiales* order. Some of these pathways lack the ABC cassette seen in *A. tumefaciens* C58.

7.1 Sulfoquinovose-binding protein (SmoF)

Solute binding proteins deliver solutes to ABC transporters to move small molecules across membranes.[72] These systems use a homodimeric nucleotide binding domain to enforce conformational changes on a heterodimeric transmembrane domain pair depending on the state of ATP binding, hydrolysis and ADP dissociation. This conformational change opens a channel to the periplasm upon ATP binding and to the cytosol upon ATP hydrolysis. The binding protein recognizes the substrate and docks to the transporter when the channel is in the periplasm-open state. Recombinant SmoF binds SQGro with K_d = 290 nM and does not bind the stereochemically-related monosaccharides D-glucose and D-glucuronic acid.[28] 3D X-ray structures reveal a globular fold featuring a pair of α/β domains **(Fig. 9a)**.[28] SQGro binds SmoF in a cleft between domains. Binding of SQGro causes a 31° hinge motion of its two domains resulting in complete enclosure of the ligand. Recognition is achieved through a sulfonate binding pocket, in which the sulfonate oxygens form hydrogen bonds to Thr220, Gyl166 and Ser43. An ordered water mediates an interaction to His13. The sugar C2-4 hydroxyls are also recognized through hydrogen-bonding interactions.

7.2 Flavin mononucleotide reductase (SmoA)

The genes encoding flavin reductase SmoA and monooxygenase SmoC are present within sulfo-SMO gene clusters in multiple organisms including *Agrobacterium* sp., *Rhizobium oryzae,* and *Aureimonas flava* suggestive that SmoA and SmoC comprise a two-component system that effects the desulfurization and oxidation of SQ.[28] Recombinantly-expressed *A. tumefaciens* SmoA is a pale yellow colour and heat-denaturing resulted in the release of FMN, identifying its preferred flavin. Kinetics with saturating FMN and NADH or NADPH revealed a preference for NADH. Thus, reduction of SmoA-bound FMN by hydride transfer from NAD(P)H supplies $FMNH_2$ to the monooxygenase partner SmoC. Similar results were obtained for the homologous enzyme from *R. oryzae*.

The SmoA-SmoC proteins share similarity with the *E. coli* SsuD-SsuE proteins of the alkylsulfonate sulfur-utilization gene cluster (*ssuEADCB*).[73] SsuD-SsuE comprise a two-component alkane sulfonate monooxygenase active on a range of linear alkane sulfonates, which is upregulated under sulfur starvation.[74] Crystal structures and biophysical analysis of solution states (apo vs FMN bound forms) of the smaller flavin reductase SsuE indicates FMN binding drives a tetramer-dimer equilibrium, possibly to enable association with the monooxygenase partner for transfer of reduced flavin. [75, 76] The mechanism and stoichiometry of association of SmoA-SmoC system is unknown.

7.3 Flavin mononucleotide-dependent sulfoquinovose monooxygenase (SmoC)

The oxidative cleavage of C—S bond of SQ is catalysed by the sulfoquinovose monooxygenase SmoC.[28] This reaction presumably involves a C4a-(hydro)peroxyflavin intermediate proposed for LadA[77] and SsuD[78] monooxygenases or an *N*5-peroxyflavin intermediate invoked for methanesulfonate monooxygenase MsuD and other monooxygenases.[79, 80] Combination of SmoC, FMN reductase SmoA, NADH or NADPH and FMN resulted in release of sulfite from SQ, with a preference for NADH. Activity appeared to be oxygen-dependent with conversion limited by the solubility of oxygen in aqueous solution. No activity was seen for SQGro or the sulfonate HEPES, thus revealing a clear preference for SQ. SmoC binds SQ with K_d = 3 μM, whereas no binding was detected for SQGro. The requirement for SmoA to reduce SmoC defines SmoC as a Category II two-component flavoprotein monooxygenase.[80]

A 3D X-ray structure of a homologue of SmoC from the equivalent pathway from the syntenic gene cluster in *Rhizobium orzyae* revealed an $\alpha_8\beta_8$ barrel with three insertion regions **(Fig. 9b)**. A low-resolution structure of SmoC from *A. tumefaciens* aligns well to this structure, showing high similarity. By overlay with the structurally-related methanesulfonate monooxygenase MsuD in complex with FMN a structural model of SmoC binding to FMN was generated. The isoalloxazine ring of $FMNH_2$ occupies a deep hydrophobic cleft present in both structures, and both structures contain an identical sulfonate binding pocket comprised of Trp206, Arg236, His238, His343. For alkylsulfonate monoxygenase SsuD, substrate binding induces conformational changes through dynamic loop movements involving salt-bridge interactions of conserved arginines and glutamates distal to the active site.[81, 82]

The mechanism of sulfonate monooxygenases remains enigmatic, with disagreement even on the identity of the oxidized form of FMN and the site of oxygen attack. Possible mechanisms are proposed based on studies with MsuDs.[79] Initially, formation of a C4a-peroxy or *N*5-peroxyflavin species is proposed to occur by reaction with oxygen on-enzyme **(Fig. 10a,c)**. One mechanism then proposes that the peroxide deprotonates C6, and the resulting

carbanion is oxidized to an α-hydroxysulfonate that undergoes elimination to produce sulfite and 6-OG **(Fig. 10a)**. An alternative mechanism suggests the terminal peroxide oxygen of either peroxyflavin species attacks the sulfonate sulfur, which then undergoes rearrangement and effect C-S bond cleavage and release of the 6-OG and sulfite **(Fig. 10b)**. Biophysical, kinetic and structural studies on the SmoA-SmoC two-component system are required to understand the kinetic mechanism, identifying protein-protein interactions and revealing residues that bind the sulfosugar and stabilize the C4a-(hydro)peroxyflavin or *N*5-peroxyflavin intermediate involved in the catalytic mechanism of SmoC.

7.4 NADPH-dependent 6-oxo-glucose reductase (SmoB)

The final step of the sulfo-SMO pathway is reduction of 6-OG to form glucose, catalysed by SmoB [28]. SmoB is an NADPH-dependent reductase from the aldose-ketose reductase (AKR) superfamily [83]. SmoB bound NADPH with K_d 2 μM and did not bind NADH. Activity for SmoB was demonstrated by incubation of SmoB with glucose, $NADP^+$ and $H_2{}^{18}O$, which resulted in incorporation of an ^{18}O-label at C6, as demonstrated by GC-MS (electron-ionisation).

Recombinant SmoB exists as a trimer in solution. The 3D structure of SmoB was obtained by X-ray crystallography and revealed a TIM-barrel fold **(Fig. 9c)**. A ternary complex of SmoB, NADPH and glucose revealed the cofactor is held in an extended, *anti*-conformation over the center of the barrel by use of a C-terminal binding site, with the nicotinamide moiety positioned 3 Å from C6 of the substrate. In the SmoB•$NADP^+$•glucose complex, glucose interacts with Lys120 (3 Å), His151 (2.8 Å) and Tyr76 (2.7 Å) within the conserved catalytic tetrad His/Tyr/Lys/Asp that is common to the AKR superfamily. The complex is consistent with a proposed mechanism involving direct hydride transfer from NADH to the aldehyde of 6-OG **(Fig. 10d)**.

8. Energy balance and carbon flux through sulfoglycolytic pathways.

Metabolism of sulfoquinovose may involve breakdown of SQDG, SQGro or SQ. The lack of lipases within sulfoglycolytic operons suggests that organisms have primarily evolved to breakdown SQGro or SQ after deacylation, which may be conducted by non-specific lipases providing an opportunistic fatty acid bounty for non-sulfoglycolytic and sulfoglycolytic organisms alike, with the metabolic yields as expected through the pathways of β-oxidation.[84] The four pathways of sulfoglycolysis involve SQ utilization with different outcomes in terms of carbon supply and production of reducing equivalents and ATP. The tier 1 pathways that utilize only half of the carbon of SQ and release C3 sulfonates are ascetic relative to the equivalent

glycolysis pathways.[25] The sulfo-EMP pathway of *E. coli* under aerobic conditions (including substrate-level phosphorylation of released DHAP) can be represented by the equation:

$$SQ + ADP + P_i \rightarrow \text{pyruvate} + DHPS + ATP$$

For the sulfo-ED pathways of *P. putida* SQ1 and *R. leguminosarum* SRDI565 the reaction is:

$$SQ + 2\ NAD^+ \rightarrow \text{pyruvate} + SL + 2\ NADH + 2\ H^+$$

While the sulfo-ED pathway of *P. putida* SQ1 appears to use passive transporters, the pathway in *R. leguminosarum* SRDI565 contains an ABC transporter, suggesting that a single molecule of ATP is required to import SQ.

