Author/s:
Li, Z;Pickles, IB;Sharma, M;Melling, B;Pallasdies, L;Codée, JDC;Williams, SJ;Overkleeft, HS;Davies, GJ

Title:
Detection of Sulfoquinovosidase Activity in Cell Lysates Using Activity-Based Probes

Date:
2024-06-21

Citation:
Li, Z., Pickles, I. B., Sharma, M., Melling, B., Pallasdies, L., Codée, J. D. C., Williams, S. J., Overkleeft, H. S. & Davies, G. J. (2024). Detection of Sulfoquinovosidase Activity in Cell Lysates Using Activity-Based Probes. Angewandte Chemie International Edition, 63 (26), https://doi.org/10.1002/anie.202401358.

Persistent Link:
https://hdl.handle.net/11343/351050

How to cite: *Angew. Chem. Int. Ed.* **2024**, *63*, e202401358
doi.org/10.1002/anie.202401358

Activity-Based Probes

Detection of Sulfoquinovosidase Activity in Cell Lysates Using Activity-Based Probes

Zirui Li[+], Isabelle B. Pickles[+], Mahima Sharma[+], Benjamin Melling, Luise Pallasdies, Jeroen D. C. Codée, Spencer J. Williams, Herman S. Overkleeft,* and Gideon J. Davies**

Abstract: The sulfolipid sulfoquinovosyl diacylglycerol (SQDG), produced by plants, algae, and cyanobacteria, constitutes a major sulfur reserve in the biosphere. Microbial breakdown of SQDG is critical for the biological utilization of its sulfur. This commences through release of the parent sugar, sulfoquinovose (SQ), catalyzed by sulfoquinovosidases (SQases). These vanguard enzymes are encoded in gene clusters that code for diverse SQ catabolic pathways. To identify, visualize and isolate glycoside hydrolase CAZY-family 31 (GH31) SQases in complex biological environments, we introduce SQ cyclophellitol-aziridine activity-based probes (ABPs). These ABPs label the active site nucleophile of this enzyme family, consistent with specific recognition of the SQ cyclophellitol-aziridine in the active site, as evidenced in the 3D structure of *Bacillus megaterium* SQase. A fluorescent Cy5-probe enables visualization of SQases in crude cell lysates from bacteria harbouring different SQ breakdown pathways, whilst a biotin-probe enables SQase capture and identification by proteomics. The Cy5-probe facilitates monitoring of active SQase levels during different stages of bacterial growth which show great contrast to more traditional mRNA analysis obtained by RT-qPCR. Given the importance of SQases in global sulfur cycling and in human microbiota, these SQase ABPs provide a new tool with which to study SQase occurrence, activity and stability.

Introduction

Sulfoquinovose (6-deoxy-6-sulfo-D-glucose; SQ) is a sulfosugar that is found primarily as the headgroup of plant, algal and cyanobacterial sulfolipids (sulfoquinovosyl diacylglycerol; SQDG).[1] SQDG is primarily located within thylakoid membranes within photosynthetic tissues. In plants and algae it is found within chloroplasts, where it maintains membrane charge and supports photosynthetic proteins.[2] SQ is produced on a scale of 10 billion tonnes per annum, making it one of the most important organosulfur species within the sulfur cycle.[1c]

The breakdown of SQ occurs through the pathways of sulfoglycolysis and SQ sulfolysis, which occur mainly in bacteria (Figure 1a).[3] Prior to breakdown, SQ must be released from SQDG, or its deacylated form, sulfoquinovosyl glycerol (SQGro), by the action of highly selective glycoside hydrolases termed sulfoquinovosidases (SQases).[4] Sulfoglycolysis involves specialized catabolic pathways that cleave the carbon chain to access some of the carbon within SQ as C2- and C3-fragments such as pyruvate and triose phosphates that enter central carbon metabolism, and the release of short-chain C2- and C3-organosulfonates that are excreted.[5] SQ sulfolytic pathways cleave the C–S bond in SQ to release glucose, which can undergo glycolysis.[3a,c,6] Gene clusters in bacteria encoding sulfoglycolytic and SQ sulfolytic pathways typically encode SQases, as these are gateway enzymes required to liberate SQ from its glycosides.

The majority of SQases are classical glycosidases that belong to glycoside hydrolase family 31 (GH31) of the carbohydrate active enzyme classification Scheme (although a new NAD^+-dependent SQase family, GH188, has recently been reported),[7] and possess a conserved sulfonate binding motif that distinguishes them from other glycosidases within this family.[4,8] These SQases use a two-step mechanism that results in the stereochemically retaining hydrolysis of the glycosidic linkage, and which has been studied in detail for the *E. coli* SQase, YihQ (Figure 1b).[4b] The first step involves the nucleophilic attack, at the anomeric carbon of

[*] Z. Li,[+] Prof. Dr. J. D. C. Codée, Prof. Dr. H. S. Overkleeft
Department of Bio-organic Synthesis, Leiden Institute of Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands
E-mail: h.s.overkleeft@lic.leidenuniv.nl
Dr. I. B. Pickles,[+] Dr. M. Sharma,[+] B. Melling, Prof. Dr. G. J. Davies
York Structural Biology Laboratory, Department of Chemistry, University of York, York, YO10 5DD, UK
E-mail: gideon.davies@york.ac.uk
L. Pallasdies, Prof. Dr. S. J. Williams
School of Chemistry and Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Victoria, 3010, Australia
E-mail: sjwill@unimelb.edu.au
[+] these authors contributed equally

