Novel sulfatase cancer therapeutics negatively impact Bacteroidota of the colonic microbiota in a non-sulfatase dependent manner

Conor J. Crawford[1], Christian Gunawan[2], Zongjia Chen[2], Dominic P Byrne[3], Cosette Darby[4], Ana S Luis[5], Stefano Elli[6], Edwin A Yates[3], David N Bolam[4], Sjoerd van der Post[5], Spencer J Williams[2]*, and Alan Cartmell[7,8,9]*

[1]Max Planck Institute of Colloids and Interfaces, Am Mühlenberg 1, 14476, Potsdam, Germany
[2]School of Chemistry and Bio21, Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Victoria3010, Australia
[3]Department of Biochemistry, Cell and Systems biology, Institute of Systems, Molecular and Integrative biology, University of Liverpool, Liverpool L69 7ZB, U.K.
[4]Biosciences Institute, Faculty of medical sciences, Newcastle university, Medical school, Framlington place, Newcastle upon tyne, NE2 4HH, U.K.
[5]Department of Medical Biochemistry and Cell Biology, University of Gothenburg, Gothenburg, Sweden
[6]Istituto di Ricerche Chimiche e Biochimiche G. Ronzoni, Milan,20133, Italy
[7]Department of Biology, University of York, Wentworth Way, York, YO10 5DD, U.K.
[8]York Structural Biology Laboratory (YSBL), University of York, Wentworth Way, York, YO10 5DD, U.K.
[9]York Biomedical Research Institute (YBRI), University of York, Wentworth Way, York, YO10 5DD, U.K.

*To whom correspondence should be addressed:
alan.cartmell@york.ac.uk
sjwill@unimelb.edu.au

Abstract

The S1 family of sulfatases is the sole family of enzymes that desulfate carbohydrates, linking them to inflammatory bowel diseases and various cancers. There is an unmet need for therapeutics against carbohydrate sulfatases, and while effective, orally available inhibitors of S1 steroid sulfatases exist, their efficacy against S1 carbohydrate sulfatases remains largely unexplored. In this study we assess a library of aryl- and carbohydrate sulfamates/sulfonates for their ability to inhibit carbohydrate sulfatases from human colonic bacteria. Surprisingly, these compounds prove ineffective. We show that arylsulfamate inhibitors inhibit the growth of anti-cancer human gut microbiota species independent of sulfatase activity. Leveraging thermal proteome profiling, we identify the non-sulfatase targets responsible for these effects. This work highlights the imperative for developing specific inhibitors targeting carbohydrate sulfatases and unveils the adverse effects of arylsulfamates, a class of drugs designed for hormone cancer treatment, on the human gut microbiota, potentially influencing their therapeutic efficacy.

Introduction

Sulfatases are enzymes that perform the enzymatic desulfation of various biomolecules. These biomolecules include small molecules, such as hormones, and larger macromolecules, such as proteins and complex carbohydrates. The SulfAtlas database classifies sulfatases into four families, S1, S2, S3, and S4[1]. Of the over 160,000 sequences catalogued in the SulfAtlas database, the majority (over 90%, or >145,000 sequences) belong to the S1 family[2]. The S1 family occurs across all domains of life and members desulfate all known classes of sulfated biomolecules[1,3,4]. The remaining approximately 10% of sulfatase sequences are distributed across the S2, S3, and S4 families. These are exclusive to bacteria and are active on alkyl- and arylsulfates[1].

The significance of the S1 sulfatase family lies in its size and distribution, with 110 subfamilies based on sequence homology. The dysfunction of S1 sulfatases is implicated in a range of human diseases. Human steroid sulfatase (STS) plays a critical role in catalysing the desulfation of the hormones dehydroepiandrosterone (DHEA) sulfate and esterone (E1) sulfate[5], as well as other oestrogen-based hormones, converting them into their unsulfated active forms[6]. Elevated levels of STS are correlated with increased hormone levels in breast cancer tumours, primarily in post-menopausal women, leading to poorer patient prognosis[5]. Furthermore, STS activity has also been linked to prostate cancer in men, influencing the production of testosterone and androstenediol from DHEA sulfate[7].

The S1 human carbohydrate sulfatases Sulf1 and Sulf2, catalyse desulfation of the glycosaminoglycan (GAG) heparan sulfate at the *O*6 position of *N*-acetyl-D-glucosamine (GlcNAc) within the extracellular matrix (ECM)[8,9]. Both Sulf1 and Sulf2 are overexpressed in most cancers, and their heightened expression is essential for metastasis in small cell lung carcinomas[10,11]. Bacterial S1 carbohydrate sulfatases are essential for some symbiotic colonic bacteria to colonise and utilise host colonic mucin and have been implicated in ulcerative colitis (UC)[12,13], and colon cancer (CC)[14,15]. Excessive activity of these bacterial sulfatases leads to over-degradation and alteration of the protective colonic mucin layer, exposing the gut epithelium to the normally symbiotic human gut microbiota (HGM). This exposure leads to gut inflammation, microbial dysbiosis, and the development of disease states such as UC[12,13,16] and CC[14,17]. Inhibitors targeting S1 sulfatases may prove effective in treating these diseases. Some promising inhibitors of the S1 family have already been developed[18,19], offering potential approaches to treat conditions like UC and CC.

The S1 family of sulfatases has a conserved alkaline phosphatase-like fold characterized by a larger N-terminal domain housing a central mixed beta sheet flanked by alpha helices. This domain is abutted by a smaller C-terminal subdomain comprising a 4 stranded antiparallel beta sheet, and a single alpha helix[20] **(Figure 1a)**. Within this structural

framework the S1 family employs a non-canonical amino acid as the catalytic nucleophile known as formylglycine (FGly). FGly is generated co-translationally from a Cys or Ser within the consensus sequence C/S-X-P/A-X-R, and is situated within an invariant sulfate (S) binding site[3] **(Figure 1a)**. The arylsulfamate class of inhibitors have been developed to target the distinctive FGly residue of the S1 family.[21].

Arylsulfamate inhibitors have the general formula Ar-O-SO_2NH_2. These inhibitors, developed by replacing the sulfate group with a sulfamate on aryl scaffolds, prove to be effective, irreversible, active site directed inhibitors of S1 steroid sulfatases. Despite their efficacy, the precise nature of the dead-end complex formed is unknown[21]. Arylsulfamates demonstrate broad effectiveness against S1 steroid sulfatases of both bacterial and eukaryotic origin[18]. The initial development of arylsulfamates targeted human STS for breast cancer treatment. Arylsulfamates mimic sulfated steroids, the natural substrates of STS. Notable arylsulfamate inhibitors of STS are Irosustat[22], estrone 3-O-sulfamate (EMATE)[23], and estradiol sulfamate (E2MATE)[24] **(Figure 1b)**. Irosustat, a first-in-class STS inhibitor has progressed into phase II clinical trials[25]. It has shown positive outcomes in treating oestrogen receptor-positive breast cancer in post-menopausal women, both in early[26] and late-stage cancers[25], and oestrogen receptor-positive endometrial cancer[27]. EMATE, the first oestrogen sulfamate tested against cancer cells, exhibits almost complete reduction in STS activity and a reduction in mammary tumour growth in rats upon both oral and subcutaneous administration[28]. E2MATE, originally developed as an estradiol prodrug, reached phase II trials for hormone replacement therapy, demonstrating effectiveness in a clinical trial[28]. Additionally, it was explored as an STS inhibitor for treatment of endometriosis, and showed a 91% reduction in endometrial STS in a phase I clinical trial of 24 healthy women of reproductive age[28]. All three inhibitors exhibit good bioavailability, are well tolerated, and have nanomolar activities against STS[22,25,27,28].

Given the invariant nature of the sulfate binding site, it was initially assumed that sulfamate based inhibitors would be universally effective across all sulfatases within the S1 family. However, limited tests have suggested arylsulfamates are ineffective inhibitors of S1 carbohydrate sulfatases[29,30], but this has yet to be systematically investigated. The specificity and a significant portion of the binding energy of carbohydrate sulfatases are dictated by their glycan binding regions, which target hydrophilic carbohydrates rather than hydrophobic steroids[14,20], and is line with low activity on arylsulfates. Consequently, efforts have been made to incorporate sulfamate groups into carbohydrate substrates of S1 carbohydrate sulfatases, aiming to achieve results similar to those observed for S1 steroid sulfatases.

While some success has been reported against the human carbohydrate sulfatase Sulf1 using sulfamate-modified disaccharide, trisaccharide, and tetrasaccharide substrate

analogues, these compounds functioned as reversible competitive inhibitors[31,32]. This suggests a fundamental change in the reactivity of the sulfamate group when transitioning from an aryl to a carbohydrate scaffold. There has not been systematic analysis of the effectiveness of sulfamate inhibitors, nor the newer sulfur-based covalent modifiers such as vinyl sulfones and fluorosulfates, against S1 carbohydrate sulfatases, nor has there been an assessment of the impact of these inhibitors on the HGM. *Bacteroides* species within the HGM encode a large number of S1 sulfatases[33,34] and may be susceptible to arylsulfamates themselves, making this point relevant for two reasons: 1) *Bacteroides* species are crucial for positive outcomes for patients undergoing chemotherapy for pancreatic ductal adenocarcinoma (PDAC)[35], which may be applicable to a broader range of cancers; 2) concentrations of orally ingested drugs are elevated in the colon compared to plasma, and recent data suggests this significantly influences the HGM[36,37].

Here, we examine the ability of a library of aryl- and carbohydrate sulfamates/sulfonates, as well as vinyl sulfones and fluorosulfates, **(Figure 1c,d)** to inhibit carbohydrate sulfatases from human colonic bacteria, demonstrating a lack of effectiveness. Furthermore, we show that arylsulfamate inhibitors, which are delivered orally to treat cancer, inhibit the growth of anti-cancer HGM Bacteroidota species in a non-sulfatase dependent manner, but these effects can be overcome through growth on select complex glycans. Finally, utilising thermal proteome profiling we identify the non-sulfatase targets through which these effects are mediated.

Results

Arylsulfamates do not inhibit S1 HGM carbohydrate sulfatases

In our initial experiments, we evaluated a panel of arylsulfamates to assess their inhibitory potential on ten recombinantly expressed, S1 carbohydrate sulfatases from the HGM members *Akkermansia muciniphila ATCC BAA-835 (A. muc*) and *Bacteroides thetaiotaomicron VPI-5482* (*B. theta*). As a positive control we conducted assays on our arylsulfamate library, which included three negative control methylated sulfamates. We included two S1 steroid sulfatases: PaAstA, belonging to S1_4 and from *Pseudomonas aeruginosa*, and the commercially available snail sulfatase from *Helix pomatia* (HpSulf), from S1_2. The efficacy of each compound was determined by measuring the sulfatase activity of either enzyme incubated in the presence of inhibitor, or following ~24h pre-incubation and jump-dilution 'washout'[38]. At a concentration of 1 mM, all the arylsulfamates were broadly equal in their ability to completely inhibit both enzymes with only minor differences observed between the two **(Figure 2, Figure S1-S4, Table S1 and S2).** Against PaAsta, partial sulfatase activity could be measured for 17 and 18 before complete inactivation **(Figure 2a, left panel and Figure S1)**, and this residual activity could be eliminated by pre-incubation

with the two compounds **(Figure 2a, right panel)**. Similarly, for HpSulf, initial rates that could be measured for 8 and 14 were blocked in the jump dilution experiment **(Figure 2a and Figure S3)**. These findings suggest different on rates between the enzymes and are consistent with the irreversible, covalent, nature of the inhibitor compounds. Importantly, near wildtype activity was observed for the *N* methyl and *N, N* dimethyl sulfamate control compounds 2, 5, and 6 **(Figure 2a, right panel)**.