The sulfo-SFT pathways of *B. aryabhattai* SOS1 and *B. megaterium* DSM1804 including oxidation of SLA to SL and substrate level phosphorylation of DHAP to produce pyruvate the reaction is:

$$SQ + 2\ NAD^+ + ADP + P_i \rightarrow \text{pyruvate} + SL + 2\ NADH + ATP + 2\ H^+$$

In the case of reduction of SLA to DHPS, such as in *E. rectale*, the sulfo-SFT process becomes:

$$SQ + ADP + P_i \rightarrow \text{pyruvate} + DHPS + ATP$$

The sulfo-SMO pathway of *A. tumefaciens* C58 is fundamentally different as the end-product of the reaction is glucose, which enters glycolysis, and can be represented as:

$$SQ + O_2 + 2\ NAD(P)H + 2\ H^+ \rightarrow \text{glucose} + SO_3^{2-} + 2\ NAD(P)^+ + H_2O$$

This analysis of the sulfo-SMO pathway excludes the investment of a single molecule of ATP during importation using the ABC transporter. For direct comparison with the tier 1 pathways, glycolytic breakdown of glucose to pyruvate must be included, as outlined in Table 1.

The metabolic logic of sulfoglycolysis changes for growth on SQGro, which is likely to be the more environmentally significant substrate. For the tier 1 sulfoglycolytic pathways, SQGro has 6 metabolizable carbons, and for the sulfo-SMP pathway 9 metabolizable carbons,

assuming that the released glycerol is converted to GAP by the action of glycerol kinase and glycerol-3-phosphate (G3P) dehydrogenase, which may act as an acceptor for SF transaldolase in the sulfo-SFT pathway or be transformed to pyruvate. Table 1 summarizes the metabolic yields using SQGro in each of these pathways. This analysis may lead to undue focus on the energetic and reducing equivalent yields of sulfoglycolysis and ignore the much greater yield from respiratory catabolism of pyruvate. Under aerobic conditions, pyruvate enters the citric acid cycle and undergoes oxidative phosphorylation leading to the net production of 32 ATP molecules. Sulfoglycolysis metabolic yields are likely to be more significant for growth under anaerobic conditions, where pyruvate may be subjected to fermentation (in the case of *E. coli* producing formate, acetate and succinate through mixed acid fermentation).[56]

Table 1. Metabolic yields for sulfoglycolytic pathways utilizing SQ or SQGro, upon conversion to pyruvate.

Pathway	Substrate	pyruvate	ATP	NAD(P)H	Product
EMP	glucose	2	2	2	
Sulfo-EMP	SQ	1	1	0	DHPS
	SQGro [a]	2	2	2	DHPS
ED	Glucose	2	1	2	
Sulfo-ED	SQ	1	0	2	SL
	SQGro [a]	2	1	4	SL
Sulfo-SFT	SQ	1	1	2	SL
	SQGro [a]	2	2	4	SL
Sulfo-SMO [b]	SQ + O_2	2	2	-1	Sulfite
	SQGro [a]	3	3	1	Sulfite

[a] Yields for SQGro assume conversion of released glycerol to pyruvate according to Gro + ADP + 2 NAD^+ + 2 H^+ → pyruvate + ATP + 2 NADH.

[b] For the sulfo-SMO pathway, the metabolic yield is calculated for glycolysis of the released glucose to pyruvate using the EMP pathway.

For bacteria grown exclusively under sulfoglycolytic conditions (ie growth on only SQ or SQGro), the carbon released from sulfoglycolysis needs to satisfy the energetic and metabolic requirements of growth. This requires partitioning of pyruvate into the citric acid cycle, fatty acid synthesis, cell wall biogenesis and the PPP. Unlike classical glycolysis pathways that produce G6P and F6P that can enter the PPP or support peptidoglycan synthesis, the sulfo-EMP and sulfo-ED pathways do not, and thus must generate these critical

molecules by gluconeogenesis. Evidence for gluconeogenesis for *R. leguminosarum* SRDI565 (which uses a sulfo-ED pathway) has been obtained by metabolomic analysis of cells grown on SQ as sole carbon source, which revealed the presence of G6P and F6P, but at levels much lower than cells grown on glucose, consistent with a more ascetic lifestyle under sulfoglycolysis.[44] Both the sulfo-SFT and sulfo-SMO pathways are fundamentally different in that they generate hexose-6-P through either transaldolase/isomerase reactions or direct glycolysis, respectively, and cells therefore do not require gluconeogenesis to supply the PPP or cell wall biogenesis. The sulfo-SMO pathway represents an assimilation pathway, that supplies glucose to the cell, whereas the sulfo-SFT pathway rewires the metabolic connections, injecting F6P directly into the glycolytic pathway.

9. Conclusions

The discovery of new pathways for sulfoglycolysis have invigorated the study of SQ and its metabolites and helped close an important gap in the biogeochemical sulfur cycle. The striking similarity of sulfo-EMP and sulfo-ED pathways with their namesake glycolytic pathways hints at an evolutionary relationship, possibly arising from gene duplication and neofunctionalization. A connection exists with the key enzyme of the sulfo-SFT pathway, sulfofructose transaldolase, and transaldolase within the PPP, which converts sedoheptulose-7-phosphate and GAP to erythrose-4-phosphate and F6P. The SQase (SmoI), SQGro binding protein (SmoF) and SmoE/G/H (ABC transporter) of the sulfo-SMO pathway share similarity with MalP (maltodextrin phosphorylase), MalE (maltose binding protein) and MalF/G/K (ABC transporter) encoded by the *mal* operon of *E. coli* that degrades maltodextin,[85] while SmoC (SQ monooxygenase) and SmoA (flavin reductase) of the SMO pathway are reminiscent of SsuD ($FMNH_2$-dependent alkylsulfonate monooxygenase) and SsuE (NADPH-dependent FMN reductase) encoded by the *ssu* operon of *E. coli* that degrades alkanesulfonates.[86] Like the *smo* gene cluster, the *ssu* operon also encodes an ABC transporter encoded by *ssuABC* that constitutes an uptake system for alkane sulfonates.[86]

The widespread distribution of sulfoglycolysis pathways in bacteria reflects the diverse niches in which photosynthetic tissues are found. These include within soils, waterways and the digestive tract of herbivores and omnivores, which appear to be replete with bacteria that catabolize only half of the SQ carbon through tier 1 pathways and require other bacteria to catabolise the released SL or DHPS through tier 2 biomineralization pathways.[1, 27, 87] The tier 1 sulfoglycolytic pathways are ‘generous’ in the sense that they support diverse microbial communities through supporting keystone degrading organisms that release C3-sulfonates. However, this support also operates in the reverse direction as biomineralized sulfur (as sulfate/sulfite) will support assimilatory sulfur metabolism by tier 1 organisms. The sulfo-SMO

pathway is distinct in that it allows the complete metabolism of SQ and release of sulfite that can enter sulfur assimilation pathways; thus this pathway may support a more selfish microbial lifestyle within competitive environments. The occurrence of the sulfo-SMO pathway within specialized mutualistic (both symbiotic and pathobiotic) bacteria that grow on or within plants suggest that it may support optimal nutrient acquisition even in the absence of other bacteria.

Sulfoglycolytic pathways are of fundamental interest as they inform our understanding of nutrient and energy acquisition from an unusual sugar with striking resemblance to glucose. However, the energy, reducing equivalents and carbon yields upon sulfoglycolysis versus glycolysis varies greatly. Understanding whether sulfoglycolysis leads to limitations in one or more of these outputs will be critical to understanding the metabolic consequences of utilizing sulfoglycolysis versus glycolysis. Our growing understanding of the structural basis for how nature recognizes the defining sulfonate group present in SQ and its metabolites is already enhancing the accuracy of bioinformatic methods and supporting new discoveries on the contribution of sulfoglycolysis to biomedically important processes.[47] Future work is needed to uncover the basis of sulfoglycolysis pathways in plants and algae to help understand SQ cycling within prototrophs.

Abbreviations:

ADP, adenosine diphosphate; ATP, adenosine triphosphate; DHAP, dihydroxyacetone phosphate; DHPS, 2,3-dihydroxypropane sulfonate; FBP, fructose bisphosphate; FMN, flavin mononucleotide; F6P, fructose-6-phosphate; G6P, glucose-6-phosphate; KDSG, 2-keto-3,6-deoxy-6-sulfogluconate; NAD(P)H, nicotinamide adenine dinucleotide (phosphate); 6-OG, 6-oxo-glucose; PPP, pentose phosphate pathway; PEP, phosphoenolpyruvate; PFK, phosphofructokinase; SF, sulfofructose; SFP, sulfofructose-1-phosphate; SG, sulfogluconate; SGL, 6-deoxy-6-sulfogluconolactone; SL, sulfolactate; SLA, sulfolactaldehyde; SQ, sulfoquinovose; SQDG, sulfoquinovosyl diacylglycerol; SQGro, sulfoquinovosyl glycerol; SQMG, sulfoquinovosyl monoglyceride; SR, sulforhamnose; sulfo-ED, sulfoglycolytic Entner-Doudoroff; sulfo-EMP, sulfoglycolytic Embden-Meyerof-Parnas; sulfo-SFT, sulfoglycolytic transaldolase; sulfo-SMO, sulfoglycolytic sulfoquinovose monooxygenase; UDP, uridine-5'-diphosphate.