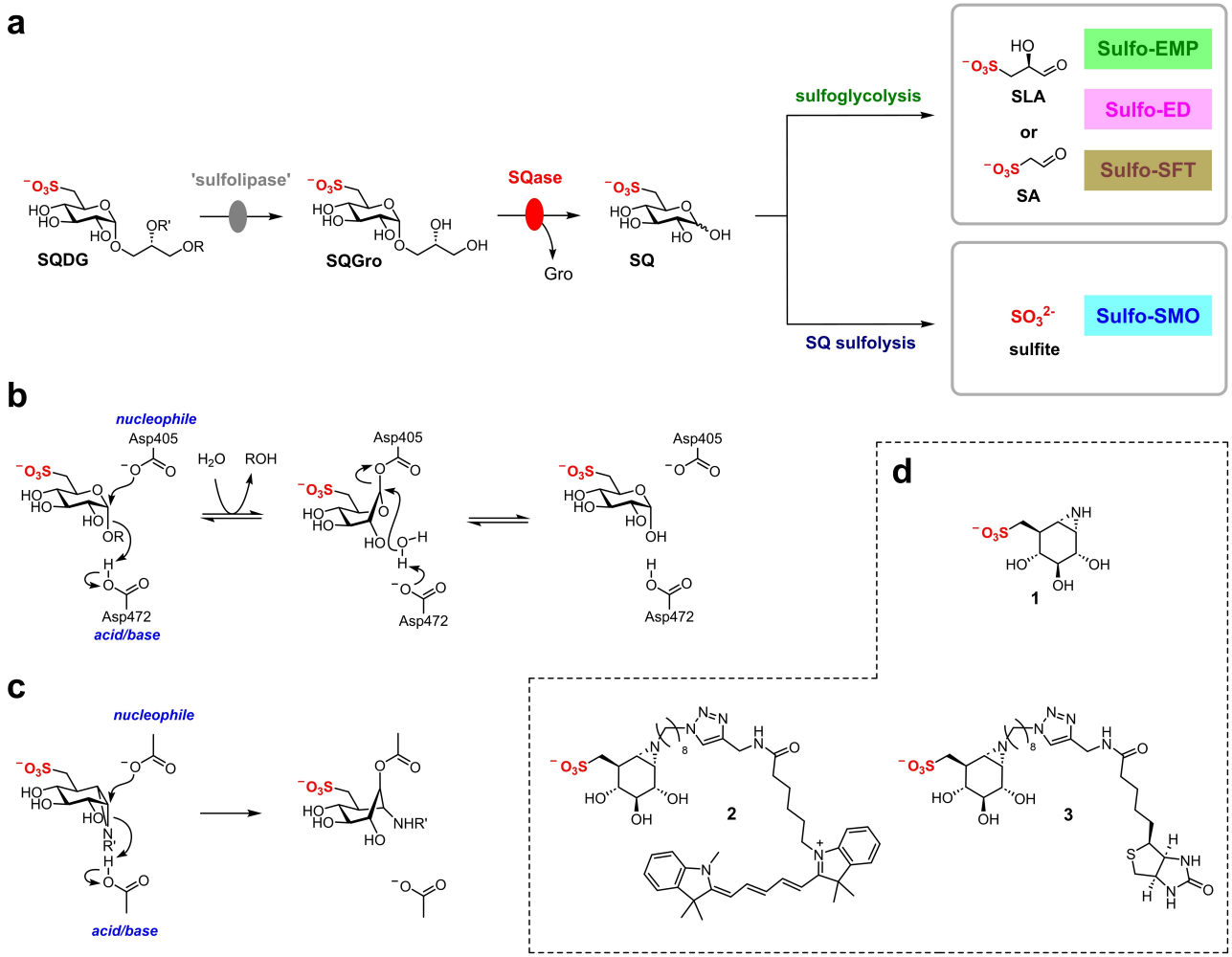


Figure 1. Design of activity-based probes targeting stereochemically retaining α-sulfoquinovosidases (SQases) of glycoside hydrolase family 31. (**a**) Summary showing the role of SQases as the vanguard enzymes for cleavage of SQ glycosides to allow SQ to enter sulfoglycolysis and SQ sulfolysis catabolism pathways. (**b**) Chemical mechanism of a SQase, with amino acid residues shown for *E. coli* SQase (YihQ). (**c**) Design of mechanism-based cyclophellitol aziridine based activity-based protein profiling probes. (**d**) Structures of parent SQ-cyclophellitol aziridine **1**, Cy5-probe **2** and biotin-probe **3**.

the substrate SQDG, by Asp405 forming a glycosyl enzyme acylal intermediate, which is hydrolysed in the second step with net retention of configuration. Asp472 acts as a general acid/base to assist the cleavage of the glycosidic bond and the nucleophilic attack by water. This mechanism is open to interrogation by activity-based probes (ABPs).

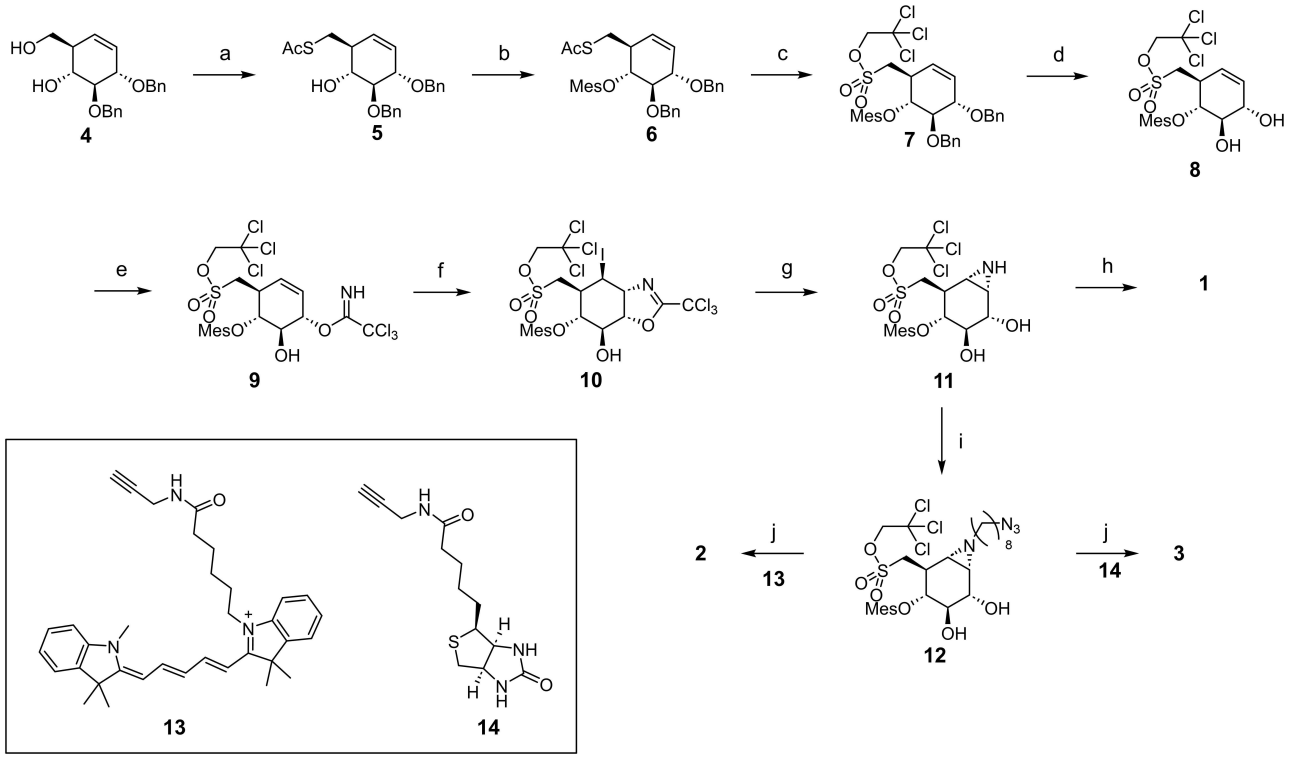
Activity-based protein profiling (ABPP) is a chemical biology technique that can be used to characterize enzyme activity in complex scenarios.[9] ABPP uses ABPs, which are reagents modified with a detectable tag that react in a known stoichiometry with an active enzyme to effect covalent and irreversible labelling (Figure 1c). The tag allows detection and quantification of the labelled enzyme in a manner proportional to its activity. ABPP is superior to alternative methods such as RNA transcript analysis and western blotting, which can detect gene expression or protein levels, respectively, but do not allow detection of enzyme activity. ABPP is complementary to chromo- and fluorogenic substrate hydrolysis in that, while the latter provides kinetic data that is proportional to enzyme catalysis, ABPP allows visualisation of active enzyme molecules and allows identification of different enzymes and isoforms in complex mixtures. ABPs have been developed for a wide range of glycosidases and have found use in a range of medical and biotechnological applications.[10] For example, acid beta-glucocerebrosidase (GBA) facilitates lysosomal recycling of glucocerebrosidase.[9c] Mutations in human GBA occur within the congenital disorder Gaucher's disease, and result in the formation of (partially) inactive protein and the accumulation of its substrate in the lysosome. Fluorescent mechanism-based inactivators of GBA have been developed based on the natural product cyclophellitol attached to a fluorescent tag, which allows visualization and quantitation of GBA activity in human subjects, and diagnosis of disease severity, as well as the effectiveness of enzyme replacement therapies.[21]

In this work we report the design of sulfonated cyclophellitol-aziridine ABPs that enable covalent labelling of

family GH31 SQases, the vanguard enzymes of sulfoglycolysis and SQ sulfolysis pathways (Figure 1d). Using X-ray crystallography, we show that these ABPs mimic the natural substrate of the enzyme and hijack the catalytic mechanism to covalently label the enzymatic nucleophile. Fluorescent and biotin tagged derivatives can be used to detect and quantify SQase activity of a range of bacterial GH31 SQases. The SQase ABPs allow detection of SQase activity in crude cell lysates and the tracking of SQase activity over the course of logarithmic and stationary phase bacterial growth on SQ.