Having confirmed the effectiveness of the arylsulfamates inhibitors against two S1 steroid sulfatases we next assayed their effectiveness against 10 bacterial carbohydrate sulfatases that target common linkages found in host glycans: four from S1_11 (BT3177$^{6S\text{-}GlcNAc}$, BT4656$^{6S\text{-}GlcNAc}$, Amuc1033$^{6S\text{-}GlcNAc}$, and Amuc1074$^{6S\text{-}GlcNAc}$) targeting *O*6 sulfated *N*-acetyl-D-glucosamine (6S-GlcNAc), four from S1_20 (BT1636$^{3S\text{-}Gal}$, BT1622$^{3S\text{-}Gal/GalNAc}$, Amuc0451$^{3S\text{-}Gal}$, and Amuc0491$^{3S\text{-}Gal/GalNAc}$) targeting *O*3 sulfated D-galactose/*N*-acetyl-D-galactosamine (3S-Gal/GalNAc), and two from S1_16 (BT3796$^{4S\text{-}Gal/GalNAc}$ and Amuc1755$^{4S\text{-}Gal}$) targeting *O*4 sulfated D-galactose/*N*-acetyl-D-galactosamine (4S-Gal/GalNAc). No significant inhibition was observed for any of the enzymes when assaying their activity with 1 mM compound, even after pre-incubation **(Figure 2b, Supplementary Figure S5-S24, Table S1 and S2)**. This indicates that the arylsulfamate compounds do not inhibit S1 HGM carbohydrate sulfatases.

Arylsulfamates affect the growth of Bacteroides thetaiotaomicron

Given that arylsulfamate inhibitors of STS are administered orally[25,27], exposing members of the HGM such as *B. theta* in any treatment regimen, we aimed to explore their impact on the growth of HGM species. We cultivated *B. theta* in brain heart infusion (BHI) media, and in minimal media supplemented with various carbon sources including D-glucose (Glc), chondroitin sulfate C (CSC), heparin (Hep), larch arabinogalactan (LAG), and potato galactan (PG), in the presence of each arylsulfamate compound at 1 mM. When DMSO was used at a final concentration of 5 % DMSO (as used in the *in vitro* assay) *B. theta* failed to grow in minimal medium supplemented with Hep or LAG, and grew poorly with CSC and Glc. Instead, a final DMSO concentration of 1 % was used, which did not affect growth of *B. theta* under any minimal media conditions tested **(Figure S25)**.

Interestingly, we observed variable bacterial growth rates with different compound and carbon source combinations **(Figure 3 and Figure S26-S31)**. Consistently however, arylsulfamates based on a napththyl scaffold (9, 10, 11, 12, 13, 14, and 16) effectively inhibited growth in all conditions. In contrast, the effect of the phenyl-based compounds (1, 2, 4, 17 and 18) varied substantially. Whilst we observed little effect of phenyl-based compounds for bacteria grown on BHI and Glc, growth on the sulfated host glycan CSC was strongly perturbed by 1 and 5. Moreover, there was an increase in lag time with 4 and 17,

and total cessation of growth for 18, on the sulfated host glycan Hep **(Figure 3)**. Similarly to CSC, large growth defects were caused for 1 and 5 when using the unsulfated plant glycan PG as a carbon source. For the unsulfated plant glycan LAG, 1 and 17 caused elongation of the lag phase whilst 4 completely inhibited growth. Interestingly, compound 3, which is a 4-phenylphenyl sulfamate, inhibited growth under all conditions tested. When *B. theta* was grown on BHI and Glc, the coumate scaffolds (5, 6, and 7) displayed little to no effect. These compounds, however, increased the lag phase and reduced growth when CSC was the carbon source but completely inhibited growth with Hep. On the plant glycan LAG 5 and 7 increased the lag phase and reduced growth whilst 6 completely inhibited growth. By contrast, when PG was the carbon source the effects of 5, 6, and 7 were relatively mild.

The variable impact of arylsulfamates on the growth of *B. theta*, observed across both sulfated and unsulfated glycan substrates, aligns with the *in vitro* biochemical data showing that these compounds do not target S1 carbohydrate sulfatases. Additionally, the fact that the 3 negative control methyl sulfamate compounds (2, 5, and 6) also influenced the growth *B. theta* provides further support that the observed growth defects are not mediated through an S1 sulfatase dependent mechanism.

Inhibition assays using cell lysates show carbohydrate sulfatase activity

To confirm the hypothesis that the growth defects observed on the sulfated host glycans were not due to S1 carbohydrate sulfatase inhibition, and that the lack of inhibition on recombinantly produced enzymes was not due to any improper FGly formation, we next grew *B. theta* to mid-exponential growth phase on chondroitin sulfate A (CSA) and Hep to stimulate upregulation of native GAG S1 carbohydrate sulfatases. Cells were then harvested by centrifugation, washed, and resuspended in PBS, before being lysed and incubated with two arylsulfamates inhibitors for 1 hour before exposure to sulfated mono- or disaccharide substrates. Compounds 3 and 12, for which kill curves were plotted **(Extended data 1a and Figure S32)**, and were effective at inhibiting growth on all carbon sources tested (**Figure 3**), did not impact the production of the final desulfated products (GlcNAc for Hep substrates and GalNAc for CSA substrates, as predicted from the combined action of endogenous S1 carbohydrate sulfatases and glycoside hydrolases), as judged by TLC and HPAEC **(Extended data 1b,c)**. This confirms that compounds 3 and 12 are not effective inhibitors of the GAG S1 carbohydrate sulfatases produced *in vivo*, and that observed effects on bacterial growth were likely due to off target interactions.

Carbohydrate sulfamates failed to inhibit the S1_20 sulfatase BT1636$^{3S\text{-}Gal}$

We hypothesised that the refractory nature of S1 carbohydrate sulfatases to arylsulfamate inhibition may be related to an inability of the compound to engage with the glycan binding

region of the enzymes. This is consistent with the inactivity of S1 carbohydrate sulfatases towards 4-nitrophenol-sulfate (4NP-SO_3), which in contrast is rapidly hydrolysed by the steroid sulfatase PaAsta, and emphasises the divergent substrate specificities (and by extension inhibitor profiles) of the two enzyme classes **(Extended data figure 2)**. We next selected the S1_20 sulfatase BT1636$^{3S\text{-}Gal}$ to test if carbohydrate sulfamates would act as effective irreversible, covalent, active site inhibitors. BT1636$^{3S\text{-}Gal}$ desulfates 3-*O* sulfated D-galactose (3S-Gal) and is essential for *B. theta* to utilise colonic mucin *O*-glycans, and competitively colonise the colon in mice[14].

The substrate analogues, 3-*O* sulfamate D-galactose (23) and 3-*O* vinylsulfonyl D-galactose (25) were synthesised, as well the control compound 2-*O* sulfamate D-galactose (24), and screened alongside aryl vinyl sulfone analogues 19 and 22, aryl fluorosulfate (20), and aryl sulfamate (21) for inhibitory activity towards BT1636$^{3S\text{-}Gal}$ **(Figure 4a)**. For the aryl compounds, 20 and 22 showed moderate inhibition after pre-incubation with BT1636$^{3S\text{-}Gal}$ **(Figure 4b, Figure S33, S34, and Table S3)**. For the carbohydrate scaffold 24, which possesses a 2-*O* sulfamate, as expected, showed no inhibition but surprisingly neither did the 3-*O* sulfamate 23. Despite the moderate inhibition activity by 22 (aryl vinyl sulfone), the D-galactose variant (25), showed no activity **(Figure 4b, Figure S33, S34, and Table S3)**. Synthesis of fluorosulfate armed D-galactose could not be achieved and therefore could not be tested. Thermal shift assays (TSA) performed using wildtype Ser77 form of BT1636$^{3S\text{-}Gal}$, which is inactive when expressed in *E. coli*, (as *E. coli* can only covert Cys to FGly), showed no change in melting temperature (indicative of compound-binding) for 23, 24, or 25 at 0.01, 0.1, or 1.0 mM compound concentrations, whilst both 21 and 22 destabilised the protein (**Figure 4c, Figure S35, and Table S4**). In contrast, the canonical substrate of BT1636$^{3S\text{-}Gal}$, 3S-Gal, potently stabilised the protein (ΔTm = 0.10, 0.98, and 3.02 ºC) at equivalent concentrations **(Figure 4c, Figure S35 and Table S4)**. Collectively, these data demonstrate that the carbohydrate sulfamate 23, and its vinylsulfonyl equivalent 25, are not effective inhibitors of BT1636$^{3S\text{-}Gal}$ and appear unable to bind to the enzyme.

The effects of arylsulfamates apply across multiple HGM Bacteroidota species

We next selected 10 additional HGM Bacteroidota species to see if the growth effects observed for *B. theta* in the presence of arylsulfamates applied more broadly. We conducted growth experiments using the arylsulfamates 3 and 12, as they inhibited the growth of *B. theta* across all conditions tested (**Figure 3**), as well the phase II arylsulfamate drugs Irosustat and estradiol sulfamate. No growth was observed for any species in the presence of arylsulfamate 3, except *Phocaeicola plebius* (*P. plebius*) which had highly variable growth **(Figure 5 and Figure S36)**. The inhibitory effect of arylsulfamate 12 treatment was milder, but all species (and in particular *B. cellulosilyticus* for which no growth could be detected)

still exhibited growth defects. Irosustat potently inhibited growth, with all but one species (*B. fragilis*) showing limited or no growth **(Figure 5 and Figure S37)**. Surprisingly, only mild growth defects could be detected with the drug estradiol sulfamate **(Figure 5 and Figure S37)**.

It is clear that the effects of arylsulfamates apply broadly across the *Bacteroides* and *Phocaeicola* genera, but the response varies significantly by arysulfamate structure, and, although less widely, by species. Estradiol sulfamate (EMATE2) is a structural mimetic of the endogenous human steroid hormone, estradiol sulfate, and this chemical resemblance to a common host product may partly explain the reduced impact on the HGM species tested here.

Thermal proteome profiling to identifies off target interactions

We next took the arylsulfamate compounds 3 and 12, and determined the effect on thermal protein stability across the proteome using thermal proteome profiling (TPP) in *B. theta* cell lysates grown on CSA as a substrate. TPP is a non-targeted mass spectrometry technique which can identify protein-compound interactions based on acquired changes in a protein's thermal stability. This is quantified as an increase in the relative abundance of peptides originating from the soluble protein fraction over a temperature gradient following incubation with a ligand[39]. Protein lysates were heat treated at 10 different temperatures ranging from 45 – 72 ºC in the presence of inhibitor 3, 12 or DMSO solvent control.

The *B. theta* lysate proteome showed a narrow melt temperature ranging between 50 and 58 ºC (90% range), much narrower than that reported for *E. coli* from whole cell TPP experiments **(Extended data 3a)**. A total of 2550 proteins were identified in the combined analysis, from which 11 S1 sulfatases, including the three upregulated in response to CSA (BT1596, BT3333, and BT3349)[40] **(Extended data figure 3b,c)** were identified. The melting temperature (Tm) was established for 2112 proteins **(Extended data figure 3d)** based on the criteria of detecting at least two unique peptides in a minimum of two replicate analysis (see methods). This allowed the determination of the melting curves for 8 out of the 11 identified S1 sulfatases **(Extended data figure 4)**. The protein stability of none of the S1 sulfatases was significantly affected upon treatment with compound 3 or 12 suggesting no direct interaction with the compounds. This is in line with the hypothesis that arylsulfamate inhibitors do not exert their effects on *B. theta* via S1 sulfatases.