Conflicts of interest

The authors declare no competing financial interests.

Acknowledgements

We thank the Australian Research Council (DP210100233, DP210100235; grants to SJW), The Leverhulme Trust (RPG-2017-190; support for A.J.D.S. and M.S.) and The Royal Society Chemistry for Ken Murray Research Professorship to G.J.D.

References

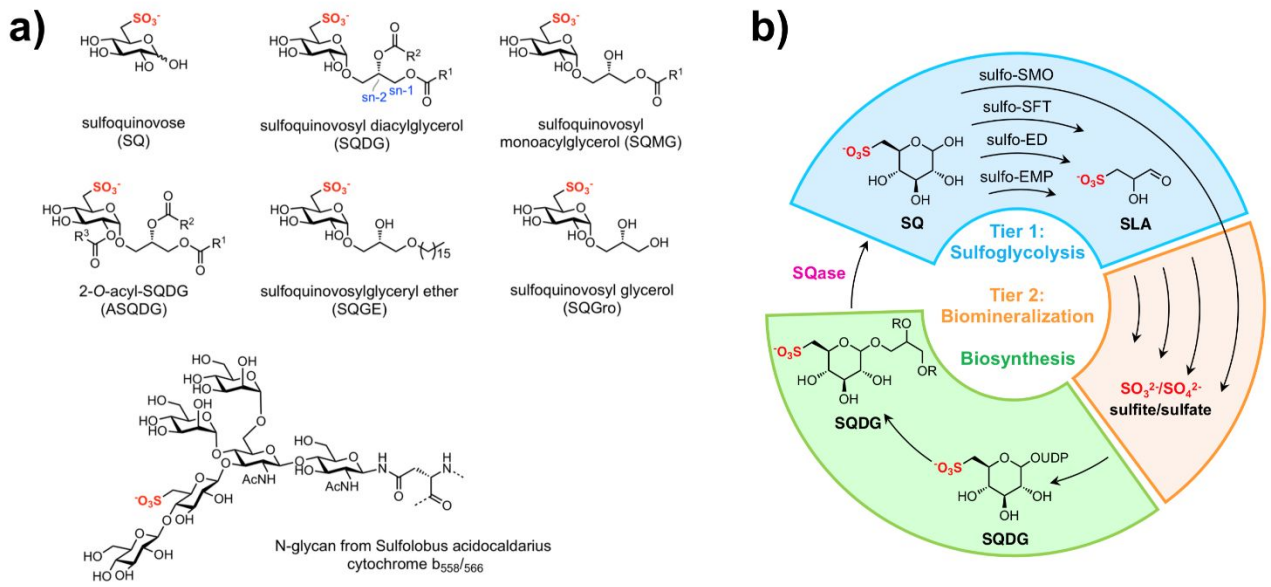


Fig. 1. Speciation and metabolism of sulfoquinovose (SQ). a) Structures of sulfoquinovose and naturally occurring glycoconjugates. b) Overview of the biosynthesis and catabolism of sulfoquinovosyl diglyceride (SQDG) and SQ.

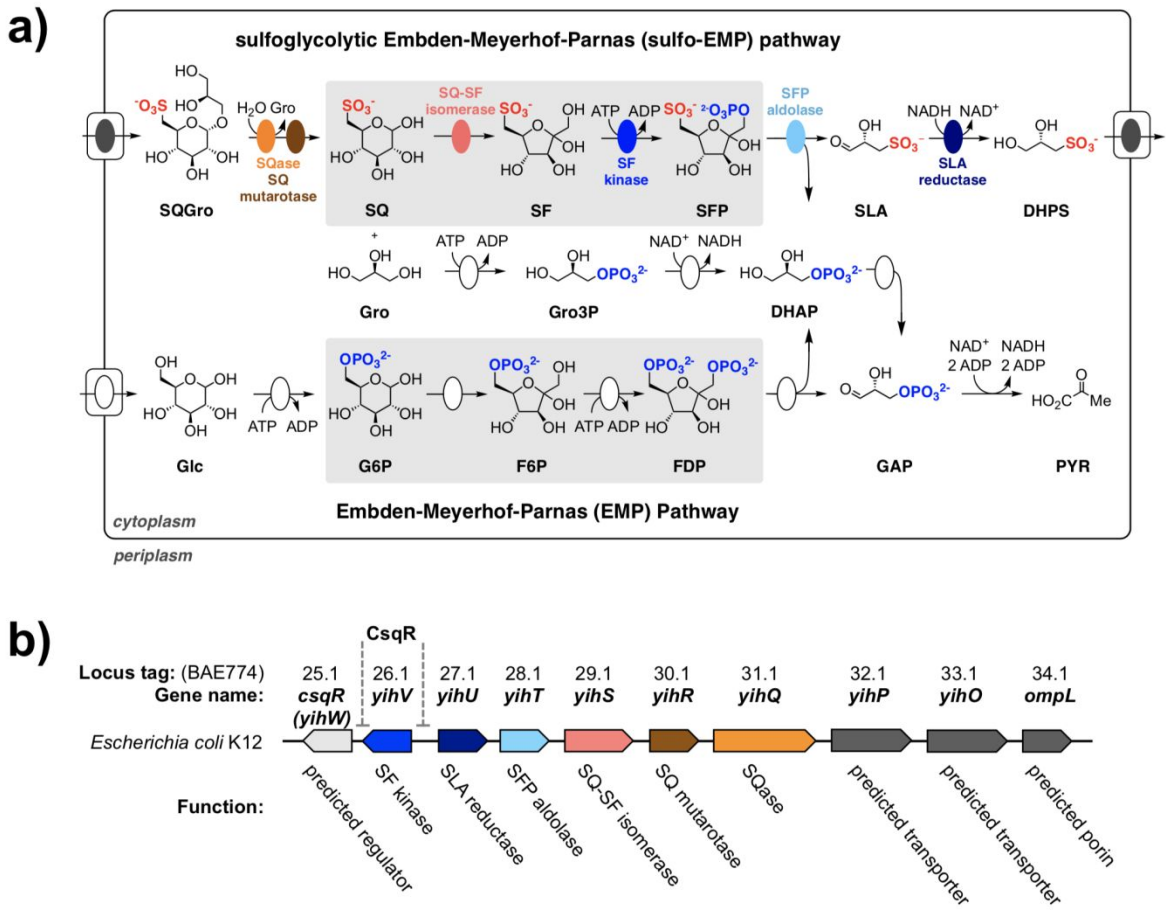


Fig. 2. The sulfoglycolytic Embden-Meyerhof-Parnas (sulfo-EMP) pathway. a) Comparison of the sulfo-EMP pathway for sulfoglycolysis of SQ by *E. coli* with the EMP pathway for glycolysis of glucose. b) sulfo-EMP gene cluster from *E. coli* K-12, showing intergenic binding sites for the transcription factor CsqR.