Results and Discussion

Structure-based design of SQ probes was aided by the 3D structure of *At*SQase in complex with natural substrate SQGro (7OFX.pdb) as a model GH31 family SQases.[6] A surface loop containing residues 278–282 in *At*SQase was seen in two stable conformations (Figure S1). Visualisation of the active site from the surface and ligand interactions diagram show that the natural glyceryl aglycone is solvent exposed in both conformations. Based on the flexibility of the loop, the depth of the binding pocket and the interactions of the sub-pocket that accommodates the glyceryl group, we concluded that affinity probes based with $(CH_2)_8$ spacer would be suitable for this study.The synthesis of the aziridine **1**, Cy5-probe **2** and biotin-probe **3** began with the preparation of intermediate **4**, derived from D-xylose, through a nine-step process previously reported (Scheme 1).[11] The introduction of a thioacetate group at C6 of intermediate **4** was accomplished through a Mitsunobu reaction with AcSH and DIAD/PPh_3, affording thioacetate **5**. Protection of the remaining free hydroxyl group with a 2,4,6-trimethylbenzoyl (mesitoyl; Mes) group[12] produced the fully protected cyclohexene **6**. Oxidation of **6** with H_2O_2/K_2CO_3, followed by treatment with $SOCl_2$ and reaction of the intermediate sulfonyl chloride with trichloroethanol, gave the protected sulfonate **7**. Debenzylation of **7** with boron trichloride furnished diol **8**, which was converted to trichloroacetimidate **9** using trichloroacetonitrile in the presence of catalytic 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU). Treatment of trichloroacetimidate **9** with *N*-iodosuccinimide (NIS) caused cyclization to dihydrooxazoline **10**. Cleavage of the cyclic imidate of **10** with methanolic HCl, followed by intramolecular substitution of the resulting ammonium salt under mild basic conditions, provided aziridine **11**. The trichloroethyl and mesitoyl groups in **11** were removed by treatment with sodium hydroxide, affording aziridine **1**. *N*-Alkylation of **11** with 8-azidooctyl trifluoromethanesulfonate gave azidoalkyl aziridine **12**. Treatment of **12** with sodium hydroxide, followed by copper(I)-catalyzed azide/alkyne click (CuAAC) ligation



Scheme 1. Synthesis of the cyclophellitol aziridine inhibitor **1** and probes **2** and **3**. Reagents and conditions: (a) AcSH, PPh_3, DIAD, THF, 0 °C, 70%. (b) MesCl, DMAP, pyridine, 60 °C, 85%. (c) (i) H_2O_2, K_2CO_3, EtOH/dioxane. (ii) $SOCl_2$, DMF. (iii) 2,2,2-trichloroethanol, Et_3N, CH_2Cl_2, 41% over 3 steps. (d) BCl_3, −78 °C, CH_2Cl_2, quant. (e) CCl_3CN, DBU, CH_2Cl_2, 0 °C, 76%. (f) NIS, CH_2Cl_2, 71%. (g) (i) 1 M HCl in MeOH, MeOH. (ii) Amberlite IRA-67, 78% over 2 steps. (h) LiOH, MeOH, 60 °C, 89%. (i) 1-azido-8-trifluoromethylsulfonyloctane, DIPEA, CH_2Cl_2, 0 °C to rt, 60%. (j) (i) LiOH, MeOH, 60 °C. (ii) **13** or **14**, $CuSO_4$, sodium ascorbate, DMF, **2**: 24%; **3**: 33%.

with Cy5-alkyne **13**[13] or biotin-alkyne **14**,[14] yielded ABPs **2** and **3**, respectively.[11b]

Armed with untagged SQ-aziridine inhibitor **1** we investigated whether this compound could be used for covalent and irreversible active site modification of SQases. We initially selected five SQases from bacterial representatives of three sulfoglycolytic pathways: *Escherichia coli*, sulfoglycolytic Embden-Meyerhof-Parnas (sulfo-EMP) pathway; *Hafnia paralvei*, sulfo-EMP pathway; *Rhizobium leguminosarum*, sulfoglycolytic Entner-Doudoroff (sulfo-ED) pathway; and *Bacillus megaterium*, sulfoglycolytic sulfofructose transaldolase (sulfo-SFT) pathway; and one SQ sulfolysis pathway: *Agrobacterium tumefaciens*, sulfolytic SQ monooxygenase (sulfo-SMO) pathway.[4,5d,6,15] The genes encoding the SQases from these organisms were cloned, expressed in *E. coli* BL21(DE3) cells and the recombinant proteins purified to homogeneity. Hereafter the SQases will be identified with the genus and species as prefixes. To identify whether aziridine **1** can interact with these proteins, we studied their interaction using nanoDSF, which measures protein unfolding by changes in intrinsic tryptophan fluorescence. Incubation of **1** with *At*SQase led to an increase in melting temperature by 18 °C, consistent with a stabilizing interaction with the protein (Figure 2a). Intact mass spectrometry of *At*SQase before and after incubation with a 10-fold excess of **1** demonstrated complete conversion to products with an increase in mass of 241 Da, consistent with covalent modification with a single molecule of **1** (Figure 2b). Similar results were obtained for the four other SQases (Figure S2).

To determine the site of reaction of **1**, we studied the 3D structure of *Bm*SQase by X-ray crystallography. Soaking experiments led to trapping of the covalent intermediate in a crystal structure of *Bm*SQase·**1** that refined to a resolution of 2.45 Å in P21 space group. Diffraction data and refinement statistics are listed in Table S1. *Bm*SQase crystallized as a homodimer (Figure 2c), which is consistent with size-exclusion chromatography-multi angle laser light scattering in solution (Figure S3). Clear, contiguous density of the aziridine ring-opened entity derived from inhibitor **1** was seen protruding from $O\delta_2$ of Asp406 and this derivative could be modelled at occupancy of 1 in all four chains in the