A total of 7 proteins for arylsulfamate 3 and 4 proteins for arylsulfamate 12 were identified as having significant changes in Tm indicating a potential interaction **(Figure 6a, b, c)**. Some of these hits can be discounted as causing the defective growth phenotype: BT0600, BT2504, and BT3464 were recently disrupted through transposon mutagenesis and no phenotype was reported for these genes when grown on a supplemented BHI

media[41]. BT0133 is a homologue of the *E. coli* protein murQ (having 45 % identity), a protein which produces 6-*O* phosphorylated GlcNAc from 6-*O* phosphorylated *N*-acetyl-D-muramic acid (MurNAc). Loss of MurQ in *E. coli* does not cause a no growth phenotype[42] and no MurNAc was present in our experiments. This leaves a total 7 unique proteins (BT0218, BT1443, BT1889, BT4322, BT4335, BT4346, and BT4480) through which an inhibitory arylsulfamate interaction could drive the observed growth defects, and of which only one was targeted by both arylsulfamate molecules **(Figure 6b,c)**; the uncharacterised protein BT4322.

BT4322 is a distant homologue of the cytoplasmic diacylglycerol kinases (DGK) YerQ and DgkB, from the gram-positive organisms *Bacillus subtilis*[43] and *Staphylococcus aureus*[44], respectively, and the human sphingosine kinase SphK1[45]. YerQ is essential for the production of the cell wall component, lipoteichoic acid (LTA), and growth of *B. subtilis*[43]. BT4322 is predicted to have the same fold and catalytic apparatus as both DgkB (PDB:2QV7) and SphK1 (PDB:3VZB), as well as similar ATP binding site motifs and lipid binding pockets **(Extended data figure 5)**. BT4322 is therefore highly likely to be a *B. theta* lipid kinase, but the nature of the lipid substrate is yet to be determined. Although LTA is specific to gram positive bacteria, BT4322 may have an analogous role to YerQ in *B. theta* membrane lipid metabolism; this would be consistent with the lack of growth phenotype in the presence of compounds 3 and 12. Furthermore, a blastp of BT4322 against the 10 other compound-sensitive bacteria (**Figure 4**) returned a single orthologue from each species, with a 100 % query coverage, and a minimum of 84 % identity for the *Bacteroides* species and 73 % for *P. plebius*, showing a high degree of conservation and lack of redundancy between these species **(Extended data figure 6 and Table S5**), and suggesting an important role for these proteins.

Diacylglycerol kinase inhibitors significantly inhibit the growth of Bacteroides species

With BT4322 being a lipid kinase we next sought to see if we could reproduce the arylsulfamate induced growth defects using commercially available sphingosine and diacylglycerol kinase inhibitors. Sphingosine kinase inhibitor (SKI) had a maximum solubility of 0.1 mM and no notable growth defect could be observed **(Figure 6d and Extended data figure 7)**. Sphingosine kinase 1 inhibitor (SK1I)-II dissolved to 0.5 mM and caused mild phenotypes in four species, manifesting as an elongated lag phase, except for *B. xylanisolvens* which showed severe defects, especially in its lag phase **(Figure 6d and Extended data figure 7)**. The solubility of diacylglcerol kinase inhibitor (DAGKI)-I and II was even lower at 125 μM and 62.5 μM, respectively. DAGKI-I was effective at inhibiting the growth of all species with only *B. cellulosilyticus*, *B. intestinalis*, and *B. oleicipnenius* showing significant growth, albeit highly variable with very extended lag phases **(Figure 6d and**

Extended data figure 8). DAGKI-II showed a milder phenotype but still inhibited the growth of all species, lowering final OD and extending the lag phase, except for *B. oleicipnenius* which reached a similar OD to non-inhibitor growths **(Figure 6d and Extended data figure 8)**. These data support the view that the arylsulfamate observed growth defects are mediated through inhibition of a diacylglycerol kinase, not a sphingosine kinase, of which BT4322 is the likely target.

Ligand docking of arylsulfamates into off-target proteins

BT4322 emerged from the TPP analyses as the top protein target through which the arylsulfamates mediated their off-target effects, and as such we performed blind ligand docking (no pocket specified) with arylsulfamates 3, 12 and Irosustat using the AF2 predicted model of BT4322. The docking results for Irosustat with BT4322 predicted binding almost exclusively to the lipid pocket, with some minor populations elsewhere. The situation was similar for arylsulfamate 3 and 12, but some binding in the ATP site was also predicted **(Extended data figure 9)**. Interestingly, when the same experiment was performed for the human sphingosine kinase, SphK1, no specific site was favoured **(Extended data 6)**. These results suggest that *in vivo* arylsulfamates bind to human S1 steroid sulfatase STS but not SphK1, but reciprocally in *Bacteroides* species they bind lipid kinases (distant SphK1 orthologues), but not carbohydrate S1 sulfatases. This would be in line with the arylsulfamates, such as Irosustat, being well tolerated in humans with relatively mild side effects.

Discussion

Arylsulfamates are effective irreversible, active site targeting, inhibitors of S1 steroid sulfatases[21] and display relatively mild side effects when administered to humans to treat hormone dependent cancers in phase II clinical trials[25,26]. S1 carbohydrate sulfatases have been implicated in driving several disease states and it was presumed that arylsulfamates would be broadly effective across the S1 family, but it is clear from the work presented here that bacterial S1 carbohydrate sulfatases of the HGM, despite sharing the invariant sulfate binding subsite (S subsite), are unaffected by arylsulfamates both *in vitro* and *in vivo*. Furthermore, the generation of an equivalent carbohydrate sulfamate not only failed to inhibit the target S1 carbohydrate sulfatase, but showed no evidence of binding **(Figure 4b,c)**. Collectively, these data demonstrate that sulfamate based inhibitors are not effective against the HGM S1 carbohydrate sulfatases tested here. The addition of an amino group to the sulfur atom must alter both its binding and/or reactivity in the S subsite. Similarly, the vinyl sulfone and fluorosulfate modifications also do not appear to be well tolerated. It is clear S1 steroid sulfatases can accommodate the changes required for arylsulfamate reactivity, but

the S1 carbohydrate sulfatases tested here cannot. With the S subsite being invariant, the apparent differing susceptibilities of these enzyme classes must be due to the distinctive chemical environments and/or structural geometries of glycan versus steroid binding subsites. Steroid binding sites have evolved to provide a hydrophobic, 'greasy', patch that can accommodate lipophilic molecules and may allow the aryl groups of arylsulfamates to adopt an alternative binding configuration relative to the natural steroid substrate, and thus facilitate binding of a sulfamate group in the S subsite. In contrast, carbohydrates are hydrophilic, and S1 carbohydrate sulfatases have polar substrate binding regions which form extensive hydrogen bonding networks[14,20,46,47]. The polar, hydrogen bond-driven nature of the glycan binding site may impose a level of rigidity on substrate interactions which may be incompatible with alternate binding modes that may be necessary to accommodate sulfamate groups.

Carbohydrate sulfamate inhibitors have been developed against the human S1 carbohydrate sulfatase Sulf1[30,31]. The authors replaced the labile sulfate in HS di-, tri, and tetrasaccharides with a sulfamate. These inhibitors have a competitive mode of action with IC_{50} values of ~700, ~30, and ~0.5 μM for the di-, tetra-, and triasaccharide sulfamates, respectively. The trisaccharide substrate has a K_M of ~26 μM, ~50 fold higher than the IC_{50} value for the trisaccharide sulfamate. However, it should be noted that the data from which the K_M was derived contain only 5 data points, 4 of which are close to, or at, V_{max} and 1 at the K_M. It is therefore likely that this K_M calculation is not accurate, and the true K_M will be substantially lower[31]. Taking this into account, and the fact HS has a disaccharide repeat, it is likely that the affinity of the trisaccharide sulfamate is largely driven by glycan binding, not the presence of the sulfamate groups, as only the trisaccharide could occupy 3 subsites spanning -1, 0, +1. The di- and tetrasaccharide sulfamates could only occupy either -1, 0 or 0, +1 subsites depending on the orientation of the bound sugar chain. Despite this, the carbohydrate sulfamate analogues did bind to the *endo*-acting Sulf1 which is in stark contrast to the *exo*-acting BT1636[3S-Gal] assayed here[31,32], suggesting the *endo*-acting S1 carbohydrate sulfatases may have increased tolerance for the binding of sulfamate groups in their S subsite.

Despite a lack of activity towards carbohydrate sulfatases, the panel of arylsulfamates tested here profoundly impacted the growth of 10 HGM *Bacteroides*, and 1 *Phocaeicola* species, both from the Bacteroidota (formerly Bacteroidetes) phylum. When considering arylsulfamates as part of hormone dependent cancer regimens, this could have a significant impact on their long-term efficacy. It was recently shown that in PDAC, the species *B. theta and B. fragilis* were required for effective response to the chemotherapy regimen FOLFIRINOX[35]. This was mediated through the production of the HGM metabolite,

3-IAA, which is found systemically, increases ROS production, and reduces autophagy in cancer cells; compromising their metabolic fitness and proliferation. The arylsulfamate drug, Irosustat, has had positive outcomes for both early and late breast cancer treatment in phase II clinical trials[25,26] but its effect on the HGM, and how this affects Irosustat's efficacy, has not been considered. Our data show that Irosustat, and arylsulfamates of varying structure, negatively impact HGM Bacteroidota species, including those identified as helpful in PDAC. Our data, also show that some structures from our library are less detrimental than others, and some of these effects can be mitigated through carbon source selection. Importantly, the phase II arylsulfamate drug estradiol sulfamate (EMATE2), tested for hormone replacement therapy and endometriosis[21,28], had very little effect on the growth of HGM Bacteroidota species. It is possible that arylsulfamate based cancer treatment could be improved by ameliorating their negative impacts on the HGM. This could be achieved through the generation of compounds with altered chemical structures and/or by co-application of selected complex glycan prebiotics.

The most likely target through which the arylsulfamates mediate their growth inhibiting effects on the HGM species was BT4322, a lipid kinase. A single homologue of BT4322, with high sequence identity, is found in the 10 other species examined here, suggesting lack of functional redundancy. Furthermore, a homologue of BT4322 has been genetically knocked out in *Bacillus subtilis*, which displayed a no growth phenotype, and is implicated in peptidoglycan metabolism and cell membrane integrity[43]. If BT4322 has a role in phospholipid biosynthesis, it may be a critical regulator of membrane integrity. Moreover, accumulation of DAG in *E. coli* membranes prevents growth under stress conditions[48], and the loss of YerQ in *B. substilis* also causes DAG accumulation in cell membranes, confirming important roles for two distantly related BT4322 homologues[43]. Unusually for bacteria, *Bacteroides spp.* also contain sphingosine lipids in their membranes[49], which may implicate BT4322 as a sphingosine kinase (given its modest homology to mammalian SphK1). However, the fact that the diacylglycerol kinase inhibitors, and not the sphingosine kinase inhibitors, caused growth defects like that observed for arylsulfamates suggests that BT4322 is a DAGK, and that accumulation of DAG in the lipid membrane is toxic for the *Bacteroides species* examined here. Furthermore, inhibition of sphingosine synthesis does not arrest growth in *B. fragilis,* but does reduce cell survival under stress conditions, and at stationary phase[49].