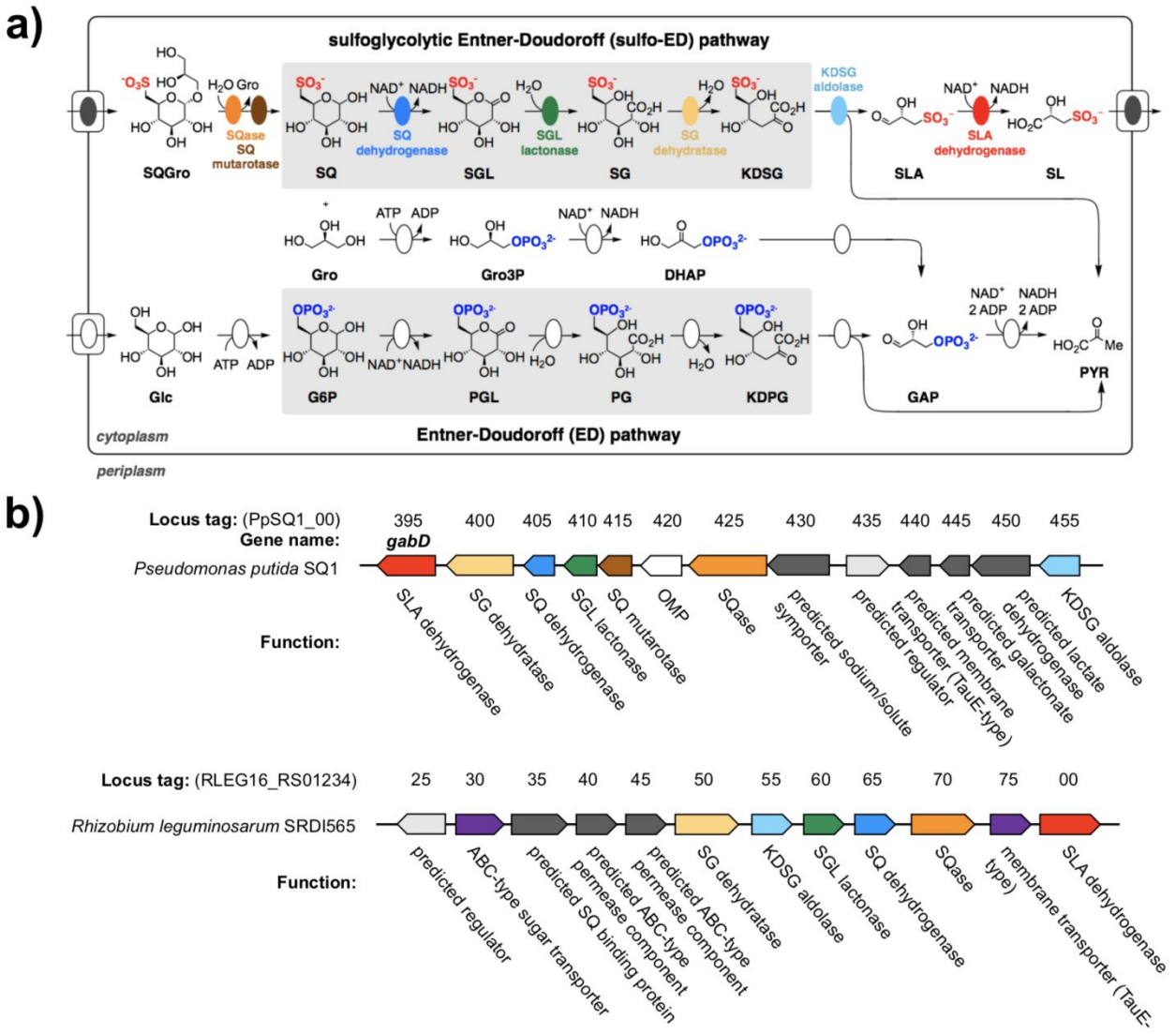


Fig. 3. The sulfoglycolytic Entner-Doudoroff (sulfo-ED) pathway. a) Comparison of the sulfo-ED pathway for sulfoglycolysis of SQ by *P. putida* SQ1 with the ED pathway for glycolysis of glucose. b) sulfo-ED gene clusters from *P. putida* SQ1 and *R. leguminosarum* SRDI565.

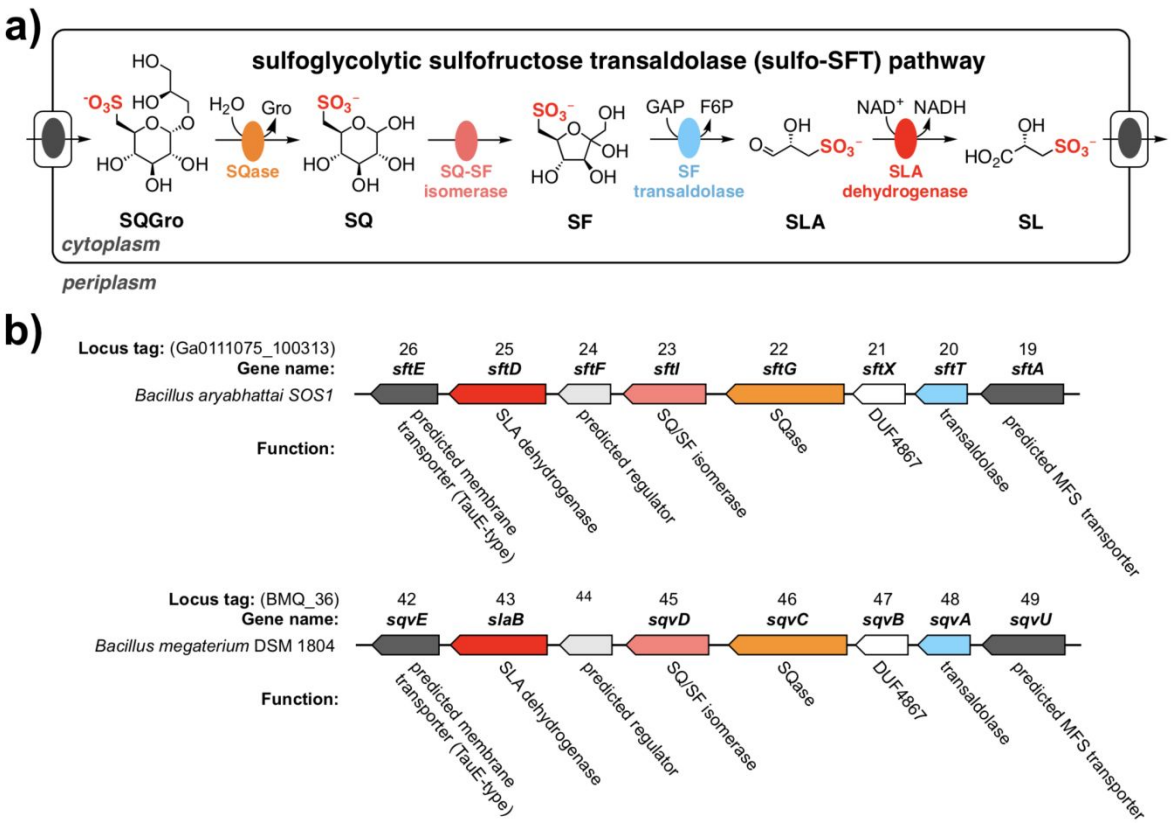


Fig. 4. The sulfoglycolytic sulfofructose transaldolase (sulfo-SFT) pathway. a) sulfo-SFT pathway for sulfoglycolysis of SQ. b) sulfo-SFT gene clusters from *B. aryabhattai* SOS1 and *B. megaterium* DSM 1804.

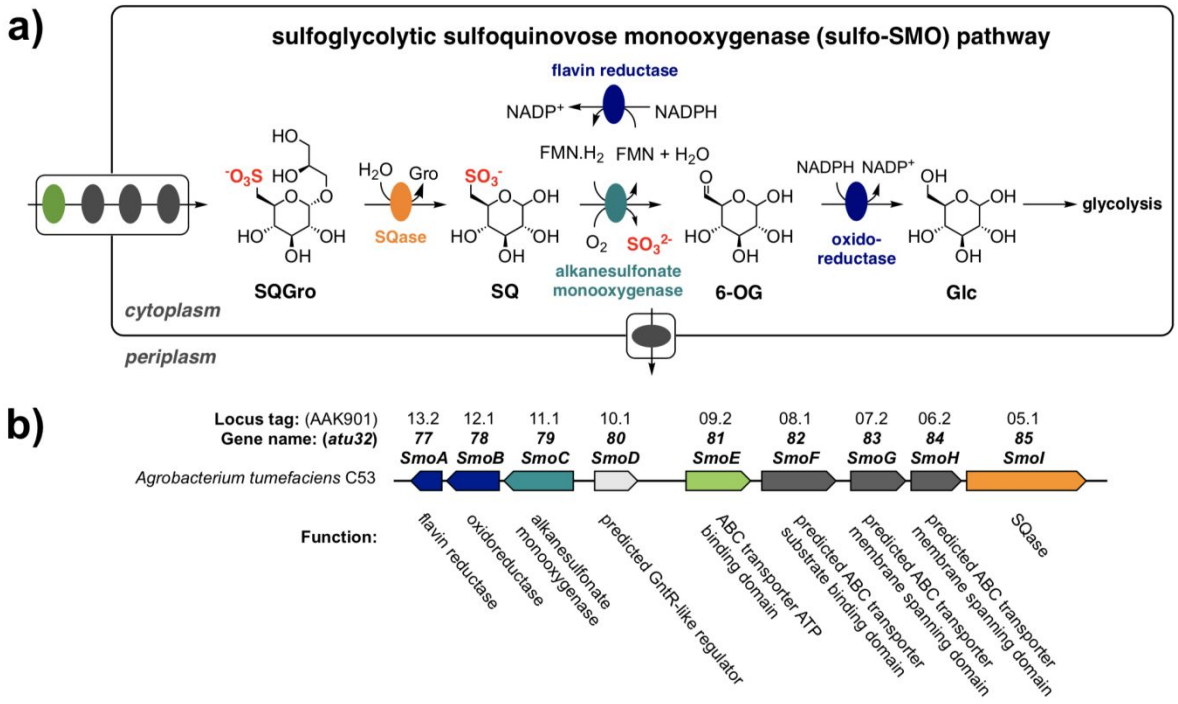


Fig. 5. The sulfoglycolytic sulfoquinovose monooxygenase (sulfo-SMO) pathway. a) sulfo-SMO pathway for sulfoglycolysis of SQ in *A. tumefaciens* C58. b) sulfo-SMO gene cluster.