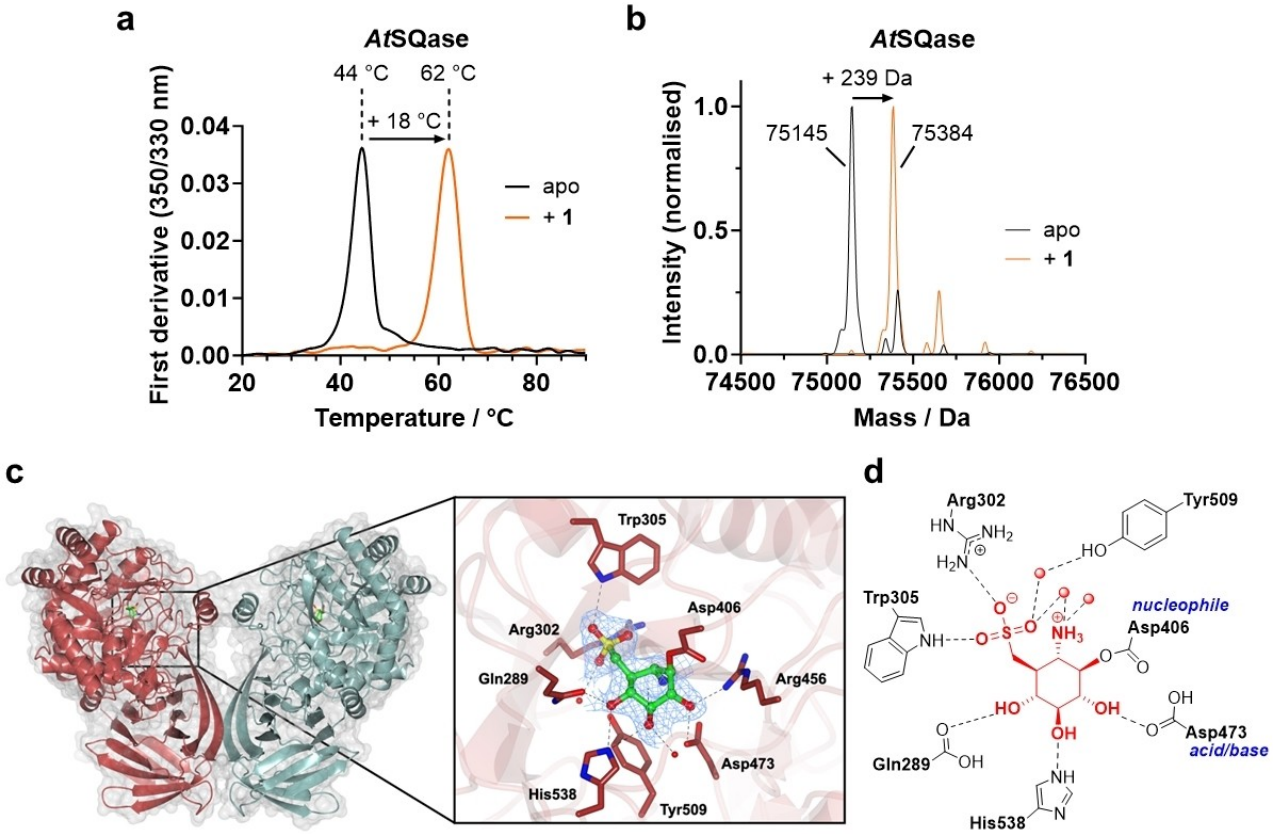


Figure 2. Cyclophellitol aziridine inhibitor **1** binds to SQase enzymes covalently. (**a**) Thermal stability curves of *At*SQase in its apo state (black line) and after incubation with 1 mM inhibitor **1** (orange), obtained using nano differential scanning fluorimetry. (**b**) Intact mass spectrum of *At*SQase in apo form (black) and bound covalently to one molecule of inhibitor **1** (orange) after treatment with 10-fold excess of compound. Expected mass calculated for apo *At*SQase: 75144, mass found: 75145 Da. Expected mass calculated for *At*SQase bound to one molecule of **1**: 75383, mass found: 75384 Da. (**c**) Overall fold of the *Bm*SQase dimer and active site interactions of bound covalent inhibitor **1**. Ribbon diagram of protein backbone is depicted in cyan and brick red, and bound ligand **1** and side chains of active site residues are shown in cylinder format. Electron density corresponds to the 2Fo − Fc map (in blue) at levels of 1 σ (0.298 e/Å^3). **(d)** Cartoon of ligand binding pocket of *Bm*SQase·**1** complex depicting hydrogen bonding interactions with active site and sulfonate binding residues. Water molecules are highlighted in red.

asymmetric unit. Consistent with the known conformational pathway[16] and previously trapped GH31 intermediates,[17] the covalently linked ligand is seen in the $^1S_3$ conformation at the active site and makes several hydrogen bonding interactions with the neighbouring residues: C2-OH binds Arg456 and Asp473 at 3.0 and 2.7 Å respectively; C3-OH interacts with His538 (3.0 Å) and a water molecule (2.6 Å); C4-OH with His538 (3.0 Å) and Gln289 (2.7 Å) and the secondary amine resultant from aziridine opening binds two water molecules (Figure 2d).

Inhibitor **1** makes identical interactions in the sulfonate pocket as seen in the Michaelis complex of nucleophile-mutant *At*SQase-D455 N with natural 2*R*′-SQGro substrate (Figure S4).[4a] The sulfonate oxyanion binds Arg302 (2.8 Å), the second sulfonate oxygen makes H-bonds to Trp305 (2.8 Å) and the third sulfonate binds Nε of Arg302 (3.0 Å) and two water molecules, one of which is in turn H-bonded to Tyr509, contributing to the high affinity of the inhibitor for GH31 SQases.

With nanoDSF, intact MS and X-ray structure analyses providing evidence of covalent labelling of the active-centre nucleophile by **1** we next explored whether Cy5-probe **2** could be used to label and visualize these SQases. After reaction optimization including variation of temperature, concentration, pH and time, conditions were established that enabled labelling of all five enzymes (1 μM protein, 5 μM **2**, pH 5.5, 37 °C, 2 h incubation; Figure 3a–c and Figure S5–8). Optimum labelling of *Ec*SQase, *Rl*SQase and *At*SQase occurred at pH 5.5, while *Hp*SQase and *Bm*SQase labelled best at pH 5 (Figure 3a). The pH dependence of labelling is consistent with the pH optimum for activity for *Ec*SQase (pH 6) but lower than that for *At*SQase (pH 8).[4] We analysed whether labelling efficacy was related to SQase sequence and the precise sequence within the mobile loop 278–282 (*E. coli* numbering) using sequence similarity network analysis and sequence motifs (Figure S9). Broadly, our results suggest that ABP labelling occurs across the major sequence-related clusters formed within the SQase subfamily, and is independent of variations in the mobile loop sequence. In order to study the specificity and detection limit of ABP **2**, *Ec*SQase was introduced into lysate from non-transformed *E. coli* BL21(DE3) cells in decreasing concentrations from 1 to 0.01 μM protein, and the mixture was labelled with 5 μM Cy5-probe **2** (Figure 3d). Only one protein was labelled in the complex lysate at a MW of ~70 kDa, corresponding to *Ec*SQase. Labelling of *Ec*SQase was observed to a detection limit of 50 nM.

We next used Cy5-probe **2** to study the expression of SQases by two organisms representing two pathways: *Escherichia coli* (sulfo-EMP) and *Pseudomonas putida* (sulfo-ED). Each bacterium was grown in minimal media, supplemented with 10 mM glucose or SQ as sole carbon source until an OD_{600} of 0.4 was reached, corresponding to approximately half stationary-phase density. Cells were sedimented, washed and resuspended, lysed at the optimum pH for labelling as demonstrated on pure protein, and clarified cell lysates were treated with 1 μM Cy5-probe **2** (Figure 4a). SDS-PAGE analysis and fluorimetry revealed a single labelled protein band only in the SQ-grown samples, suggesting selectivity for the SQases over all other soluble proteins, and showing that growth under sulfoglycolytic conditions leads to detectable SQase levels. To confirm the identity of the labelled band, biotin-tagged ABP **3** was used in a pull-down experiment. *P. putida* cell lysate was labelled with 10 μM **3**, excess probe removed by acetone precipitation and labelled proteins pulled down using magnetic streptavidin beads. On-bead digest using trypsin was followed by LC–MS/MS analysis of the resultant peptides. Peptides from *Pp*SQase were only observed when the probe was present and not in the DMSO control (Figure 4b and Table S2). The probe labelled no additional proteins that were not similarly observed in the DMSO control, showing the probe is specific for *Pp*SQase.

To probe activity levels of *Pp*SQase at different time points along the growth curve of *P. putida* grown on SQ, cells were harvested at various A_{600} points on their growth curves and their lysates labelled with ABP **2** (Figure 4c,d). Quantification of labelling by fluorimetry revealed labelled protein levels increased between time points 1 and 4 (early to late log/early stationary phase) and dropped sharply at late stationary phase (time point 5) (Figure 4e). Gene expression levels of *ppsqase* at the same five time points was monitored using quantitative polymerase chain reaction (qPCR) analysis of reverse transcribed RNA extracted from cells at different time points (Figure 4e and Figure S10,11).[5b] Levels of *ppsqase* were greatest in early logarithmic phase (time point 1) and decreased over the 20 hours of growth until no mRNA was observed at late stationary phase (time point 5). No *ppsqase* expression was observed in the glucose control.