The work presented here reveals three significant findings: first is that aryl- and carbohydrate sulfamates and sulfonates are not effective inhibitors of *exo*-acting S1 HGM carbohydrate sulfatases. Currently, no inhibitors of these enzymes have been developed and is an area where further research is needed. Secondly, it is clear that the arylsulfamates have unintended off-target effects which negatively impact the HGM and this could have

implications for the evolution of arylsulfamates as effective pharmacological agents for the treatment of hormone dependent cancers/conditions. Thirdly, select complex carbohydrates can protect the HGM *Bacteroides* species from the harmful effects of arylsulfamates opening up future paired drug-glycan co-treatment strategies.

Acknowledgements
CJC was funded by MSCA grant MARINEGLYCAN (101029842). PHS acknowledges funding from the Max Planck Society. SVDP was supported by the Swedish Research Council (2020-02536) This project has received funding from the Academy of Medical Sciences/Wellcome Trust through the Springboard Grant (SBF005\1065 163470), the Royal Society through their research grant (RGS\R2\212050), and Wellcome through the career development award (225897/Z/22/Z) awarded to AC.

Author contributions
AC wrote the draft manuscript. AC, DPB, carried out enzymology experiments. AC, and DPB analysed enzymological data. CJC, EAY, and SJW synthesised compounds. SVDP carried out and analysed proteomic experiments and data. AC carried out bacterial growth experiments. AC supervised and provided funding for the project. All authors read and approved the manuscript.

Data availability statement
Source Data for all experiments, along with corresponding statistical test values, where appropriate, are provided within the paper and in Supplementary information. The mass spectrometry data has been deposited to the ProteomeXchange consortium via PRIDE with the dataset identified PXDXXXXXX. All data is available on request from the authors.

Code availability statement
No new codes were developed or compiled in this study

Competing interest statement
The authors declare no competing interests.

Methods

Recombinant protein production.
Catalytically active recombinant S1 sulfatases (in which Ser, within the C/S-X/P/A-X-R consensus sequence, for *B. thetaiotaomicron* sulfatases was mutated to Cys, to permit conversion of the thiol to the formylglycine (FGly) nucleophile by the *E.coli* machinery[3]), or catalytically-inactive sulfatases (in which Cys in the consensus sequence is mutated to Ser, thus preventing formation of formylglycine) were expressed in the *E. coli* strain TUNER (Novagen), and cultured to mid-exponential phase in LB supplemented with 50 µg/mL kanamycin at 37 °C, in an orbital shaker set to 180 rpm. Cells were then cooled to 16 °C, and recombinant gene expression was induced by the addition of 0.1 mM isopropyl β-D-1-thiogalactopyranoside for 16 h at 16 °C, and 180 rpm. The cells were then centrifuged at 5,000 × g and pellets resuspended in 20 mM HEPES, pH 7.4, with 500 mM NaCl before being sonicated on ice. Recombinant protein was then purified by immobilized metal ion

affinity chromatography using a cobalt-based matrix (Talon, Clontech) and, after a wash with resuspension buffer, eluted with a step gradient of 10, 50, and 100 mM imidazole in resuspension buffer. Proteins were then analysed by SDS-PAGE and appropriately pure fractions were pooled and dialysed into 10 mM HEPES pH 7.0 with 150 mM NaCl. Protein concentrations were determined by measuring absorbance at 280 nm using the molar extinction coefficient calculated by ProtParam on the ExPasy server (web.expasy.org/protparam/).

Spectrophotometric based sulfatase assays

Production of *para*-nitrophenol from *para*-nitrophenol sulfate by steroid sulfatases was monitored at A_{400nm} using a HIDEX sense or spectra max plus 384 (molecular devices) plate reader in a 96 well plate format. Reactions were monitored at ambient tempertature (20-25 °C) and were 100 µl in volume in the presence of 1 mM substrate with 100 mM Bis-Tris-Propane, pH 7.0 with 5 % (v/v) DMSO, 150 mM NaCl, and 5 mM $CaCl_2$.

Microfluidic-based de-sulfation assay

Reducing end BODIPY (maximal excitation/emission coefficient of ~503/511 nm) labelled sulfated substrates were detected using the EZ Reader II platform (Ret biochem) *via* LED-induced fluorescence, as described previously[29]. Non-radioactive microfluidic mobility shift carbohydrate sulfation assays were optimised in solution with a 12-sipper chip coated with CR8 reagent and performed using a PerkinElmer EZ Reader II system employing EDTA-based separation buffer and real-time kinetic evaluation of substrate desulfation. Pressure and voltage settings were adjusted manually (1.8 psi, upstream voltage: 2250 V, downstream voltage: 500 V) to afford optimal separation in the reaction mixture of the sulfated substrate and unsulfated glycan product, with a sample (sip) time of 0.2 s, and total assay times appropriate for the experiment. Individual de-sulfation assays were carried out at 28°C and were pre-assembled in a 384-well plate in a volume of 80 µl in the presence of substrate concentrations of 1 µM with 100 mM Bis-Tris-Propane, MES, or Tris, dependent on the pH optimum of the sulfatase being assayed, and 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. The amount of de-sulfation was directly calculated in real-time using EZ Reader II software by measuring the sulfated carbohydrate□:□unsulfated carbohydrate ratio at each time-point during the assay. The activity of sulfatase enzymes was quantified in ‘kinetic mode’ by monitoring the amount of unsulfated glycan generated over the assay time, relative to control assay with no enzyme; with sulfate elimination from the substrate limited to ~20% to prevent loss assay linearity via substrate

depletion. k_{cat}/K_M values, using the equation $V_0=(V_{max}/K_M)/S$, were determined by linear regression analysis with GraphPad Prism software.

Differential scanning fluorimetry

Thermal shift/stability assays (TSAs) were performed using a StepOnePlus Real-Time PCR machine (LifeTechnologies) and SYPRO-Orange dye, at a 1:1000 dilution, (excitation 470 nm, emission maximum 570 nm, Invitrogen) with thermal ramping between 20 and 95°C in 0.3°C step intervals per data point to induce denaturation of purified, folded, inactive, wild-type (Ser77) BT1636$^{S1–20}$ in the presence or absence of substrate or putative inhibitors. The melting temperature (Tm) corresponding to the midpoint for the protein unfolding transition was calculated by fitting the sigmoidal melt curve using the Boltzmann equation in GraphPad Prism, with R^2 values of ≥0.99, as described in[29]. Data points after the fluorescence intensity maximum were excluded from the fitting. Changes in the unfolding transition temperature compared with the control curve (ΔT_m) were calculated for each ligand. A positive ΔT_m value indicates that the ligand stabilises the protein from thermal denaturation, and confirms binding to the protein. All TSA experiments were conducted using a final protein concentration of 5 µM in 100 mM Bis-Tris-Propane (BTP), pH 7.0, and 150 mM NaCl, supplemented with the appropriate ligand concentration. Three independent assays were performed for each protein and protein ligand combination.

Bacterial growth experiments

Growth of various *Bacteroides* species in brain heart infusion media (BHI [sigma 53286-500G]) was carried out by dissolving BHI in 18.2Ω deionised water (37g per litre), and then sterilised by autoclaving. The media was then allowed to cool prior to adding haematin (1.2 mg/ml in 0.2 His-HCl pH8.0) in a 1:1000 dilution. 5 mls of BHI media was then inoculated with *Bacteroides* species and grown overnight, and diluted 1:20 into 96 well plate growths (10 µl in 200) to monitor growth. Growth of the *B. thetaiotaomicron* VPI-5482 on specified glycan sources was achieved by mixing 1:1 the 0.22 micron filter-sterilised polysaccharides (10.2 mg/ml in H_2O) with 0.22 micron filter-sterilised 2 x minimal media (per 50 ml was: 0.1g of ammonium sulfate and sodium carbonate, 0.05g cysteine, 10 ml 1 M potassium phosphate pH 7.2, 0.1 ml 1 $mg.ml^{-1}$ vitamin K, 1 ml of 0.4 $mg.ml^{-1}$ iron sulfate, 0.4 ml of 0.25 mg.ml resazurin , 0.05 ml of 0.01 $mg.ml^{-1}$ $vitamin_{B12}$, 5 ml of mineral salts for defined media, and 100 µl of 1.2 mg/ml haematin in 0.2 M His-HCl pH8.0); this gives a final glycan concentration of 5 mg/ml in a 1 x minimal media. All growth experiments were monitored continuously in 96-well plates using a cerillo stratus plate reader within a Don whitely VA-500, A85, or A35 cabinet configured for anaerobic conditions (80 % N_2, 10% CO_2, 10 % H_2).

Where appropriate, either 1 % DMSO (control growths) or 1 % DMSO and 1 mM arylsulfamates was added to growth conditions. For sphingosine kinase inhibitor (SKI, 567731) and sphingosine kinase 1 inhibitor II (SK1-II, 567741) a maximum of 0.1 and 0.5 mM, respectively, could be dissolved in BHI with 1 % DMSO. For diacylglycerol kinase inhibitors i (D5794-5MG) and ii (D5919-5MG) a maximum concentration of 125 µM and 62.5 µM, respectively, could be dissolved in BHI with 1 % DMSO but had were hazy; this did not affect measurement readings at OD_{600nm}. Overnight growth cultures were diluted 1:20 into 96 well plate growths (10 µl in 200 µl). Growth curves presented are averages of three technical replicates.

Thin layer chromatography (TLC)

B. thetaiotaomicron VPI-5482 was grown in 10 ml minimal media with 10 mg/ml (w/v) appropriate GAG (either CSA or Heparin) as the sole carbon source to mid-exponential phase in glass test tubes. Cells were harvested by centrifugation at 5,000 × g for 10 min at room temperature and washed 2 x with 5 mL PBS (pH 7.2) before being resuspended in 1 mL PBS. Cells washed in PBS were lysed and the lysates assayed against 5 mM of the appropriate substrate at 37 °C for up to 24 h. Assays were analysed by TLC, and 2 µL each sample was spotted onto silica plates and resolved in butanol:acetic acid:water (2:1:1) running buffer. The plates were dried, and the sugars were visualized using diphenylamine stain (DPA: 1 ml of 37.5% HCl, 2 ml of aniline, 10 ml of 85% H_3PO_3, 100 ml of ethyl acetate, and 2 g of diphenylamine) by heating using a heat gun set to 450°C on medium flow. Arylsulfates and their pNP products were visualised using 1 M NaOH and then heated as described for DPA stains.

High performance anion exchange chromatography (HPAEC)

Sonicated bacterial cell lysates grown on the glycosaminoglycans CSA and heparin were treated with 1 mM of arylsulfamates 3 and 12 for 1 hour then challenged with CSA or heparin mono- and disaccharides by mixing 1:1, to a final concentration of 5 mM, to see if sulfatase activity was still present. Analysis of the sugar products was performed by HPAEC using an ICS-6000 with pulsed amperometic detection using a gold work electrode and pDH reference electrode with a standard quad carbohydrate waveform. Separation was done using a Dionex PA-200 analytical column preceded by a PA-200 guard column. Samples were resolved in an isocratically in 100 mM NaOH for 20 mins, then clean with 500 mM NaOH for 10 mins, finally ran back into 100 mM NaOH for 5 mins, with a flow rate of 0.5 ml/min.

Synthesis of chemical compunds

Full detail of the synthetic procedures can be found in the supporting information.