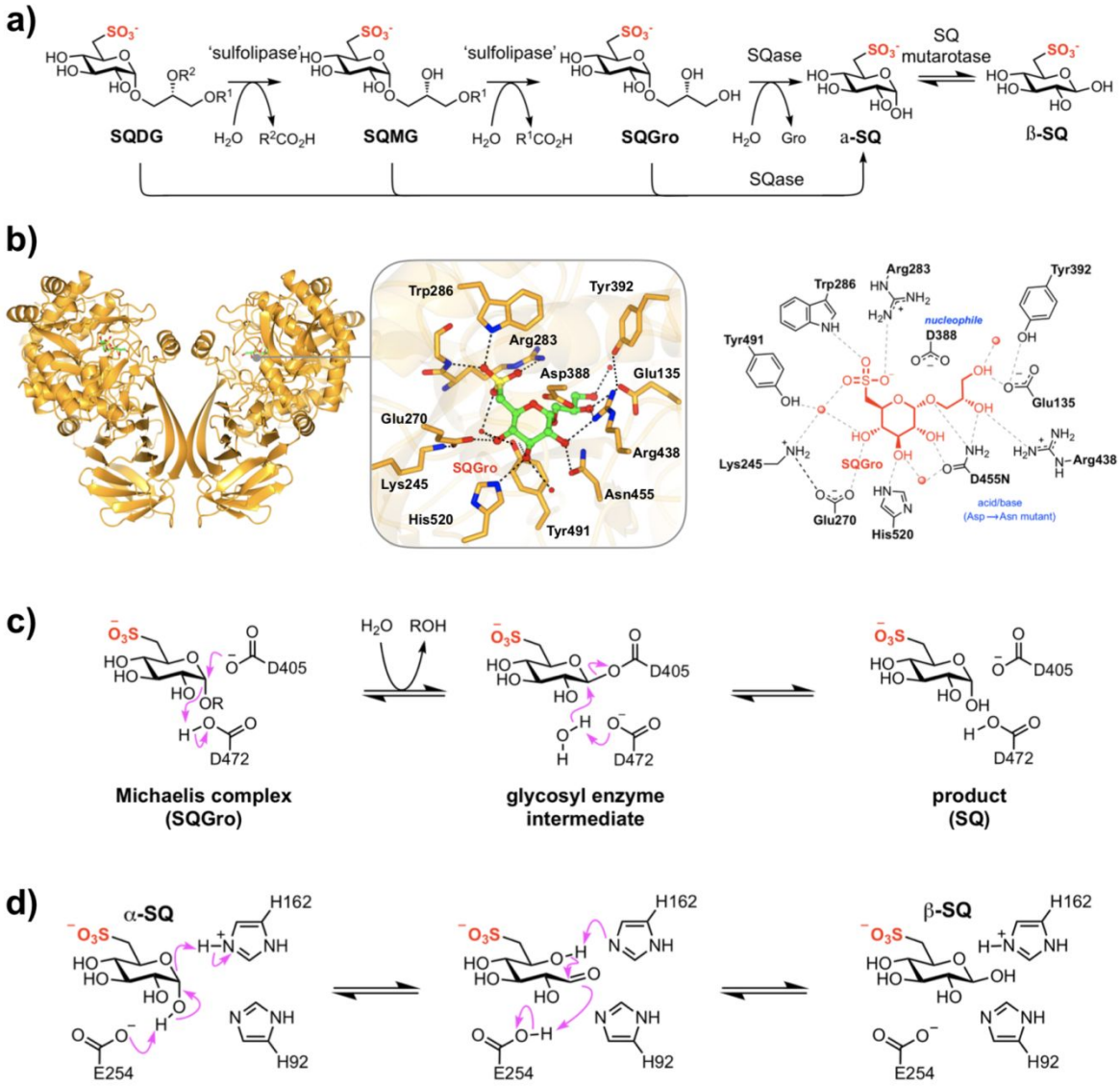


Fig. 6. SQDG catabolism preparatory phase. a) The initial steps before entering sulfoglycolysis for breakdown of SQDG involve delipidation, glycoside cleavage, and SQ mutarotation. b) Sulfoquinovosidases are a 'gateway' enzyme that liberates SQ for sulfoglycolysis. 3D structure of pseudo Michaelis complex of SQGro-bound to inactive acid/base mutant of SmoI sulfoquinovosidase (SQase) from *A. tumefaciens*. Close-up view of the active site (centre) and cartoon (right) showing conserved RWY sulfonate binding motif and catalytic residues of SQases. c) Catalytic retaining mechanism of SQases produce α-SQ; residue numbering for *A. tumefaciens* SmoI. d) Catalytic mechanism proposed for SQ mutarotase; residue numbering for *H. seropedicaea* SQM.

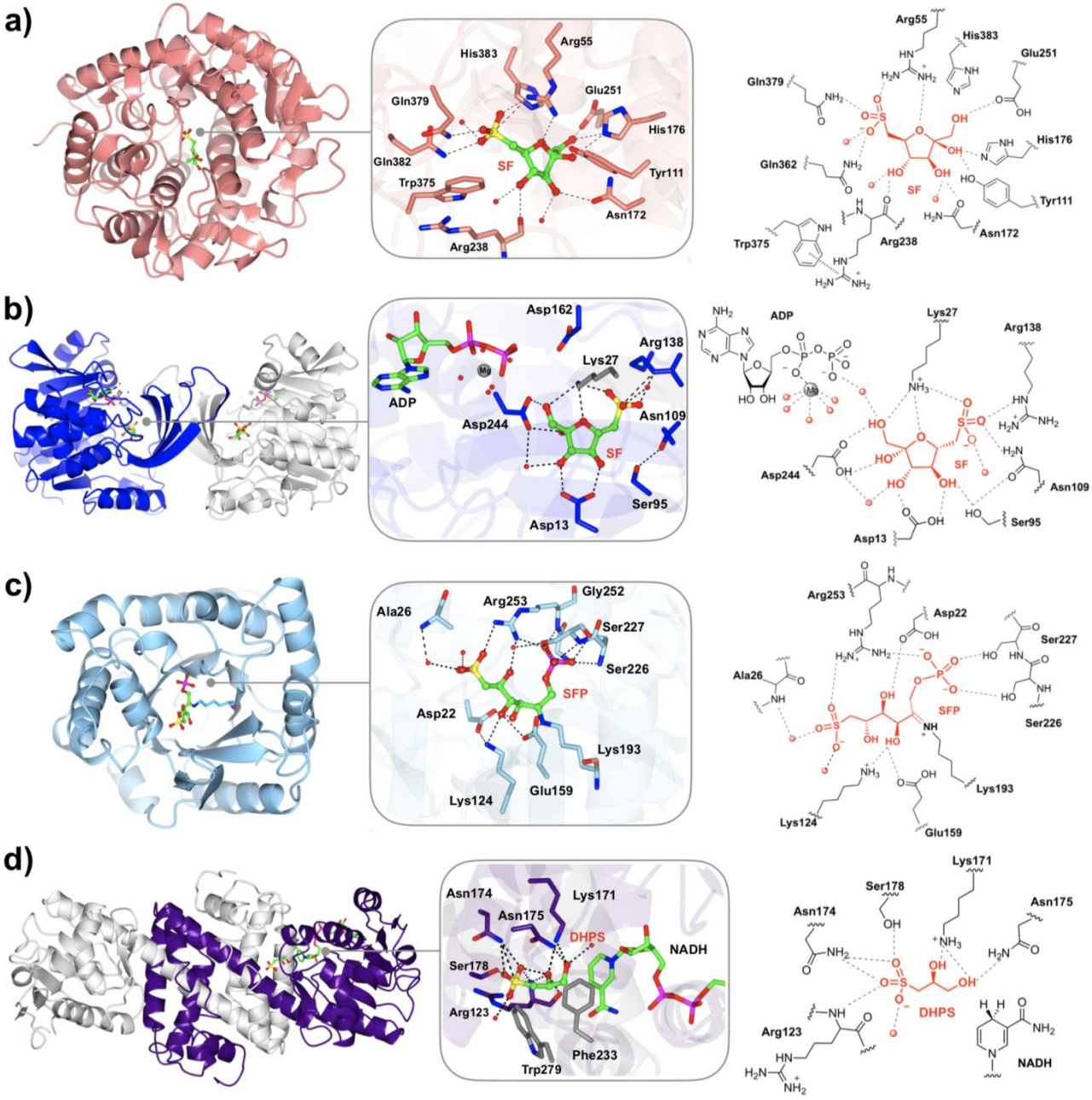


Fig. 7. 3D structures of four enzymes mediating SQ breakdown through the sulfo-EMP pathway. a) 3D structure of *S. enterica* SQ/SF isomerase (YihS) showing SF-bound active site view (left) and cartoon (right). b) Crystal structure of *E. coli* SF kinase (YihV) bound to ADP, Mg^{2+} and SF depicting the β-barrel dimer motif in blue and grey, active site (centre) and cartoon (right). c) 3D structure of *S. enterica* SFP aldolase (YihT) bound to SFP as a Schiff base with active site (centre) and cartoon (right). d) 3D structure of *E. coli* SLA reductase (YihU) in complex with NADH and DHPS showing active site view (centre) and cartoon (right). The intimate dimer pair with reciprocal domain-sharing is shown in purple and grey; the major solution state is a tetramer.