Discussion and Conclusions

Kinetic studies of *Ec*SQase and *At*SQase reveal a high selectivity for PNPSQ versus PNPGlc within the SQases of GH family 31.[4a] Consequently, we developed SQase-directed probes for GH31 SQases by modifying the natural product cyclophellitol into the corresponding sulfonate. Additionally, we varied the epoxide to an aziridine, a modification we have previously demonstrated to allow attachment of a fluorescent or biotin tag through modification of the aziridine nitrogen.[10a,18] Prior methods for the synthesis of SQ analogues typically involved introducing the sulfonate either directly by substitution of a leaving group with sulfite, or in two steps by oxidation of a divalent sulfur (for instance, sulfide or thioacetate).[1b,12,19] However, we found these approaches to be incompatible with the alkene present in cyclophellitol or other functionalities en route to the aziridine. Instead, by introducing the sulfonate group early in the sequence, and subsequently protecting it as a chloroethyl ester, we were able to achieve the necessary functional group transformations to obtain the aziridine.[11] The trichloroethyl sulfonate of intermediate **11** was cleaved, along with other protecting groups to give **1**.[11b] Furthermore, alkylation of **11** gave compound **12**. Copper(I) catalyzed azide-alkyne cycloaddition of **12** with Cy5 and

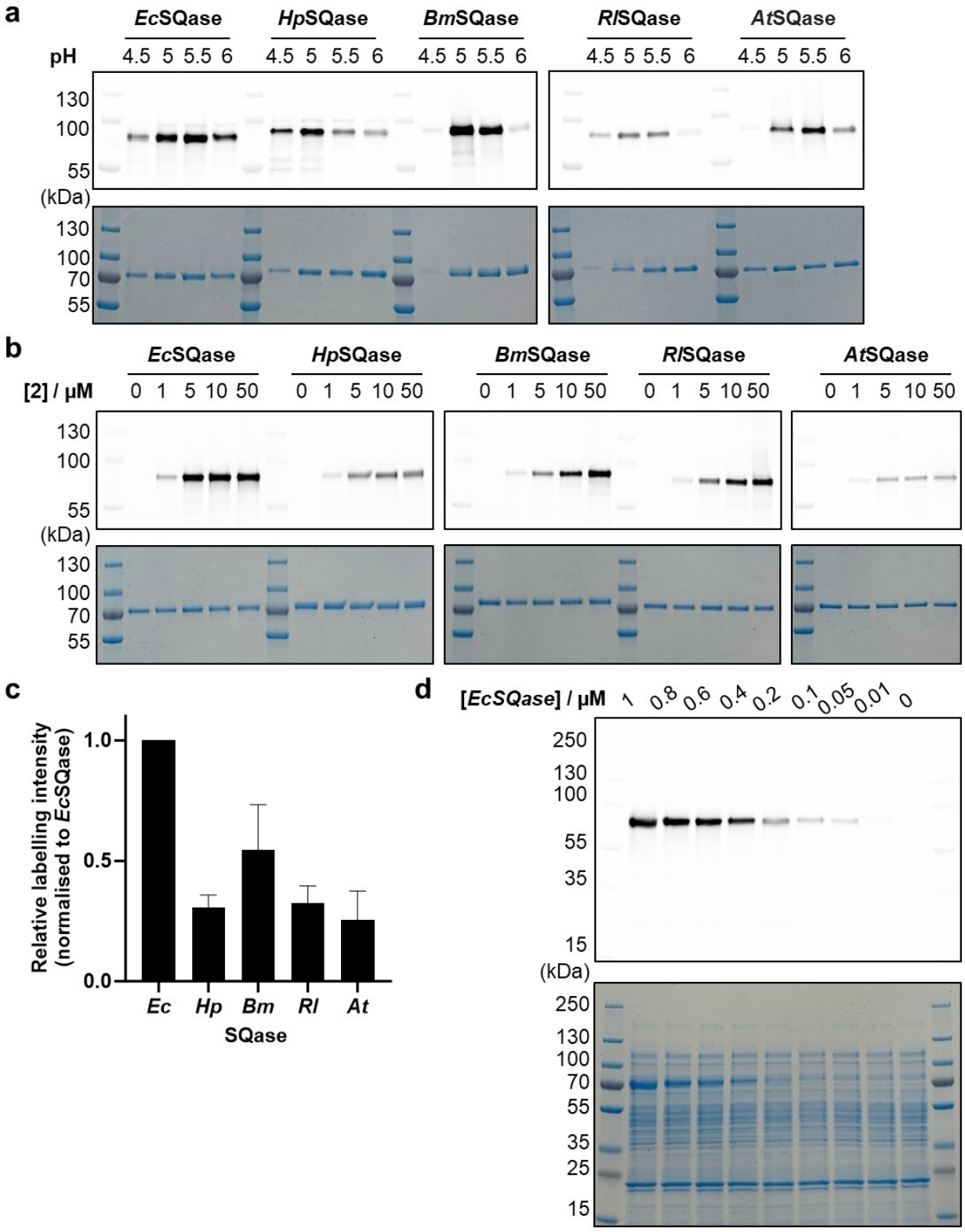


Figure 3. ABP **2** labels pure SQase enzymes selectively in a pH and concentration dependent manner. (**a**) 1 µM SQase enzyme from different bacteria labelled with 5 µM ABP **2** at different pH points. (**b**) 1 µM enzyme labelled in a concentration dependent manner from 1–50 µM ABP **2**. (**c**) Relative labelling intensities of *Hp*SQase, *Bm*SQase, *Rl*SQase and *At*SQase normalised to the signal for *Ec*SQase. The band intensity on the Cy5 fluorescent gel scan corresponding to each enzyme labelled under optimal conditions (pH 5.5, 5 µM **2**, 1 µM protein, 37 °C, 2 hours) was quantified and the average of four repeats plotted ± standard error mean. (**d**) Doping of purified *Ec*SQase into *E. coli* lysate, labelled with ABP **2**. In (a, b + d), F and C denote the fluorescent gel scan and the Coomassie stained gel, respectively. Full length gels are shown in Figure S8 in the Supporting Information.