Thermal proteome profiling

A 300 ml culture of minimal media (see 'Bacterial growth experiments' section) supplemented with 5 mg/ml chondroitin sulfate A (CSA), stored in the anaerobic overnight, and then inoculated with 10 mls of *B. thetaiotaomicron* VPI-5482 was grown overnight in BHI media. This culture was then grown to a target OD_{600nm} of ~0.5-0.6 at which point cells were centrifuged at 5000 x g and washed with PBS, this process was repeated two more times. Cells were then treated as in *Mateus et al*[39]; briefly, cells were then resuspended in lysis buffer (50µg/ml lysozyme, 1×protease inhibitor(Roche), 250 U/ml benzonase, and 1 mM $MgCl_2$ in PBS) to give an OD_{600nm} of 50. Samples were then sonicated on ice for 15 seconds three times. The lysate was then split into three: One sample was a control (1 % (v/v) DMSO) and experimental samples contained 1 mM arylsulfamate 3 or 1 mM arylsulfamate 12 (1 % (v/v) DMSO). Next 20 µl of each condition were aliquoted into a 96 well PCR plate across 10 wells in quadruplicate. The plates were then subjected to a temperature gradient from 45 – 72 °C (45, 46.8, 49.1, 52.1, 55.9, 60.3, 64.4, 67.4, 69.9, 72°C) for 3 min, followed by 3 min at room temperature. NP-40 was then added to samples to a final concentration of 0.8 % (v/v). Samples were then transferred to a 0.22 µm 96 well filter plate (Millipore: MSGVS2210) and centrifuged at 500 x g for 5 min to remove protein aggregates. Following filtration, the flowthrough was combined 1:1 with 400 mM Tris pH 8.0 and 4 % SDS (w/v). The protein concentration for each sample was determined by Bradford assay, and a volume corresponding to 10 µg for the lowest-temperature condition was used as input for all other samples in a set for analysis. Protein digestion was performed using a modified solid-phase-enhanced sample preparation (SP3) based method adapted to a 96 well format in 0.45µM filter plates (Millipore, MSRPN04)[50,51]. Premixed magnetic SpeedBeads (Cytiva, (1:1, GE45152105050250 and GE 65152105050250)) were added to the samples, combined with two volumes of ethanol and incubated for 10 min while mixing at 150 RPM. Samples were transferred to the filter plates and washed four times with 200 µL of 70% ethanol (v/v) by centrifugation for 2 min. at 2000 x g. Beads were resuspended in 40 µL 0.1 mM HEPES pH 8.5 containing trypsin/Lys-c (Thermo Scientific), 1.25 mM TCEP and 5 mM chloroacetamide (Sigma-Aldrich) and incubated at room temperature overnight while mixing 500 RPM. Peptides were collected by centrifugation, followed by a second elution with 2% DMSO. TMT 10-plex 17 labelling (Thermo Scientific) was performed at a 2:1 ratio for 1 hr while mixing at 500 RPM. The 17 labelling reaction was quenched with 0.4% hydroxylamine (v/v) for 15 min before each sample was combined and dried under vacuum. Peptides were resolved in 0.1% trifluoroacetic acid (TFA) and fractionated using 0.1% triethylamine with increasing concentrations of acetonitrile by high pH reverse phase into 8 fractions using C18

desalting columns (Pierce, 89852) pooled into 4 fractions and dried under a vacuum. Lyophilized samples are resolved in 0.1% TFA and analyzed by LC-MS/MS using a nano HPLC system (EASY-nLC 1200, Thermo Scientific) coupled to a Q-Exactive HF-X mass spectrometer (Thermo Scientific). Peptides were separated using in-house packed columns (150 x 0.075 mm) packed with Reprosil-Pur C18-AQ 3 µm particles (Dr. Maisch). Elution was performed with a 5 to 45% gradient (A: 0.1% formic acid, B :0.1% formic acid ,80% Acetontrile) in 135 minutes at 250 nl. Full mass spectra were acquired over a mass range of minimum 400 m/z and maximum 1600 m/z, with a resolution of 60,000 at 200 m/z. The isolation window for fragmentation spectra collection was set to 1Da and a fixed first mass of 100 m/z with a resolution of 30,000 at 200 m/z. The top 15 most intense peaks with a charge state ≥2 – 5 were selected for fragmentation by HCD with a normalized collision energy of 32% and subsequent excluded for selection for 30 seconds. Thermo .RAW files were converted into MZml using msConvert (proteowizard) and FragPipe (v20.0) configured with Msfragger (v3.8) and Philosopher (v5.0)[52,53] was used for protein identification and TMT quantification. Database searches were performed against the *Bacteroides thetaiotaomicron* reference proteome (4782 entries, UP000001414, 2023_10) with the following settings, parent ion mass and fragment ion mass accuracy after recalibration was set to 20 ppm. Enzyme specificity was set to trypsin, fixed modifications were set for cysteine carbamidomethylation and TMT modified lysine, variable modifications considered were methionine oxidation, protein n-terminal acetylation and TMT modification on peptide n-terminal or serine. The identified proteins and peptides were both filtered at a 1% and only identifications based on ≥2 unique peptides with a minimum ion purity of 0.75 were considered for quantification and thermal proteome profiling. Further data analysis was performed in R environment to combine the search outcome, normalize reporter ion channels. Meltome curve fitting analysis and statistical analysis were executed using the TPP-TR workflow in the TPP package[54].

Ligand docking experiments

All docking experiments were conducted using SwissDock[55] on the Expasy server using default parameter; no binding site was specified. The alphafold model of BT4322 (Q89ZQ4) and the crystal structure of SphK1 (3VZB) were downloaded, and where appropriate heteroatoms removed, and uploaded to the Swiss dock as PDB files. Small molecules to be docked where built in Jligand[56], the generated PDB or CIF files converted to MOL2 files using the openbabel server and uploaded to SwissDock as MOL2 files.

Figure 1 General example of the conserved S1 sulfatase fold and consensus active site sequence.

a. The 3-D structure of the human lysosomal sulfatase GALNS (PDB:4FDJ), the N-terminal domain is coloured red and yellow for alpha helices and beta sheets, with the C-terminal subdomain coloured cyan and magenta for alpha helices and beta sheets; a zoom in of the consensus sequence of GALNS with a converted formylglycine, and with the general consensus pictured to the right; **b.** Structure of phase I and II arylsulfamate drugs developed to inhibit STS **c.** Chemical structures of aryl- and carbohydrate sulfamate inhibitors used in the study. **d.** Chemical structures of phenylvinyl sulfone, phenylsulfonyl fluoride, phenylvinyl sulfonate, and carbohydrate vinyl sulfone.

Figure 2 The effectiveness of arylsulfamate inhibitors against S1 steroid and carbohydrate sulfatases.

a. The enzymes PaAsta and HpSulf, belonging to S1_4 (green bar) and S1_2 (yellow bar), respectively, assayed against the panel of arylsulfamates (1-19 listed in Figure 2).; **b.** The S1_11 sulfatases (BT3177$^{6S\text{-}GlcNAc}$, BT4656$^{6S\text{-}GlcNAc}$, Amuc_1033$^{6S\text{-}GlcNAc}$, Amuc_1074$^{6S\text{-}GlcNAc}$; blue bar above), the S1_20 sulfatases (BT1622$^{3S\text{-}Gal/GalNAc}$, BT1636$^{3S\text{-}Gal}$, Amuc_0451$^{3S\text{-}Gal}$, Amuc_0491$^{3S\text{-}Gal}$; pink bar above), and the S1_16 enzymes (BT3796$^{4S\text{-}Gal/GalNAc}$ and BT1755$^{4S\text{-}Gal}$; orange bar above) against the panel of arylsulfamates (1-19 listed in **Figure 2**). For both a and b the left two panels display relative activity to wildtype protein with the arylsulfamate compound in the assay, whilst the left two panels are overnight preincubation of the enzyme with selected arylsulfamates then dilution of this preincubation mix into the reaction. The numbers indicate which arylsulfamates from **Figure 1** were used and in all instances 1 mM of compound was deployed. Wt indicates reaction with enzyme and no arylsulfamate. Kinetic curves from which the data are derived are shown in supplemental figures S1-S24.

Figure 3 Growth of *Bacteroides thetaiotaomicron* on varied carbon sources in the presence of arylsulfamate inhibitors

Bacteroides thetaiotaomicron (*B.theta*) grown on varied carbon sources; two rich carbon sources, brain heart infusion media (BHI) and glucose (Glc); two sulfated host glycan carbon sources, chondroitin sulfate c (CSC) and Heparin (Hep), and two plant glycan carbon sources potato galactan (PG) and Larch arabinogalactan (LAG), in the presence of arylsulfamates inhibitors, labelled 1-19 (see **Figure 2** for structures). Complex glycans and arylsulfamates inhibitors were used at a concentration of 5 mg/ml and 1 mM, respectively. All carbon source were utilised in a prepared minimal media except BHI which was purchased from sigma and simply dissolved in water. The traces are the mean of 3 independent growths but errors bars have been omitted for clarity; individual graphs with error bars are presented in supplemental figures S25-S30.

Figure 4. The effectiveness of carbohydrate sulfamate substrate mimics against the S1 carbohydrate sulfatase BT1636$^{3S\text{-}Gal}$.

a. Structure of the carbohydrate sulfonate/sulfamates and their aryl equivalents**. b.** The graphs show the effectiveness of carbohydrate sulfamates, carbohydrate sulfovinyl, and arylsulfonates against BT1636$^{3S\text{-}Gal}$; the left graph shows relative activity to wildtype with the compounds included in the assay, whilst the right graph show the relative activity after the enzyme had been preincubated with the compounds overnight then diluted into the reaction mixture; **c.** Thermostability shift assays (TSA) using inactive BT1636$^{3S\text{-}Gal}$ against the substrate 3S-Gal, and the putative inhibitor structures carbohydrate sulfamates, carbohydrate sulfovinyl, and arylsulfonates. All compounds in **b** were used at a concentration of 1 mM.

Figure 5. Arylsulfamates broadly affect *Bacteroides* spp. of the HGM

Bacteroides species were grown in BHI media in the presence, or absence, of 1 mM of the listed arylsulfamate. The data plotted are the maximal OD_{600nm} observed for arylsulfamate treated growths relative to the untreated controls which were both ran in triplicate and

randomly paired then averages derived. For all the estradiol growths a single experiment was conducted except for *B. theta* which was ran in triplicate; this was due to limited material.

Figure 6. Global protein thermal stability reveals off target interactions of arysulfamate inhibitors 3 and 12.
a. Table of proteins which displayed an altered Tm in the presence of arylsulfamates 3 and 12. **b.** Proteins identified with a significance change in melting point upon inhibitor 3 treatment (n=2). **c.** Proteins identified with a significance change in melting point upon inhibitor 12 treatment (n=2). A dashed box indicates BT4322 is the only protein common across both groups. **d.** *Bacteroides* species were grown in BHI media in the presence, or absence, of 0.125 mM or 0.0625 mM diacylglycerol kinase inhibitor I and II, respectively. The data plotted are the maximal OD_{600nm} observed for inhibitor treated growths relative to the untreated controls which were both ran in triplicate and randomly paired then averages derived.