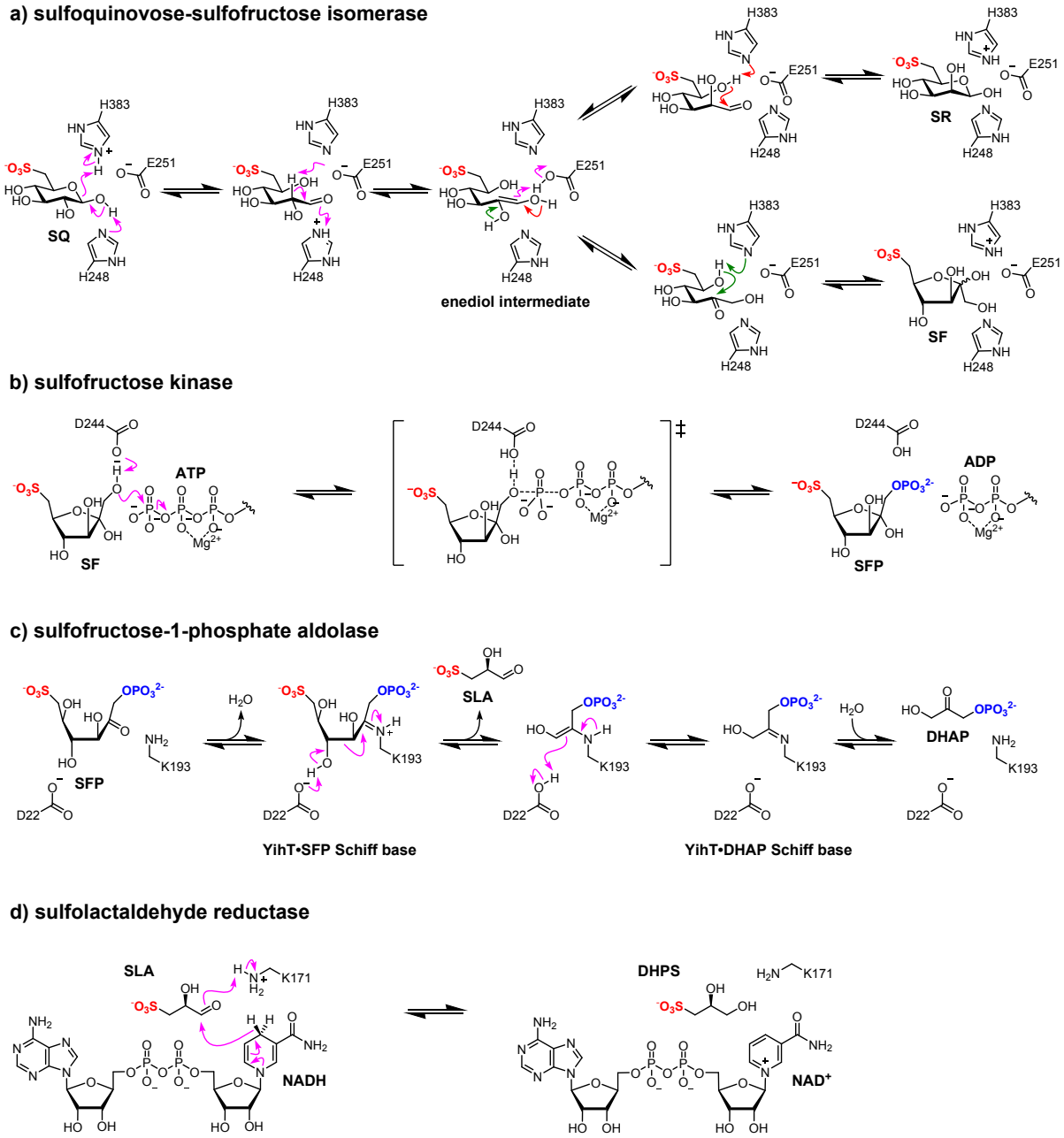


Fig. 8. Mechanism of enzymes in the sulfo-EMP pathway. a) Mechanism for interconversion of SQ, SR and SF catalyzed by SQ-SF isomerase. b) Mechanism for ATP-dependent SF kinase. c) Mechanism for SFP aldolase, a class I aldolase. d) Mechanism for NADH-dependent SLA reductase.

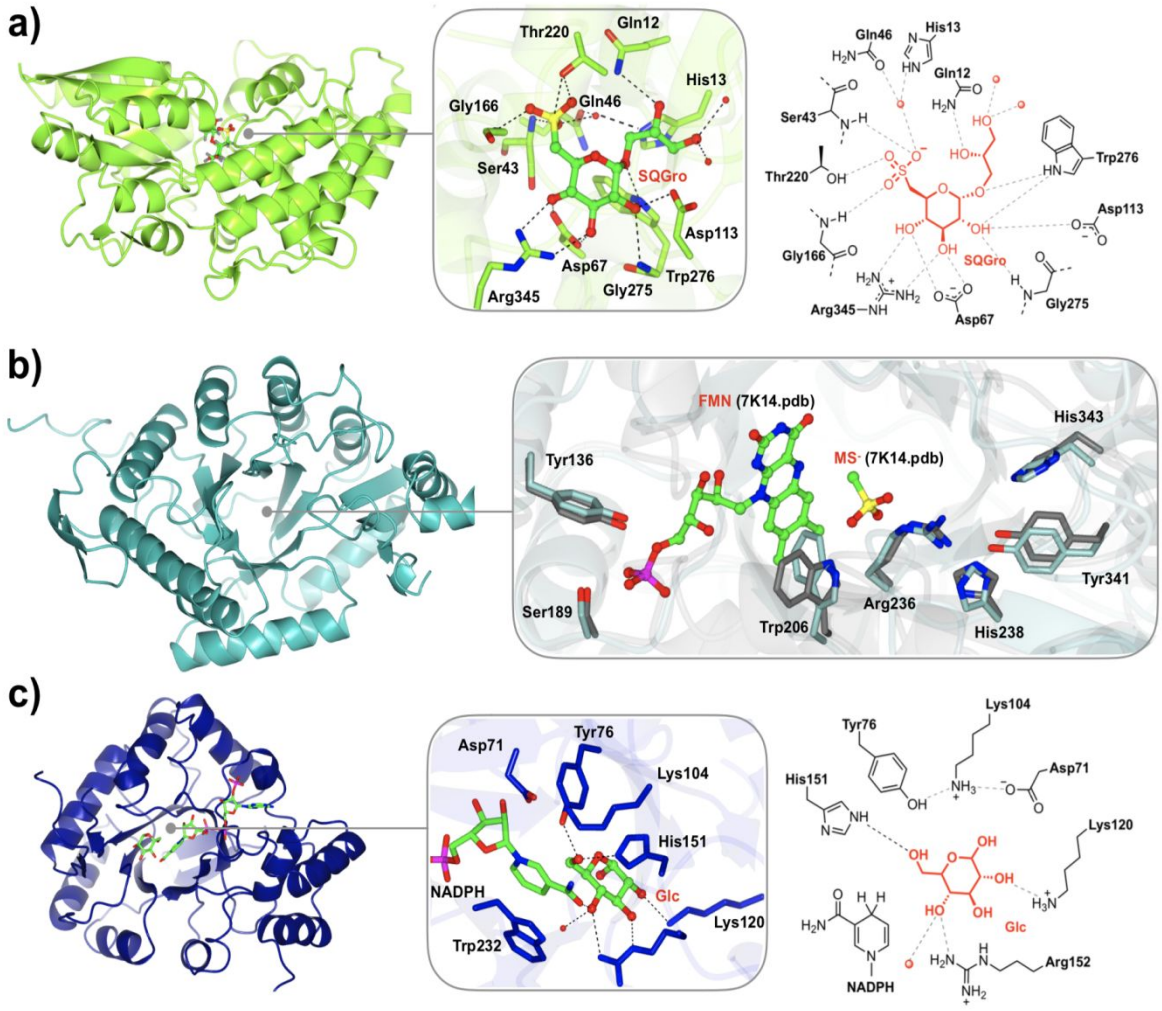


Fig. 9. 3D structures of proteins mediating SQ transport and breakdown through the sulfo-SMO pathway. a) X-ray structure of *A. tumefaciens* SmoF SQ binding protein (in green) showing SQGro-bound active site view (centre) and cartoon (right). b) Overlay of X-ray structure of apo *R. oryzae* SmoC SQ monooxygenase (in cyan) vs a ternary complex of a model alkanesulfonate monooxygenase (7K14, in grey) showing putative sulfosugar and FMN binding residues in an active site view. c) Crystal structure of *A. tumefaciens* SmoB 6-OG reductase (in dark blue) showing ternary complex with NADPH and glucose.