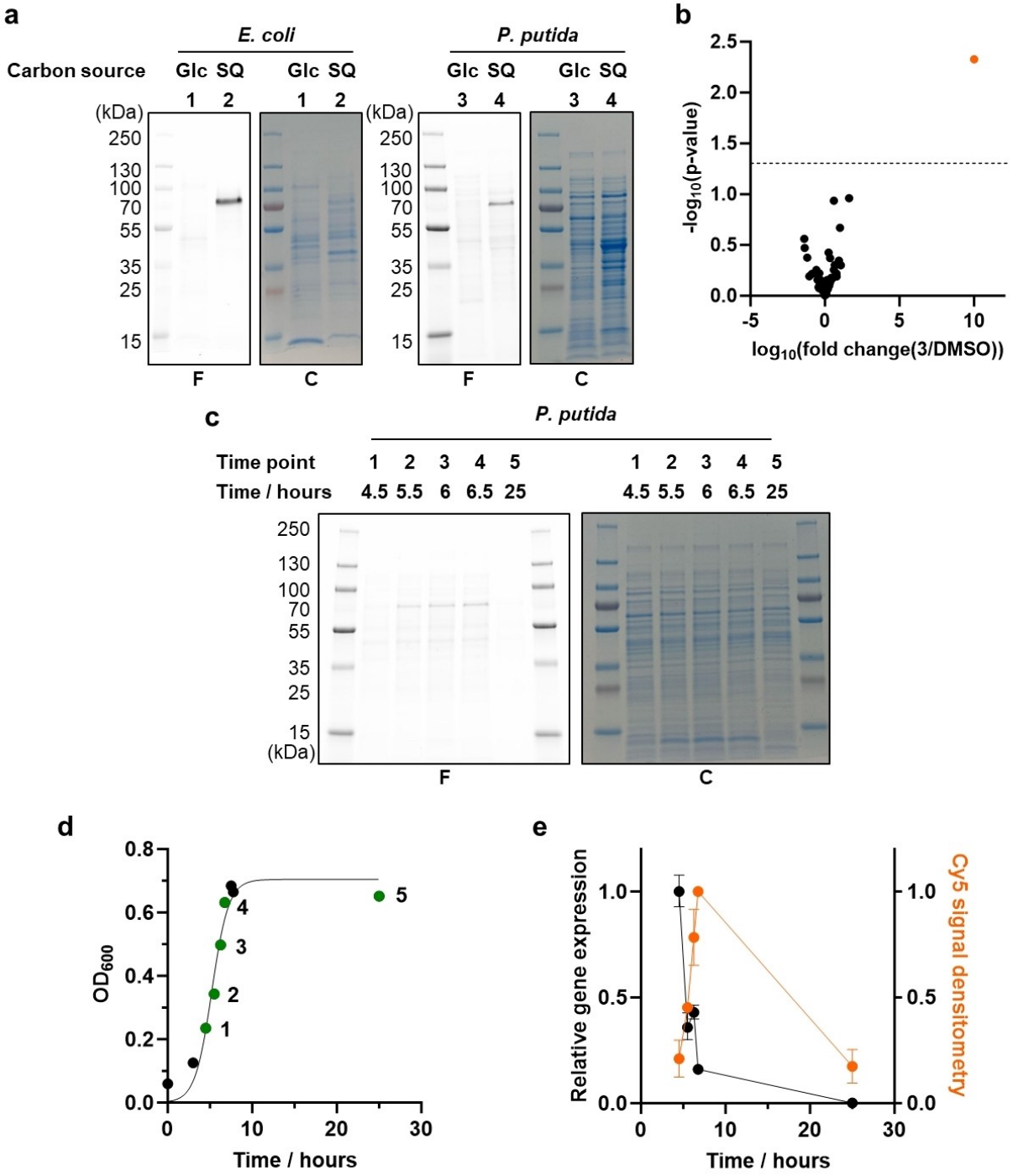


Figure 4. Selective labelling of endogenously expressed SQase protein in whole cell lysates. (**a**) Whole cell lysates of *E. coli* and *P. putida* grown on either glucose or SQ as sole carbon source, labelled with 1 μM ABP **2**. (**b**) Volcano plot of LCMS/MS analysis of pull-down experiments using *P. putida* lysates with ABP **3** versus DMSO treatment ($n=3$). A Student's t-test was performed with a maximum p-value of 0.05 (equating to a $-\log_{10}$(p-value) of 1.3; highlighted as dotted line on the plot). *Pp*SQase is indicated in orange. (**c**) Whole cell lysates of *P. putida* grown on SQ harvested at different points along the growth curve, labelled with 1 μM ABP **2**. (**d**) Growth curve of *P. putida* on SQ with harvest time-points corresponding to (c) in green. (**e**) Relative gene expression levels studied using two-step rt-qPCR of RNA extracted from the time points indicated in (c+d), shown in black, along with densitometry analysis of the Cy5 fluorescent bands at ~80 kDa in (c) normalized to the most intense band, shown in orange. Gene expression data is the average of three repeat measurements from the same sample in (c+d) (each repeat contained triplicate wells of each time-point) ± SEM of the fold change converted from the logarithmic value to the linear. Densitometry data is plotted as the average of two labelling repeats from the same growths ± SEM where each repeat has been normalized to the most intense point. In (a+c), F and C denote the fluorescent gel scan and the Coomassie stained gel, respectively.

biotin tags, followed by deprotection, afforded compounds **2** and **3**, respectively.

Labelling *At*SQase with aziridine **1** caused a notable increase in protein melting temperature, indicative of the formation of a stable, and stabilized, complex. Covalent labelling in a 1:1 stoichiometry was demonstrated by mass spectrometry analysis of the resulting complexes after incubation of five purified SQases with **1**. The 3D X-ray structure of the *Bm*SQase·**1** complex showed that labelling occurred at the catalytic nucleophile. This observation aligns with other studies of cyclophellitol aziridine inactivators of retaining glycoside hydrolases, underscoring that the sulfonated cyclophellitol aziridine hijacks the catalytic mechanism of the enzyme to achieve covalent labelling.

The Cy5-probe **2** enabled labelling and visualization of five distinct SQases, after SDS-PAGE, with a limit of detection of 50 nM protein, and labelling was selective and effective when SQases were introduced into an *E. coli* cell lysate. This success prompted efforts to use probe **2** in a more complex scenario, investigating protein abundance in two bacteria that utilize different sulfoglycolytic pathways: *E. coli* (sulfo-EMP) and *P. putida* (sulfo-ED). In both cases, treatment of cell lysates from cells grown on SQ with probe **2** revealed labelling of a band consistent with the predicted mass of the SQase from these two organisms, and SQase bands were absent in lysates from cells grown on Glc. Biotin-probe **3** was used to capture the labelled protein from within the *P. putida* lysate, and proteomics analysis verified specific labelling of *Pp*SQase.

Little is known about the abundance of sulfoglycolytic enzymes throughout various stages of bacterial growth. To understand how SQase abundance relates to gene expression in response to growth of *P. putida* on SQ, we compared the levels of *Pp*SQase detected with Cy5-probe **2** with levels of *ppsqase* transcripts detected by qPCR analysis. While *ppsqase* expression is greatest at the earliest sampled point in the growth curve, it declines as log phase progresses and cannot be detected in late stationary phase. In contrast, levels of active *Pp*SQase continue to rise and peak at late exponential/early stationary phase, and only decline in late stationary phase. This demonstrates that *Pp*SQase is a relatively long-lived protein, persisting longer in the cell than the mRNA encoding this protein, and highlighting the ability of ABPs to monitor levels of active SQase enzyme.

In conclusion, our study underscores the adaptability of cyclophellitol-based ABPs for functional studies of retaining glycosidases. Installation of the sulfonate group alongside an (tagged) aziridine was achieved through judicious use of a sulfonate protecting group, and provided inhibitor **1** and ABPs **2** and **3**. These tools proved highly selective for SQase labelling even within bacterial cell lysates, and provided unique insights into SQ metabolism. Our ABPP approach provides insights that complement molecular biology approaches involving transcript analysis and proteomics. Specifically, rt-qPCR of bacteria grown on SQ demonstrates increased SQase gene expression in early-mid log phase growth, but shows a sharp drop in SQase gene expression in late log/early stationary phase. Our ABPP methodology unveils SQase activity in later stages, revealing a long-lived SQase protein. This dynamic insight is unattainable with transcript analyses methods, because mRNA encoding the enzyme is absent in later stages of growth. In summary, the ABP technology offers a versatile means to probe family GH31 SQase activity that is independent of degradation pathway, shedding light on how microorganisms catabolize a key organosulfur species. We propose that this technology can be applied to explore SQase activity in diverse microorganisms or environmental samples.

Supporting Information

The authors have cited additional references within the Supporting Information.[4,5b,6,11,13–14,20]

Author Contributions

G.J.D., H.S.O. and S.J.W. conceived the project, obtained funding, planned and wrote the paper. Z.L. performed chemical syntheses, supervised by J.D.C. and H.S.O. M.S. and B.M., supervised by G.J.D., initiated the ABPP work in York, M.S. solving the 3D structure and performing biochemical analyses. L.P. conducted bioinformatics analysis. I.B.P. performed proteomics, labelling, pulldown, mass spectrometry and qPCR work, supervised by G.J.D.