Extended data figure 1. Cell lysates of *Bacteroides thetaiotaomicron* incubated with two arylsulfamates inhibitors
a. The effects of varied concentrations of arylsulfamate 3 and 12 on the growth of *B. theta* in BHI media. **b.** Thin layer chromatography (TLC) analysis of cell lysates from *B. theta* cells grown to mid-exponential phase on CSA as a carbon source, and incubated with arylsulfamates 4 and 13 **(Figure 1)**, mixed 1:1 with the sulfated substrates: *O*2 sulfated uronic acid β1,3 linked to *O*4 sulfated *N*-acetylgalactosmaine (UA2S-GalNAc4S), uronic acid β1,3 linked to *O*4 sulfated *N*-acetylgalactosmaine (UA-GalNAc4S), and *O*6 sulfated *N*-acetylgalactosmaine (GalNAc6S). Below are high performance anion exchange chromatography (HPAEC) traces showing the production of GalNAc in all instances except the negative control; **c.** TLC analysis of cell lysates from *B. theta* cells grown to mid-exponential phase on Heparin as a carbon source, and incubated with arylsulfamates 4 and 13 **(Figure 1)**, mixed 1:1 with the sulfated substrates: *O*2 sulfated uronic acid β1,4 linked to *O*4 sulfated *N*-acetylgalactosmaine (UA2S-GalNAc4S), *O*6 sulfated *N*-acetylglucosmaine (GlcNAc6S), and *N*-sulfated glucosamine (GlcNS). Below are high performance anion exchange chromatography (HPAEC) traces showing the production of GlcNAc in all instances except the negative control. All reactions were in PBS, with 5 % DMSO, and incubated at 37ºC overnight.

Extended data figure 2. Activity of carbohydrate sulfatases on *para*-nitrophenol sulfate.
Sulfatases were incubated with 5 mM *para*-nitrophenol sulfate overnight in 100 mM BTP pH 7.0 with 150 mM NaCl and 5 mM $CaCl_2$. Reactions were then ran on silica based thin layer chromatography using a resolving buffer of butanol:acetic acid:water (2:1:1) and develop by soaking the plates with 1 M NaOH and heating with a heat gun. It can be seen only the steroid sulfatase *Pa*Asta shows the production of a *para*-nitrolphenol product.

Extended data figure 3. Global features of the thermal proteome profiling data
a. Distribution of the melting temperature (Tm) of *Bacteroides thetaiotaomicron* VPI-5482 (*B. theta*) lysate proteome and the whole cell *E.coli* proteome. **b.** Ranked abundance plot of all identified proteins in the 1% (v/v) DMSO treated controls, S1 sulfatase are highlighted. **c.** Relative abundance of 11 S1 sulfatases identified by mass spectrometry for *B. theta* cultured in the presence of chondroitin sulfate A (n=6). **d.** Melting curve profiles for all identified *B. theta* proteins in control (1% (v/v) DMSO) treated lysate. The thermal stability profiles for each individual S1 sulfatases are highlighted in blue.

Extended data figure 4. Global protein thermal stability analysis confirms that arysulfamate inhibitors 3 and 12 do not directly interact with S1 sulfatases.

Individual melt curve for selected S1 sulfatase for control and arysulfamate inhibitors 3 and 12 treated lysates (n=2 per treatment group).

Extended data figure 5. Structure of BT4322, SaDgKB, and Sphk1 and comparison of the ATP helix

a. Tertiary structures of BT4322, SaDgKB, and SphK1 shown as with α helices, β sheets, and loops coloured cyan, magenta, and pink. The 'ATP helix' is coloured in green. **b.** An expanded view of the ATP helix and its sequence.

Extended data figure 6. Sequence alignment of putative DAGKs from select Bacteroides species.

Sequences were aligned using the MAFFT online server and visualised using Jalview.

Extended data figure 7. The effects of sphigosine kinase inhibitors on the growth of HGM Bacteroidota species

Bacteroidota species were grown in BHI, 1 % DMSO, with 0.1 mM of sphingosine kinase inhibitor (SKI) and 0.5 mM sphingosine kinase 1 inhibitor II and the effects on growth observed. Data are technical triplicates with the standard error of the mean.

Extended data figure 8. The effects of diacylglycerol kinase inhibitors on the growth of HGM Bacteroidota species

Bacteroidota species were grown in BHI, 1 % DMSO, with 0.125 mM of diacylglycerol kinase inhibitor I (DAGKI-i) and 0.0625 mM diacylglycerol kinase inhibitor II (DAGKI-ii) and the effects on growth observed. Data are technical triplicates with the standard error of the mean.

Extended data figure 9. Blind docking of arylsulfamates onto BT4322 and SphK1.

The results of blind docking (no site specified) arylsulfamate inhibitors into the alphafold model of BT4322 (Q89ZQ4) and the crystal structure of Sphk1 (3VZB). It can be observed all arylsulfamates preferentially cluster in, or near, the lipid binding pocket. For SpKH1 there appears to be no single preferred clustering site.

Supplementary figure 1. Kinetic curves of *Pa*Asta assayed against *para*-nitrophenol sulfate with, and without, arylsulfamate inhibitors

*Pa*Asta, at a concentration of 25 nM, was assayed against 1 mM *para*-nitrophenol with and without 1 mM of various arysulfamate inhibitors included in the assay. The assay was performed in 100 mM of Bis-Tris-propane pH 7.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 2. Kinetic curves of *Pa*Asta assayed against *para*-nitrophenol sulfate after pre-incubation of the enzyme with arylsulfamate inhibitors

*Pa*Asta, which had been incubated for ~24h with 1 mM of the appropriate arylsulfamate inhibitor, was assayed using a concentration of 25 nM against 1 mM para-nitrophenol. The assay was performed in 100 mM of Bis-Tris-propane pH 7.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 3. Kinetic curves of *Hp*Sulf assayed against para-nitrophenol sulfate with, and without, arylsulfamate inhibitors

*Hp*Sulf, at a concentration of 5 μg/ml, was assayed against 1 mM para-nitrophenol with and without 1 mM of various arysulfamate inhibitors included in the assay. The assay was performed in 100 mM of Bis-Tris-propane pH 7.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 4. Kinetic curves of *Hp*Sulf assayed against para-nitrophenol sulfate after pre-incubation of the enzyme with arylsulfamate inhibitors

*Hp*Sulf, which had been incubated for ~24h with 1 mM of the appropriate arylsulfamate inhibitor, was assayed using a concentration of 50 µg/ml against 1 mM para-nitrophenol. The assay was performed in 100 mM of Bis-Tris-propane pH 7.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 5. Kinetic curves of BT3177$^{6S\text{-}GlcNAc}$ assayed against BODIPY labelled 6S-*N*-acetylglucosamine with, and without, arylsulfamate inhibitors

BT3177$^{6S\text{-}GlcNAc}$, at a concentration of 446.5 nM, was assayed against 1 µM BODIPY labelled 6S-*N*-acetylglucosamine with and without 1 mM of various arysulfamate inhibitors included in the assay. The assay was performed in 100 mM of Bis-Tris-propane pH 8.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 6. Kinetic curves of BT3177$^{6S\text{-}GlcNAc}$ assayed against BODIPY labelled 6S-*N*-acetylglucosamine after pre-incubation of the enzyme with arylsulfamate inhibitors

BT3177$^{6S\text{-}GlcNAc}$, which had been incubated for ~24h with 1 mM of the appropriate arylsulfamate inhibitor, was assayed using a concentration of 470 nM against 1 µM BODIPY labelled 6S-*N*-acetylglucosamine. The assay was performed in 100 mM of Bis-Tris-propane pH 8.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 7. Kinetic curves of BT4656$^{6S\text{-}GlcNAc}$ assayed against BODIPY labelled 6S-*N*-acetylglucosamine with, and without, arylsulfamate inhibitors

BT4656$^{6S\text{-}GlcNAc}$, at a concentration of 200 nM, was assayed against 1 µM BODIPY labelled 6S-*N*-acetylglucosamine with and without 1 mM of various arysulfamate inhibitors included in the assay. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 8. Kinetic curves of BT4656$^{6S\text{-}GlcNAc}$ assayed against BODIPY labelled 6S-*N*-acetylglucosamine after pre-incubation of the enzyme with arylsulfamate inhibitors

BT4656$^{6S\text{-}GlcNAc}$, which had been incubated for ~24h with 1 mM of the appropriate arylsulfamate inhibitor, was assayed using a concentration of 200 nM against 1 µM BODIPY labelled 6S-*N*-acetylglucosamine. The assay was performed in MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 9. Kinetic curves of Amuc1074$^{6S\text{-}GlcNAc}$ assayed against BODIPY labelled 6S-*N*-acetylglucosamine after pre-incubation of the enzyme with arylsulfamate inhibitors

Amuc1074$^{6S\text{-}GlcNAc}$, at a concentration of 6.4 µM, was assayed against 1 µM BODIPY labelled 6S-*N*-acetylglucosamine with and without 1 mM of various arysulfamate inhibitors included in the assay. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 10. Kinetic curves of Amuc1074$^{6S\text{-}GlcNAc}$ assayed against BODIPY labelled 6S-*N*-acetylglucosamine with, and without, arylsulfamate inhibitors

Amuc1074$^{6S\text{-}GlcNAc}$, which had been incubated for ~24h with 1 mM of the appropriate arylsulfamate inhibitor, was assayed using a concentration of 6.08 μM against 1 μM BODIPY labelled 6S-*N*-acetylglucosamine. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 11. Kinetic curves of Amuc1033$^{6S\text{-}GlcNAc}$ assayed against BODIPY labelled 6S-*N*-acetylglucosamine with, and without, arylsulfamate inhibitors

Amuc1033$^{6S\text{-}GlcNAc}$, at a concentration of 4.6 μM, was assayed against 1 μM BODIPY labelled 6S-*N*-acetylglucosamine with and without 1 mM of various arysulfamate inhibitors included in the assay. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 12. Kinetic curves of Amuc1033$^{6S\text{-}GlcNAc}$ assayed against BODIPY labelled 6S-*N*-acetylglucosamine after pre-incubation of the enzyme with arylsulfamate inhibitors

Amuc1033$^{6S\text{-}GlcNAc}$, which had been incubated for ~24h with 1 mM of the appropriate arylsulfamate inhibitor, was assayed using a concentration of 4.37 μM against 1 μM BODIPY labelled 6S-*N* acetylglucosamine. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 13. Kinetic curves of BT1636$^{3S\text{-}Gal}$ assayed against BODIPY labelled 3S-galactose with, and without, arylsulfamate inhibitors

BT1636$^{3S\text{-}Gal}$, at a concentration of 400 nM, was assayed against 1 μM BODIPY labelled 3S-galactose with and without 1 mM of various arysulfamate inhibitors included in the assay. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 14. Kinetic curves of BT1636$^{3S\text{-}Gal}$ assayed against BODIPY labelled 3S-galactose sulfate after pre-incubation of the enzyme with arylsulfamate inhibitors

BT1636$^{3S\text{-}Gal}$, which had been incubated for ~24h with 1 mM of the appropriate arylsulfamate inhibitor, was assayed using a concentration of 200 nM against 1 μM BODIPY labelled 3S-galactose. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 15. Kinetic curves of BT1622$^{3S\text{-}Gal/GalNAc}$ assayed BODIPY labelled 3S-*N*-acetylgalactosamine with, and without, arylsulfamate inhibitors

BT1622$^{3S\text{-}Gal/GalNAc}$, at a concentration of 6.5 μM, was assayed against 1 μM BODIPY labelled 3S-*N*-acetylgalactosamine with and without 1 mM of various arysulfamate inhibitors included in the assay. The assay was performed in 100 mM of BTP pH 8.5 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 16. Kinetic curves of BT1622$^{3S\text{-}Gal/GalNAc}$ assayed against BODIPY labelled 3S-*N*-acetylgalactosamine after pre-incubation of the enzyme with arylsulfamate inhibitors