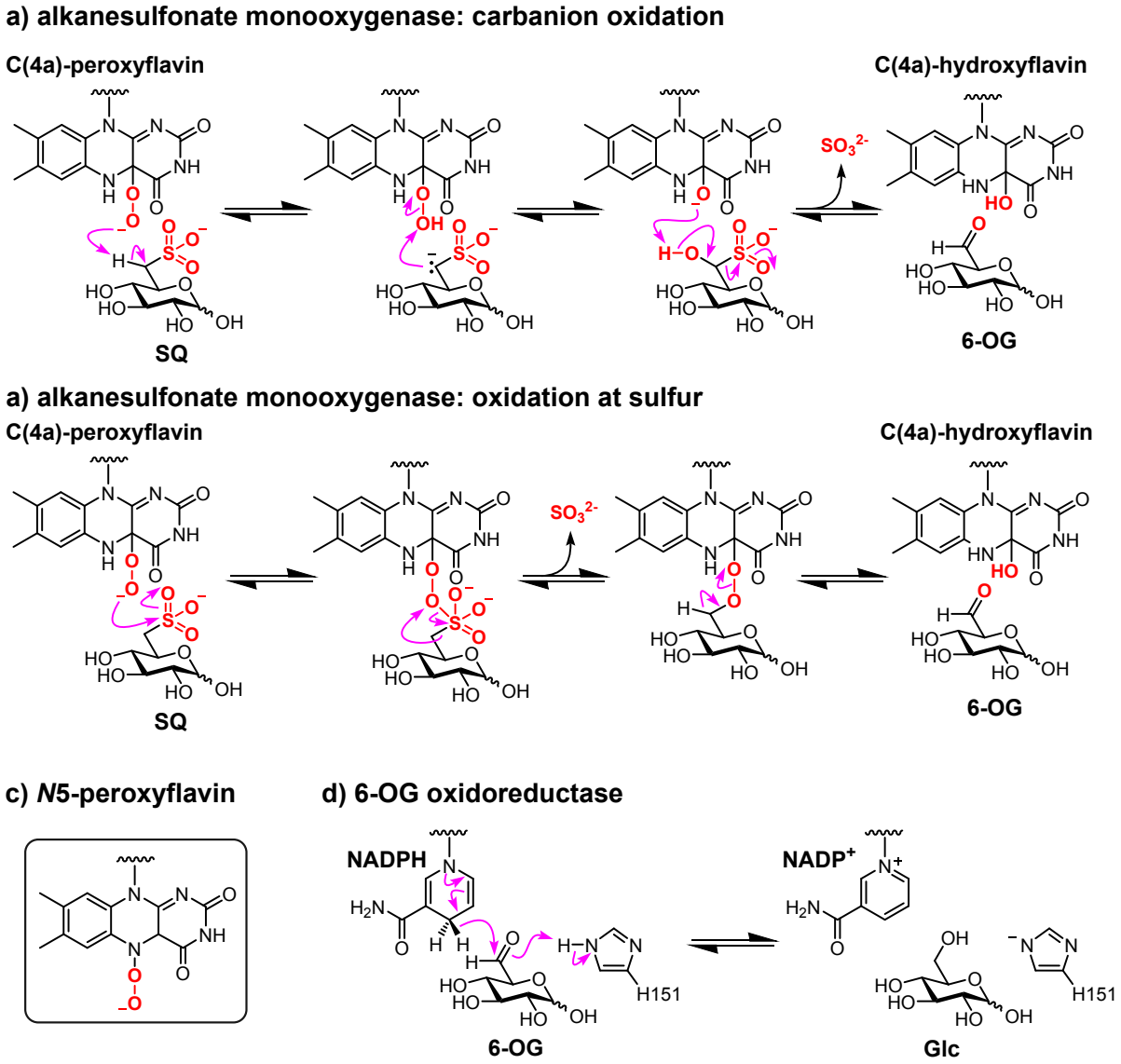


Fig. 10. Mechanism of enzymes in the sulfo-SMO pathway. Proposed mechanisms for conversion of SQ to 6-OG catalyzed by FMN-dependent SQ monooxygenase via a C4-peroxyflavin intermediate. a) Carbanion mechanism. b) Oxidation at sulfur mechanism. c) Structure of an alternative *N*5-peroxyflavin. d) Mechanism for NADPH-dependent 6-OG reductase.

1. E. D. Goddard-Borger and S. J. Williams, *Biochem. J.*, 2017, **474**, 827–849.
2. C. Benning, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 1998, **49**, 53-75.
3. A. A. Benson, H. Daniel and R. Wiser, *Proc. Natl. Acad. Sci. USA*, 1959, **45**, 1582-1587.
4. T. Yagi and A. A. Benson, *Biochim. Biophys. Acta*, 1962, **57**, 601-603.
5. Y. H. N. Fusetani, *Agric. Biol. Chem.*, 1975, **39**, 2021-2025.
6. N. Oku, A. Hasada, K. Kimura, H. Honoki, R. Katsuta, A. Yajima, T. Nukada, K. Ishigami and Y. Igarashi, *Chem. Asian J.*, 2021, **16**, 1493-1498.
7. W. R. Riekhof, M. E. Ruckle, T. A. Lydic, B. B. Sears and C. Benning, *Plant Physiol.*, 2003, **133**, 864-874.
8. U. Zähringer, H. Moll, T. Hettmann, Y. A. Knirel and G. Schäfer, *Eur. J. Biochem.*, 2000, **267**, 4144-4149.
9. K. Kobayashi, *J. Plant Res.*, 2016, **129**, 565-580.
10. N. Mizusawa and H. Wada, *Biochim. Biophys. Acta*, 2012, **1817**, 194-208.
11. Y. Nakajima, Y. Umena, R. Nagao, K. Endo, K. Kobayashi, F. Akita, M. Suga, H. Wada, T. Noguchi and J.-R. Shen, *J. Biol. Chem.*, 2018, **293**, 14786-14797.
12. G. Hölzl and P. Dörmann, *Annu. Rev. Plant Biol.*, 2019, **70**, 51-81.
13. M. Frentzen, *Curr. Opin. Plant Biol.*, 2004, **7**, 270-276.
14. B. Kalisch, P. Dörmann and G. Hölzl, *Subcell. Biochem.*, 2016, **86**, 51-83.
15. J. L. Harwood and R. G. Nicholls, *Biochem. Soc. Trans.*, 1979, **7**, 440-447.
16. E. Celik, M. Maczka, N. Bergen, T. Brinkhoff, S. Schulz and J. S. Dickschat, *Org. Biomol. Chem.*, 2017, **15**, 2919-2922.
17. J. S. Dickschat, P. Rabe and C. A. Citron, *Org. Biomol. Chem.*, 2015, **13**, 1954-1968.
18. K. Thume, B. Gebser, L. Chen, N. Meyer, D. J. Kieber and G. Pohnert, *Nature*, 2018, **563**, 412-415.
19. A. A. Benson and I. Shibuya, *Fed. Proc.*, 1961, **20**, 79.
20. A. B. Roy, A. J. Ellis, G. F. White and J. L. Harwood, *Biochem. Soc. Trans.*, 2000, **28**, 781-783.
21. A. B. Roy, M. J. Hewlins, A. J. Ellis, J. L. Harwood and G. F. White, *Appl. Environ. Microbiol.*, 2003, **69**, 6434-6441.
22. K. Denger, T. Huhn, K. Hollemeyer, D. Schleheck and A. M. Cook, *FEMS Microbiol. Lett.*, 2012, **328**, 39-45.
23. K. Denger, M. Weiss, A. K. Felux, A. Schneider, C. Mayer, D. Spiteller, T. Huhn, A. M. Cook and D. Schleheck, *Nature*, 2014, **507**, 114-117.
24. A. K. Felux, D. Spiteller, J. Klebensberger and D. Schleheck, *Proc. Natl. Acad. Sci. USA*, 2015, **112**, E4298-4305.
25. B. Frommeyer, A. W. Fiedler, S. R. Oehler, B. T. Hanson, A. Loy, P. Franchini, D. Spiteller and D. Schleheck, *iScience*, 2020, **23**, 101510.
26. Y. Liu, Y. Wei, Y. Zhou, E. L. Ang, H. Zhao and Y. Zhang, *Biochem. Biophys. Res. Commun.*, 2020, **533**, 1109-1114.
27. Y. Wei and Y. Zhang, *Annu. Rev. Biochem.*, 2021, **90**, 817-846.
28. M. Sharma, 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., *ChemRxiv*, 2021, 10.33774/chemrxiv-32021-33708j33701.
29. I. Shibuya, T. Yagi and A. A. Benson, *Plant Cell Physiol.*, 1963, 627-636.