Acknowledgements

We thank the Royal Society (Ken Murray Research Professorship to G.J.D.), the Biotechnology and Biological Sciences Research Council (BBSRC) (grant BB/W003805/1 to G.J.D. and M.S.), the Netherlands Organization for Scientific Research (NWO TOP grant 2018–714.018.002 to H.S.O.), the European Research Council (ERC-2011-AdG-290836 "Chembiosphing" to H.S.O. and ERC-2020-SyG-951231 "Carbocentre" to G.J.D. and H.S.O.), and the Australian Research Council (DP210100233, DP210100235, DP240100126 to S.J.W.). Mass spectrometry was performed at the York Centre of Excellence in Mass Spectrometry, which benefited from capital investment through Science City York, supported by Yorkshire Forward with funds from the Northern Way Initiative, and from the EPSRC (EP/K039660/1; EP/M028127/1). We thank Hans van den Elst (University of Leiden) and Andrew Leech, Adam Dowle, Chris Taylor, Chloë Baldreki, Sally James, Lesley Gilbert and Samantha Donninger for technical support and advice (all University of York Bioscience Technology Facility). We thank Diamond Light Source for beamtime (proposals 24948, and 18598), and Johan Turkenburg, Sam Hart and the staff of beamline I03 for assistance with crystal testing and data collection.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

Atomic coordinate files and structure factors have been deposited in the Protein DataBank (PDB) with accession code 8R56 (*Bm*SQase·**1**). Data collection and refinement statistics are presented in Supporting Information Table S1. Proteomics data has been deposited to the PRIDE partner repository and can be accessed with the data set identifier PXD047091.

Keywords: Sulfur cycle · sulfoquinovose · activity-based probe · enzymes · hydrolases

[1] a) A. A. Benson, H. Daniel, R. Wiser, *Proc. Natl. Acad. Sci. USA* **1959**, *45*, 1582–1587; b) E. D. Goddard-Borger, S. J. Williams, *Biochem. J.* **2017**, *474*, 827–849; c) J. L. Harwood, R. G. Nicholls, *Biochem. Soc. Trans.* **1979**, *7*, 440–447.

[2] a) G. Hölzl, P. Dörmann, *Annu. Rev. Plant Biol.* **2019**, *70*, 51–81; b) N. Mizusawa, H. Wada, *Biochim. Biophys. Acta Bioenerg.* **2012**, *1817*, 194–208.

[3] a) J. Liu, Y. Wei, K. Ma, J. An, X. Liu, Y. Liu, E. L. Ang, H. Zhao, Y. Zhang, *ACS Catal.* **2021**, *11*, 14740–14750; b) A. J. D. Snow, L. Burchill, M. Sharma, G. J. Davies, S. J. Williams, *Chem. Soc. Rev.* **2021**, *50*, 13628–13645; c) Z. Ye, Y. Wei, L. Jiang, Y. Zhang, *iScience* **2023**, *26*, 107803.

[4] a) 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, S. J. Williams, *ACS Cent. Sci.* **2018**, *4*, 1266–1273; b) G. Speciale, Y. Jin, G. J. Davies, S. J. Williams, E. D. Goddard-Borger, *Nat. Chem. Biol.* **2016**, *12*, 215–217.

[5] a) K. Denger, M. Weiss, A.-K. Felux, A. Schneider, C. Mayer, D. Spiteller, T. Huhn, A. M. Cook, D. Schleheck, *Nature* **2014**, *507*, 114–117; b) A.-K. Felux, D. Spiteller, J. Klebensberger, D. Schleheck, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E4298–E4305; c) B. Frommeyer, A. W. Fiedler, S. R. Oehler, B. T. Hanson, A. Loy, P. Franchini, D. Spiteller, D. Schleheck, *iScience* **2020**, *23*, 101510; d) Y. Liu, Y. Wei, Y. Zhou, E. L. Ang, H. Zhao, Y. Zhang, *Biochem. Biophys. Res. Commun.* **2020**, *533*, 1109–1114.

[6] M. Sharma, J. P. Lingford, M. Petricevic, A. J. D. Snow, Y. Zhang, M. A. Järvå, J. W.-Y. Mui, N. E. Scott, E. C. Saunders, R. Mao, R. Epa, B. M. da Silva, D. E. V. Pires, D. B. Ascher, M. J. McConville, G. J. Davies, S. J. Williams, E. D. Goddard-Borger, *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2116022119.

[7] A. Kaur, I. B. Pickles, M. Sharma, N. Madeido Soler, N. E. Scott, S. J. Pidot, E. D. Goddard-Borger, G. J. Davies, S. J. Williams, *J. Am. Chem. Soc.* **2023**, *145*, 28216–28223.

[8] T. Arumapperuma, J. Li, B. Hornung, N. M. Soler, E. D. Goddard-Borger, N. Terrapon, S. J. Williams, *J. Biol. Chem.* **2023**, *299*, 103038.

[9] a) Y. Liu, M. P. Patricelli, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 14694–14699; b) D. Leung, C. Hardouin, D. L. Boger, B. F. Cravatt, *Nat. Biotechnol.* **2003**, *21*, 687–691; c) M. D. Witte, W. W. Kallemeijn, J. Aten, K.-Y. Li, A. Strijland, W. E. Donker-Koopman, A. M. C. H. van den Nieuwendijk, B. Bleijlevens, G. Kramer, B. I. Florea, B. Hooibrink, C. E. M. Hollak, R. Ottenhoff, R. G. Boot, G. A. van der Marel, H. S. Overkleeft, J. M. F. G. Aerts, *Nat. Chem. Biol.* **2010**, *6*, 907–913.

[10] a) W. W. Kallemeijn, K.-Y. Li, M. D. Witte, A. R. A. Marques, J. Aten, S. Scheij, J. Jiang, L. I. Willems, T. M. Voorn-Brouwer, C. P. A. A. van Roomen, R. Ottenhoff, R. G. Boot, H. van den Elst, M. T. C. Walvoort, B. I. Florea, J. D. C. Codée, G. A. van der Marel, J. M. F. G. Aerts, H. S. Overkleeft, *Angew. Chem. Int. Ed.* **2012**, *51*, 12529–12533; b) L. I. Willems, H. S. Overkleeft, S. I. van Kasteren, *Bioconjugate Chem.* **2014**, *25*, 1181–1191; c) L. Wu, J. Jiang, Y. Jin, W. W. Kallemeijn, C.-L. Kuo, M. Artola, W. Dai, C. van Elk, M. van Eijk, G. A. van der Marel, J. D. C. Codée, B. I. Florea, J. M. F. G. Aerts, H. S. Overkleeft, G. J. Davies, *Nat. Chem. Biol.* **2017**, *13*, 867–873; d) A. R. A. Marques, L. I. Willems, D. Herrera Moro, B. I. Florea, S. Scheij, R. Ottenhoff, C. P. A. A. van Roomen, M. Verhoek, J. K. Nelson, W. W. Kallemeijn, A. Biela-Banas, O. R. Martin, M. B. Cachón-González, N. N. Kim, T. M. Cox, R. G. Boot, H. S. Overkleeft, J. M. F. G. Aerts, *ChemBioChem* **2017**, *18*, 402–412; e) L. Wu, Z. Armstrong, S. P. Schröder, C. de Boer, M. Artola, J. M. F. G. Aerts, H. S. Overkleeft, G. J. Davies, *Curr. Opin. Chem. Biol.* **2019**, *53*, 25–36; f) S. P. Schröder, C. de Boer, N. G. S. McGregor, R. J. Rowland, O. Moroz, E. Blagova, J. Reijngoud, M. Arentshorst, D. Osborn, M. D. Morant, E. Abbate, M. A. Stringer, K. B. R. M. Krogh, L. Raich, C. Rovira, J.-G. Berrin, G. P. van Wezel, A. F. J. Ram, B. I. Florea, G. A. van der Marel, J. D. C. Codée, K. S. Wilson, L. Wu, G. J. Davies, H. S. Overkleeft, *ACS Cent. Sci.* **2019**, *5*, 1067–1078; g) N. G. S. McGregor, M. Artola, A. Nin-Hill, D. Linzel, M. Haon, J. Reijngoud, A. Ram, M.-N. Rosso, G. A. van der Marel, J. D. C. Codée, G. P. van Wezel, J.-G. Berrin, C. Rovira, H. S. Overkleeft, G. J. Davies, *J. Am. Chem. Soc.* **2020**, *142*, 4648–4662; h) Z. Armstrong, C.-L. Kuo, D. Lahav, B. Liu, R. Johnson, T. J. M. Beenakker, C. de Boer, C.-S. Wong, E. R. van Rijssel, M. F. Debets, B. I. Florea, C. Hissink, R. G. Boot, P. P. Geurink, H. Ovaa, M. van der Stelt, G. M. van der Marel, J. D. C. Codée, J. M. F. G. Aerts, L. Wu, H. S. Overkleeft, G. J. Davies, *J. Am. Chem. Soc.* **2020**, *142*, 13021–13029; i) C. de Boer, N. G. S. McGregor, E. Peterse, S. P. Schröder, B. I. Florea, J. Jiang, J. Reijngoud, A. F. J. Ram, G. P. van Wezel, G. A. van der Marel, J. D. C. Codée, H. S. Overkleeft, G. J. Davies, *RSC Chemical Biology* **2020**, *1*, 148–155; j) Y. Chen, Z. Armstrong, M. Artola, B. I. Florea, C.-L. Kuo, C. de Boer, M. S. Rasmussen, M. Abou Hachem, G. A. van der Marel, J. D. C. Codée, J. M. F. G. Aerts, G. J. Davies, H. S. Overkleeft, *J. Am. Chem. Soc.* **2021**, *143*, 2423–2432.