BT1622$^{3S\text{-}Gal/GalNAc}$, which had been incubated for ~24h with 1 mM of the appropriate arylsulfamate inhibitor, was assayed using a concentration of 2 μM against 1 μM BODIPY labelled 3S-*N*-acetylgalactosamine. The assay was performed in 100 mM of BTP pH 8.5

with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 17. Kinetic curves of Amuc0451$^{3S\text{-}Gal}$ assayed against BODIPY labelled 3S-galactose with, and without, arylsulfamate inhibitors

Amuc0451$^{3S\text{-}Gal}$, at a concentration of 15.6 µM, was assayed against 1 µM BODIPY labelled 3S-galactose with and without 1 mM of various arysulfamate inhibitors included in the assay. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 18. Kinetic curves of Amuc0451$^{3S\text{-}Gal}$ assayed against BODIPY labelled 3S-galactose after pre-incubation of the enzyme with arylsulfamate inhibitors

Amuc0451$^{3S\text{-}Gal}$, which had been incubated for ~24h with 1 mM of the appropriate arylsulfamate inhibitor, was assayed using a concentration of 15.6 µM against 1 µM BODIPY labelled 3S-galactose. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 19. Kinetic curves of Amuc0491$^{3S\text{-}Gal/GalNAc}$ assayed against BODIPY labelled 3S-*N*-acetylgalactosamine with, and without, arylsulfamate inhibitors

Amuc0491$^{3S\text{-}Gal/GalNAc}$, at a concentration of 3.4 µM, was assayed against 1 µM BODIPY labelled 3S-*N*-acetylgalactosamine with and without 1 mM of various arysulfamate inhibitors included in the assay. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 20. Kinetic curves of Amuc0491$^{3S\text{-}Gal/GalNAc}$ assayed against BODIPY labelled 3S-*N*-acetylgalactosamine after pre-incubation of the enzyme with arylsulfamate inhibitors

Amuc0491$^{3S\text{-}Gal/GalNAc}$, which had been incubated for ~24h with 1 mM of the appropriate arylsulfamate inhibitor, was assayed using a concentration of 3.2 µM against 1 µM BODIPY labelled 3S-*N*-acetylgalactosamine. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 21. Kinetic curves of BT3796$^{4S\text{-}Gal/GalNAc}$ assayed against BODIPY labelled 4S-galactose with, and without, arylsulfamate inhibitors

BT3796$^{4S\text{-}Gal/GalNAc}$, at a concentration of 630 nM, was assayed against 1 µM BODIPY labelled 4S-galactose with and without 1 mM of various arysulfamate inhibitors included in the assay. The assay was performed in 100 mM of BTP pH 8.5 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 22. Kinetic curves of BT3796$^{4S\text{-}Gal/GalNAc}$ assayed against BODIPY labelled 4S-galactose after pre-incubation of the enzyme with arylsulfamate inhibitors

BT3796$^{4S\text{-}Gal/GalNAc}$, which had been incubated for ~24h with 1 mM of the appropriate arylsulfamate inhibitor, was assayed using a concentration of 630 nM against 1 µM BODIPY labelled 4S-galactose. The assay was performed in 100 mM of BTP pH 8.5 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 23. Kinetic curves of Amuc1755$^{4S\text{-}Gal}$ assayed against BODIPY labelled 4S-galactose with, and without, arylsulfamate inhibitors

Amuc1755$^{4S\text{-}Gal}$, at a concentration of 2.3 μM, was assayed against 1 μM BODIPY labelled 4S-galactose with and without 1 mM of various arysulfamate inhibitors included in the assay. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 24. Kinetic curves of Amuc1755$^{4S\text{-}Gal}$ assayed against BODIPY labelled 4S-galactose after pre-incubation of the enzyme with arylsulfamate inhibitors
Amuc1755$^{4S\text{-}Gal}$, which had been incubated for ~24h with 1 mM of the appropriate arylsulfamate inhibitor, was assayed using a concentration of 2.185 μM against 1 μM BODIPY labelled 4S-galactose. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 25. The effects of DMSO on the growth of *Bacteroides thetaiotaomicron* VPI-5482
Bacteroides thetaiotaomicron VPI-5482 was grown in rich media (BHI), or minimal media supplemented 5 mg/ml of an appropriate polysaccharide with either 5 % or 1 % DMSO. Data are technical triplicates with the standard error of the mean. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 26. The effects of arylsulfamates on the growth of *Bacteroides thetaiotaomicron* VPI-5482 grown in BHI media
Bacteroides thetaiotaomicron VPI-5482 was grown in rich media (BHI), with 1 % DMSO, and with or without 1 mM of the appropriate arylsulfamate. Data are technical triplicates with the standard error of the mean. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 27. The effects of arylsulfamates on the growth of *Bacteroides thetaiotaomicron* VPI-5482 grown in minimal media with D-glucose
Bacteroides thetaiotaomicron VPI-5482 was grown in minimal media supplemented with 5 mg/ml D-glucose (Glc), 1 % DMSO, and with or without 1 mM of the appropriate arylsulfamate. Data are technical triplicates with the standard error of the mean. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 28. The effects of arylsulfamates on the growth of *Bacteroides thetaiotaomicron* VPI-5482 grown in minimal media with Chondroitin sulfate C (CSC)
Bacteroides thetaiotaomicron VPI-5482 was grown in minimal media supplemented with 5 mg/ml Chondroitin sulfate C (CSC), 1 % DMSO, and with or without 1 mM of the appropriate arylsulfamate. Data are technical triplicates with the standard error of the mean. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 29. The effects of arylsulfamates on the growth of *Bacteroides thetaiotaomicron* VPI-5482 grown in minimal media with Heparin
Bacteroides thetaiotaomicron VPI-5482 was grown in minimal media supplemented with 5 mg/ml Heparin (Hep), 1 % DMSO, and with or without 1 mM of the appropriate arylsulfamate. Data are technical triplicates with the standard error of the mean. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 30. The effects of arylsulfamates on the growth of *Bacteroides thetaiotaomicron* VPI-5482 grown in minimal media with Potato galactan
Bacteroides thetaiotaomicron VPI-5482 was grown in minimal media supplemented with 5 mg/ml Potato galactan (PG), 1 % DMSO, and with or without 1 mM of the appropriate arylsulfamate. Data are technical triplicates with the standard error of the mean. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 31. The effects of arylsulfamates on the growth of *Bacteroides thetaiotaomicron* VPI-5482 grown in minimal media with Larch arabinogalactan

Bacteroides thetaiotaomicron VPI-5482 was grown in minimal media supplemented with 5 mg/ml Larch arabinogalactan (LAG), 1 % DMSO, and with or without 1 mM of the appropriate arylsulfamate. Data are technical triplicates with the standard error of the mean. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 32. The effects of varied concentration arylsulfamates 3 and 12 on the growth of *Bacteroides thetaiotaomicron* VPI-5482 grown in BHI media

Bacteroides thetaiotaomicron VPI-5482 was grown in BHI, 1 % DMSO, with varied concentrations of arylsulfamate 3 and 12. Data are technical triplicates with the standard error of the mean. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 33. Kinetic curves of BT1636$^{3S\text{-}Gal}$ assayed against BODIPY labelled 3S-galactose with, and without, carbohydrate and aryl- sulfonates/sulfamates

BT1636$^{3S\text{-}Gal}$, at a concentration of 270 nM, was assayed against 1 μM BODIPY labelled 3S-galactose with and without 1 mM of various arysulfamate inhibitors included in the assay. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 34. Kinetic curves of BT1636$^{3S\text{-}Gal}$ assayed against BODIPY labelled 3S-galactose sulfate after pre-incubation of the enzyme with carbohydrate and aryl- sulfonates/sulfamates

BT1636$^{3S\text{-}Gal}$, which had been incubated for ~24h with 1 mM of the appropriate arylsulfamate inhibitor, was assayed using a concentration of 270 nM against 1 μM BODIPY labelled 3S-galactose. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Numbers indicate arylsulfamate compound as in Figure 1.

Supplementary figure 35. Thermal melt curves of BT1636$^{3S\text{-}Gal}$ incubated substrate and various carbohydrate and aryl- sulfonates/sulfamates

BT1636$^{3S\text{-}Gal}$, at a concentration of 5 μM, was incubated with no compound, and varying concentrations and the effect on its melting temperature monitored, a shift in melting temperature is indicative of an interaction. The assays were performed in 100 mM of BTP pH 7.0 with 5 % DMSO, 150 mM NaCl. Assays were performed in triplicate.

Supplementary figure 36. The effects of arylsulfamates 3 and 12 on the growth of HGM Bacteroidota species

Bacteroidota species were grown in BHI, 1 % DMSO, with 1 mM of arylsulfamate 3 and 12 and the effects on growth observed. Data are technical triplicates with the standard error of the mean.

Supplementary figure 37. The effects of the phase I/II arylsulfamate drugs Irosustat and estradiol sulfamate on the growth of Bacteroidota species

Bacteroidota species were grown in BHI, 1 % DMSO, with 1 mM of Irosustat and estradiol sulfamate and the effects on growth observed. Growths with Irosustat are technical triplicates with the standard error of the mean. Growths with estradiol sulfamate represent single growth experiments with the exception of *Bacteroides thetaiotaomicron* VPI-5482 where the data is from triplicate growths with standard error of the mean.

Supplementary table 1. Kinetic rates of sulfatase when arylsulfamates were included in the assay conditions.

The concentration of arylsulfamate was always 1 mM. For the steroid sulfatases the substrate *para*-nitrophenol sulfate was at a concentration of 1 mM, whilst for carbohydrate sulfatases substrate concentration was 1 μM. NA means no activity when compared to the control. NT means not tested. Numbers in the left most column refer to the arylsulfamate in Figure 1c added to the reaction; control represents substrate only whilst WT is substrate plus enzyme without an arylsulfamate inhibitor. Rates are μM product/min.

Supplementary table 2. Kinetic rates of sulfatase when sulfatases were pre-incubated with arylsulfamates then assayed.

The concentration of arylsulfamate used for preincubation was always 1 mM and the enzyme concentration 25 - 50 μM. For the steroid sulfatases the substrate *para*-nitrophenol sulfate was at a concentration of 1 mM, whilst for carbohydrate sulfatases substrate concentration was 1 μM. NA means no activity when compared to the control. NT means not tested. Numbers in the left most column refer to the arylsulfamate in Figure 1c added to the reaction; control represents substrate only whilst WT is substrate plus enzyme without an arylsulfamate inhibitor. Rates are μM product/min.

Supplementary table 3. Kinetic rate of BT1636$^{3S\text{-}Gal}$ assayed against BODIPY labelled 3S-galactose sulfate with various carbohydrate and aryl- sulfonates/sulfamates.

BT1636$^{3S\text{-}Gal}$ was assayed against 1 μM BODIPY labelled 3S-galactose with 1 mM of **with** various carbohydrate and aryl- sulfonates/sulfamates the assay, or after BT1636$^{3S\text{-}Gal}$ had been incubated with 1 mM of the same compounds; concentrations of 300 nM and 285 nM of BT1636$^{3S\text{-}Gal}$ were used, respectively. The assay was performed in 100 mM of MES pH 6.0 with 5 % DMSO, 150 mM NaCl, 0.02% (v/v) Brij-35 and 5 mM $CaCl_2$. Assays were performed in triplicate. Rates are fluorescence of product/min. The values in green indicate 5 % was not present.

Supplementary table 4. Thermal melt values of BT1636$^{3S\text{-}Gal}$ in the absence and presence of various carbohydrate and aryl- sulfonates/sulfamates.