30. R. F. Lee and A. A. Benson, *Biochem. Biophys. Acta*, 1972, **261**, 35-37.
31. S. Scholz, M. Serif, D. Schleheck, M. D. J. Sayer, A. M. Cook and F. C. Küpper, *Botanica Marina*, 2021.
32. T. C. Strickland and J. W. Fitzgerald, *Soil Biol. Biochem.*, 1983, **15**, 347-349.
33. H. L. Martelli and A. A. Benson, *Biochim. Biophys. Acta*, 1964, **93**, 169-171.
34. M. G. Wolfersberger and R. A. Pieringer, *J. Lipid Res.*, 1974, **15**, 1-10.
35. D. Dougal Burns, T. Galliard and J. L. Harwood, *Phytochemistry*, 1980, **19**, 2281-2285.
36. G. Hazlewood and R. M. C. Dawson, *Microbiology (Reading, England)*, 1979, **112**, 15-27.
37. I. Shibuya and E. Hase, *Plant Cell Physiol.*, 1965, **6**, 267-283.
38. S. D. Gupta and P. S. Sastry, *Arch. Biochem. Biophys.*, 1987, **259**, 510-519.
39. I. Shibuya and A. A. Benson, *Nature*, 1961, **192**, 1186-1187.
40. H. Daniel, M. Miyano, R. O. Mumma, T. Yagi, M. Lepage, I. Shibuya and A. A. Benson, *J. Am. Chem. Soc.*, 1961, **83**, 1765-1766.
41. V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho and B. Henrissat, *Nucleic acids research*, 2014, **42**, D490-495.
42. The CAZypedia Consortium, *Glycobiology*, 2018, **28**, 3-8.
43. G. Speciale, Y. Jin, G. J. Davies, S. J. Williams and E. D. Goddard-Borger, *Nat. Chem. Biol.*, 2016, **12**, 215-217.
44. J. Li, R. Epa, N. E. Scott, D. Skoneczny, M. Sharma, A. J. D. Snow, J. P. Lingford, E. D. Goddard-Borger, G. J. Davies, M. J. McConville and S. J. Williams, *Appl. Environ. Microbiol.*, 2020, **86**, e00750-00720.
45. S. R. Eddy, *Bioinformatics*, 1998, **14**, 755-763.
46. P. Rahfeld, J. F. Wardman, K. Mehr, D. Huff, C. Morgan-Lang, H. M. Chen, S. J. Hallam and S. G. Withers, *J. Biol. Chem.*, 2019, **294**, 16400-16415.
47. B. T. Hanson, K. Dimitri Kits, J. Löffler, A. G. Burrichter, A. Fiedler, K. Denger, B. Frommeyer, C. W. Herbold, T. Rattei, N. Karcher, N. Segata, D. Schleheck and A. Loy, *ISME J.*, 2021.
48. P. Abayakoon, Y. Jin, J. P. Lingford, M. Petricevic, A. John, E. Ryan, J. Wai-Ying Mui, D. E. V. Pires, D. B. Ascher, G. J. Davies, E. D. Goddard-Borger and S. J. Williams, *ACS Cent. Sci.*, 2018, **4**, 1266-1273.
49. Y. Zhang, J. W. Mui, T. Arumaperuma, J. P. Lingford, E. D. Goddard-Borger, J. M. White and S. J. Williams, *Org. Biomol. Chem.*, 2020, **18**, 675-686.
50. P. Abayakoon, J. P. Lingford, Y. Jin, C. Bengt, G. J. Davies, S. Yao, E. D. Goddard-Borger and S. J. Williams, *Biochem. J.*, 2018, **475**, 1371-1383.
51. J. M. Bailey, P. H. Fishman and P. G. Pentchev, *J. Biol. Chem.*, 1967, **242**, 4263-4269.
52. J. M. Bailey, P. G. Pentchev and P. H. Fishman, *Fed. Proc.*, 1967, **26**, 854.
53. J. B. Thoden, J. Kim, F. M. Raushel and H. M. Holden, *J. Biol. Chem.*, 2002, **277**, 45458-45465.
54. J. A. Beebe and P. A. Frey, *Biochemistry*, 1998, **37**, 14989-14997.
55. D. L. Gibson, A. P. White, S. D. Snyder, S. Martin, C. Heiss, P. Azadi, M. Surette and W. W. Kay, *J. Bacteriol.*, 2006, **188**, 7722-7730.
56. A. Burrichter, K. Denger, P. Franchini, T. Huhn, N. Müller, D. Spiteller and D. Schleheck, *Front. Microbiol.*, 2018, **9**.
57. T. Conway, *FEMS Microbiol. Rev.*, 1992, **9**, 1-27.

58. M. Sharma, P. Abayakoon, R. Epa, Y. Jin, J. P. Lingford, T. Shimada, M. Nakano, J. W. Y. Mui, A. Ishihama, E. D. Goddard-Borger, G. J. Davies and S. J. Williams, *ACS Cent. Sci.*, 2021, **7**, 476-487.
59. M. Sharma, P. Abayakoon, J. P. Lingford, R. Epa, A. John, Y. Jin, E. D. Goddard-Borger, G. J. Davies and S. J. Williams, *ACS Catalysis*, 2020, **10**, 2826-2836.
60. T. Itoh, B. Mikami, W. Hashimoto and K. Murata, *J. Mol. Biol.*, 2008, **377**, 1443-1459.
61. P. Abayakoon, R. Epa, M. Petricevic, C. Bengt, J. W. Y. Mui, P. L. van der Peet, Y. Zhang, J. P. Lingford, J. M. White, E. D. Goddard-Borger and S. J. Williams, *J. Org. Chem.*, 2019, **84**, 2901-2910.
62. V. Guixé and J. Babul, *J. Biol. Chem.*, 1985, **260**, 11001-11005.
63. R. L. Zheng and R. G. Kemp, *J. Biol. Chem.*, 1992, **267**, 23640-23645.
64. J. Park and R. S. Gupta, *Cell. Mol. Life Sci.*, 2008, **65**, 2875-2896.
65. T. Gefflaut, C. Blonski, J. Perie and M. Willson, *Prog. Biophys. Mol. Biol.*, 1995, **63**, 301-340.
66. N. Saito, M. Robert, H. Kochi, G. Matsuo, Y. Kakazu, T. Soga and M. Tomita, *J. Biol. Chem.*, 2009, **284**, 16442-16451.
67. R. K. Njau, C. A. Herndon and J. W. Hawes, *Chem-Biol Interact.*, 2001, **130-132**, 785-791.
68. J. Osipiuk, M. Zhou, S. Moy, F. Collart and A. Joachimiak, *J. Struct. Funct. Genomics*, 2009, **10**, 249-253.
69. T. Shimada, K. Yamamoto, M. Nakano, H. Watanabe, D. Schleheck and A. Ishihama, *Microbiology (Reading, England)*, 2019, **165**, 78-89.
70. A.-K. Felux, P. Franchini and D. Schleheck, *Stand. Genomic Sci.*, 2015, **10**, 42-42.
71. J. Liu, Y. Wei, L. Lin, L. Teng, J. Yin, Q. Lu, J. Chen, Y. Zheng, Y. Li, R. Xu, W. Zhai, Y. Liu, Y. Liu, P. Cao, E. L. Ang, H. Zhao, Z. Yuchi and Y. Zhang, *Proc. Natl. Acad. Sci. USA*, 2020, **117**, 15599-15608.
72. A. L. Davidson, E. Dassa, C. Orelle and J. Chen, *Microbiol. Mol. Biol. Rev.*, 2008, **72**, 317-364.
73. E. Eichhorn, J. R. van der Ploeg and T. Leisinger, *J. Biol. Chem.*, 1999, **274**, 26639-26646.
74. H. R. Ellis, *Bioorg. Chem.*, 2011, **39**, 178-184.
75. C. M. Driggers, P. V. Dayal, H. R. Ellis and P. A. Karplus, *Biochemistry*, 2014, **53**, 3509-3519.
76. K. Abdurachim and H. R. Ellis, *J. Bacteriol.*, 2006, **188**, 8153-8159.
77. L. Li, X. Liu, W. Yang, F. Xu, W. Wang, L. Feng, M. Bartlam, L. Wang and Z. Rao, *J. Mol. Biol.*, 2008, **376**, 453-465.
78. E. Eichhorn, C. A. Davey, D. F. Sargent, T. Leisinger and T. J. Richmond, *J. Mol. Biol.*, 2002, **324**, 457-468.
79. E. Romero, J. R. Gómez Castellanos, G. Gadda, M. W. Fraaije and A. Mattevi, *Chem. Rev.*, 2018, **118**, 1742-1769.
80. M. Toplak, A. Matthews and R. Teufel, *Arch. Biochem. Biophys.*, 2021, **698**, 108732.
81. A. Thakur, S. Somai, K. Yue, N. Ippolito, D. Pagan, J. Xiong, H. R. Ellis and O. Acevedo, *Biochemistry*, 2020, **59**, 3582-3593.
82. J. Xiong and H. R. Ellis, *Biochim. Biophys. Acta*, 2012, **1824**, 898-906.
83. J. M. Jez, M. J. Bennett, B. P. Schlegel, M. Lewis and T. M. Penning, *Biochem. J.*, 1997, **326**, 625-636.
84. J. K. Hiltunen and Y.-M. Qin, *Biochim. Biophys. Acta*, 2000, **1484**, 117-128.
85. R. Dippel and W. Boos, *J. Bacteriol.*, 2005, **187**, 8322.

86. J. R. van Der Ploeg, R. Iwanicka-Nowicka, T. Bykowski, M. M. Hryniewicz and T. Leisinger, *J. Biol. Chem.*, 1999, **274**, 29358-29365.
87. M. A. Moran and B. P. Durham, *Nature reviews. Microbiology*, 2019, **17**, 665-678.