[11] a) F. G. Hansen, E. Bundgaard, R. Madsen, *J. Org. Chem.* **2005**, *70*, 10139–10142; b) K.-Y. Li, J. Jiang, M. D. Witte, W. W. Kallemeijn, H. van den Elst, C.-S. Wong, S. D. Chander, S. Hoogendoorn, T. J. M. Beenakker, J. D. C. Codée, J. M. F. G. Aerts, G. A. van der Marel, H. S. Overkleeft, *Eur. J. Org. Chem.* **2014**, *2014*, 6030–6043.

[12] N. W. McGill, S. J. Williams, *J. Org. Chem.* **2009**, *74*, 9388–9398.

[13] O. Kaczmarek, H. A. Scheidt, A. Bunge, D. Föse, S. Karsten, A. Arbuzova, D. Huster, J. Liebscher, *Eur. J. Org. Chem.* **2010**, *2010*, 1579–1586.

[14] P.-C. Lin, S.-H. Ueng, S.-C. Yu, M.-D. Jan, A. K. Adak, C.-C. Yu, C.-C. Lin, *Org. Lett.* **2007**, *9*, 2131–2134.

[15] 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, S. J. Williams, *Appl. Environ. Microbiol.* **2020**, *86*, e00750–00720.

[16] G. J. Davies, A. Planas, C. Rovira, *Acc. Chem. Res.* **2012**, *45*, 308–316.

[17] A. L. Lovering, S. S. Lee, Y.-W. Kim, S. G. Withers, N. C. J. Strynadka, *J. Biol. Chem.* **2005**, *280*, 2105–2115.

[18] a) B. Chandrasekar, T. Colby, A. Emran Khan Emon, J. Jiang, T. N. Hong, J. G. Villamor, A. Harzen, H. S. Overkleeft, R. A. L. van der Hoorn, *Mol. Cell. Proteomics* **2014**, *13*, 2787–2800; b) L. I. Willems, T. J. M. Beenakker, B. Murray, S. Scheij, W. W. Kallemeijn, R. G. Boot, M. Verhoek, W. E. Donker-Koopman, M. J. Ferraz, E. R. van Rijssel, B. I. Florea,

J. D. C. Codée, G. A. van der Marel, J. M. F. G. Aerts, H. S. Overkleeft, *J. Am. Chem. Soc.* **2014**, *136*, 11622–11625; c) J. Jiang, W. W. Kallemeijn, D. W. Wright, A. M. C. H. van den Nieuwendijk, V. C. Rohde, E. C. Folch, H. van den Elst, B. I. Florea, S. Scheij, W. E. Donker-Koopman, M. Verhoek, N. Li, M. Schürmann, D. Mink, R. G. Boot, J. D. C. Codée, G. A. van der Marel, G. J. Davies, J. M. F. G. Aerts, H. S. Overkleeft, *Chem. Sci.* **2015**, *6*, 2782–2789.

[19] S. C. Miller, *J. Org. Chem.* **2010**, *75*, 4632–4635.

[20] a) V. B. Chen, W. B. Arendall III, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. Sect. D* **2010**, *66*, 12–21; b) P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D* **2004**, *60*, 2126–2132; c) P. Evans, *Acta Crystallogr. Sect. D* **2006**, *62*, 72–82; d) W. Kabsch, *Acta Crystallogr. Sect. D* **2010**, *66*, 125–132; e) L. Käll, J. D. Storey, W. S. Noble, *Bioinformatics* **2008**, *24*, i42–i48; f) A. A. Lebedev, P. Young, M. N. Isupov, O. V. Moroz, A. A. Vagin, G. N. Murshudov, *Acta Crystallogr. Sect. D* **2012**, *68*, 431–440; g) F. Long, R. A. Nicholls, P. Emsley, S. Grazulis, A. Merkys, A. Vaitkus, G. N. Murshudov, *Acta Crystallogr. Sect. D* **2017**, *73*, 112–122; h) S. McNicholas, E. Potterton, K. S. Wilson, M. E. M. Noble, *Acta Crystallogr. Sect. D* **2011**, *67*, 386–394; i) D. Mellacheruvu, Z. Wright, A. L. Couzens, J.-P. Lambert, N. A. St-Denis, T. Li, Y. V. Miteva, S. Hauri, M. E. Sardiu, T. Y. Low, V. A. Halim, R. D. Bagshaw, N. C. Hubner, A. al-Hakim, A. Bouchard, D. Faubert, D. Fermin, W. H. Dunham, M. Goudreault, Z.-Y. Lin, B. G. Badillo, T. Pawson, D. Durocher, B. Coulombe, R. Aebersold, G. Superti-Furga, J. Colinge, A. J. R. Heck, H. Choi, M. Gstaiger, S. Mohammed, I. M. Cristea, K. L. Bennett, M. P. Washburn, B. Raught, R. M. Ewing, A.-C. Gingras, A. I. Nesvizhskii, *Nat. Methods* **2013**, *10*, 730–736; j) A. I. Nesvizhskii, A. Keller, E. Kolker, R. Aebersold, *Anal. Chem.* **2003**, *75*, 4646–4658; k) A. Vagin, A. Teplyakov, *J. Appl. Crystallogr.* **1997**, *30*, 1022–1025; l) G. Winter, *J. Appl. Crystallogr.* **2010**, *43*, 186–190; m) N. Oberg, R. Zallot, J. A. Gerlt, *J. Mol. Biol.* **2023**, *435*, 168018; n) R. Zallot, N. Oberg, J. A. Gerlt, *Biochemistry* **2019**, *58*, 4169–4182; o) G. E. H. G. C. Crooks, J.-M. Brenner, *Genome Research* **2004**, *14*; p) G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr. Sect. D* **1997**, *53*, 240–255; q) P. Shannon, M. A. O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, **2003**, *13*; r) K. Tamura, G. Stecher, S. Kumar, *Mol. Biol. Evol.* **2021**, *38*, 3022–3027.

[21] a) E. van Meel, E. Bos, M. J. C. van den Lienden, H. S. Overkleeft, S. I. van Kasteren, A. J. Koster, J. M. F. G. Aerts, *Traffic* **2019**, *20*, 346–356.

Manuscript received: January 19, 2024
Accepted manuscript online: April 22, 2024
Version of record online: May 21, 2024