BT1636$^{3S\text{-}Gal}$, at a concentration of 5 μM, was incubated with no compound, and varying concentrations and the effect on its melting temperature monitored, a shift in melting temperature is indicative of an interaction. The assays were performed in 100 mM of BTP pH 7.0 with 5 % DMSO, 150 mM NaCl. Assays were performed in triplicate.

Supplementary table 5. Conservation of BT4322 in select Bacteroidota species of the human gut microbiota.

BT4322 was used as the query against the genomes of the 10 species from Figure 5 against which arylsulfamates 3 and 12 were tested. The BLASTp search was done using default settings with the blastp algorithm accessed through the NCBI interface.

References:

1 Barbeyron, T. *et al.* Matching the Diversity of Sulfated Biomolecules: Creation of a Classification Database for Sulfatases Reflecting Their Substrate Specificity. *PloS one* **11**, e0164846 (2016). https://doi.org:10.1371/journal.pone.0164846

2 Stam, M. *et al.* SulfAtlas, the sulfatase database: state of the art and new developments. *Nucleic Acids Res* **51**, D647-D653 (2023). https://doi.org:10.1093/nar/gkac977

3 Hanson, S. R., Best, M. D. & Wong, C. H. Sulfatases: structure, mechanism, biological activity, inhibition, and synthetic utility. *Angewandte Chemie* **43**, 5736-5763 (2004). https://doi.org:10.1002/anie.200300632

1167 4 Luis, A. S., Yates, E. A. & Cartmell, A. Functions and specificity of bacterial
1168 carbohydrate sulfatases targeting host glycans. *Essays Biochem* **67**, 429-442 (2023).
1169 https://doi.org:10.1042/EBC20220120
1170 5 Stanway, S. J. *et al.* Steroid sulfatase: a new target for the endocrine therapy of
1171 breast cancer. *Oncologist* **12**, 370-374 (2007).
1172 https://doi.org:10.1634/theoncologist.12-4-370
1173 6 Reed, M. J., Purohit, A., Woo, L. W., Newman, S. P. & Potter, B. V. Steroid sulfatase:
1174 molecular biology, regulation, and inhibition. *Endocr Rev* **26**, 171-202 (2005).
1175 https://doi.org:10.1210/er.2004-0003
1176 7 Armstrong, C. M. *et al.* Steroid Sulfatase Stimulates Intracrine Androgen Synthesis
1177 and is a Therapeutic Target for Advanced Prostate Cancer. *Clin Cancer Res* **26**, 6064-
1178 6074 (2020). https://doi.org:10.1158/1078-0432.CCR-20-1682
1179 8 Ai, X. *et al.* Substrate specificity and domain functions of extracellular heparan
1180 sulfate 6-O-endosulfatases, QSulf1 and QSulf2. *The Journal of biological chemistry*
1181 **281**, 4969-4976 (2006). https://doi.org:10.1074/jbc.M511902200
1182 9 Lamanna, W. C. *et al.* Heparan sulfate 6-O-endosulfatases: discrete in vivo activities
1183 and functional co-operativity. *The Biochemical journal* **400**, 63-73 (2006).
1184 https://doi.org:10.1042/BJ20060848
1185 10 Hur, K. *et al.* Up-regulated expression of sulfatases (SULF1 and SULF2) as prognostic
1186 and metastasis predictive markers in human gastric cancer. *The Journal of pathology*
1187 **228**, 88-98 (2012). https://doi.org:10.1002/path.4055
1188 11 Vicente, C. M., Lima, M. A., Yates, E. A., Nader, H. B. & Toma, L. Enhanced
1189 tumorigenic potential of colorectal cancer cells by extracellular sulfatases. *Molecular*
1190 *cancer research : MCR* **13**, 510-523 (2015). https://doi.org:10.1158/1541-7786.MCR-
1191 14-0372
1192 12 Van der Sluis, M. *et al.* Muc2-deficient mice spontaneously develop colitis, indicating
1193 that MUC2 is critical for colonic protection. *Gastroenterology* **131**, 117-129 (2006).
1194 https://doi.org:10.1053/j.gastro.2006.04.020
1195 13 Tsai, H. H., Dwarakanath, A. D., Hart, C. A., Milton, J. D. & Rhodes, J. M. Increased
1196 faecal mucin sulphatase activity in ulcerative colitis: a potential target for treatment.
1197 *Gut* **36**, 570-576 (1995). https://doi.org:10.1136/gut.36.4.570
1198 14 Luis, A. S. *et al.* A single sulfatase is required to access colonic mucin by a gut
1199 bacterium. *Nature* **598**, 332-337 (2021). https://doi.org:10.1038/s41586-021-03967-
1200 5
1201 15 Velcich, A. *et al.* Colorectal cancer in mice genetically deficient in the mucin Muc2.
1202 *Science* **295**, 1726-1729 (2002). https://doi.org:10.1126/science.1069094
1203 16 Bloom, S. M. *et al.* Commensal Bacteroides species induce colitis in host-genotype-
1204 specific fashion in a mouse model of inflammatory bowel disease. *Cell host &*
1205 *microbe* **9**, 390-403 (2011). https://doi.org:10.1016/j.chom.2011.04.009
1206 17 Hickey, C. A. *et al.* Colitogenic Bacteroides thetaiotaomicron Antigens Access Host
1207 Immune Cells in a Sulfatase-Dependent Manner via Outer Membrane Vesicles. *Cell*
1208 *host & microbe* **17**, 672-680 (2015). https://doi.org:10.1016/j.chom.2015.04.002
1209 18 Bojarova, P. & Williams, S. J. Aryl sulfamates are broad spectrum inactivators of
1210 sulfatases: effects on sulfatases from various sources. *Bioorg Med Chem Lett* **19**,
1211 477-480 (2009). https://doi.org:10.1016/j.bmcl.2008.11.059

1212 19 Thomas, M. P. & Potter, B. V. Discovery and Development of the Aryl O-Sulfamate
1213 Pharmacophore for Oncology and Women's Health. *Journal of medicinal chemistry*
1214 **58**, 7634-7658 (2015). https://doi.org:10.1021/acs.jmedchem.5b00386
1215 20 Luis, A. S. *et al.* Sulfated glycan recognition by carbohydrate sulfatases of the human
1216 gut microbiota. *Nature chemical biology* **18**, 841-849 (2022).
1217 https://doi.org:10.1038/s41589-022-01039-x
1218 21 Potter, B. V. L. SULFATION PATHWAYS: Steroid sulphatase inhibition via aryl
1219 sulphamates: clinical progress, mechanism and future prospects. *J Mol Endocrinol*
1220 **61**, T233-T252 (2018). https://doi.org:10.1530/JME-18-0045
1221 22 Palmieri, C., Januszewski, A., Stanway, S. & Coombes, R. C. Irosustat: a first-
1222 generation steroid sulfatase inhibitor in breast cancer. *Expert Rev Anticancer Ther*
1223 **11**, 179-183 (2011). https://doi.org:10.1586/era.10.201
1224 23 Elger, W. *et al.* Estrogen sulfamates: a new approach to oral estrogen therapy.
1225 *Reprod Fertil Dev* **13**, 297-305 (2001). https://doi.org:10.1071/rd01029
1226 24 Woo, L. W. *et al.* Steroidal and nonsteroidal sulfamates as potent inhibitors of
1227 steroid sulfatase. *Journal of medicinal chemistry* **41**, 1068-1083 (1998).
1228 https://doi.org:10.1021/jm970527v
1229 25 Palmieri, C. *et al.* IRIS study: a phase II study of the steroid sulfatase inhibitor
1230 Irosustat when added to an aromatase inhibitor in ER-positive breast cancer
1231 patients. *Breast Cancer Res Treat* **165**, 343-353 (2017).
1232 https://doi.org:10.1007/s10549-017-4328-z
1233 26 Palmieri, C. *et al.* IPET study: an FLT-PET window study to assess the activity of the
1234 steroid sulfatase inhibitor irosustat in early breast cancer. *Breast Cancer Res Treat*
1235 **166**, 527-539 (2017). https://doi.org:10.1007/s10549-017-4427-x
1236 27 Pautier, P. *et al.* A Phase 2, Randomized, Open-Label Study of Irosustat Versus
1237 Megestrol Acetate in Advanced Endometrial Cancer. *Int J Gynecol Cancer* **27**, 258-
1238 266 (2017). https://doi.org:10.1097/IGC.0000000000000862
1239 28 Thomas, M. P. & Potter, B. V. Estrogen O-sulfamates and their analogues: Clinical
1240 steroid sulfatase inhibitors with broad potential. *J Steroid Biochem Mol Biol* **153**,
1241 160-169 (2015). https://doi.org:10.1016/j.jsbmb.2015.03.012
1242 29 Byrne, D. P., London, J. A., Eyers, P. A., Yates, E. A. & Cartmell, A. Mobility shift-based
1243 electrophoresis coupled with fluorescent detection enables real-time enzyme
1244 analysis of carbohydrate sulfatase activity. *The Biochemical journal* **478**, 735-748
1245 (2021). https://doi.org:10.1042/BCJ20200952
1246 30 Schelwies, M. *et al.* Glucosamine-6-sulfamate analogues of heparan sulfate as
1247 inhibitors of endosulfatases. *Chembiochem : a European journal of chemical biology*
1248 **11**, 2393-2397 (2010). https://doi.org:10.1002/cbic.201000401
1249 31 Chiu, L. T. *et al.* Trisaccharide Sulfate and Its Sulfonamide as an Effective Substrate
1250 and Inhibitor of Human Endo-O-sulfatase-1. *J Am Chem Soc* **142**, 5282-5292 (2020).
1251 https://doi.org:10.1021/jacs.0c00005
1252 32 Tseng, K. Y., Tzeng, Z. H., Cheng, T. R., Liang, P. H. & Hung, S. C. Design and Synthesis
1253 of 1-O- and 6'-C-Modified Heparan Sulfate Trisaccharides as Human Endo-6-O-
1254 Sulfatase 1 Inhibitors. *Front Chem* **10**, 947475 (2022).
1255 https://doi.org:10.3389/fchem.2022.947475
1256 33 Berteau, O., Guillot, A., Benjdia, A. & Rabot, S. A new type of bacterial sulfatase
1257 reveals a novel maturation pathway in prokaryotes. *The Journal of biological*
1258 *chemistry* **281**, 22464-22470 (2006). https://doi.org:10.1074/jbc.M602504200

1259 34 Benjdia, A., Martens, E. C., Gordon, J. I. & Berteau, O. Sulfatases and a radical S-
1260 adenosyl-L-methionine (AdoMet) enzyme are key for mucosal foraging and fitness of
1261 the prominent human gut symbiont, Bacteroides thetaiotaomicron. *The Journal of*
1262 *biological chemistry* **286**, 25973-25982 (2011).
1263 https://doi.org:10.1074/jbc.M111.228841
1264 35 Tintelnot, J. *et al.* Microbiota-derived 3-IAA influences chemotherapy efficacy in
1265 pancreatic cancer. *Nature* **615**, 168-174 (2023). https://doi.org:10.1038/s41586-023-
1266 05728-y
1267 36 Maier, L. *et al.* Extensive impact of non-antibiotic drugs on human gut bacteria.
1268 *Nature* **555**, 623-628 (2018). https://doi.org:10.1038/nature25979
1269 37 Ricaurte, D. *et al.* High-throughput transcriptomics of 409 bacteria-drug pairs reveals
1270 drivers of gut microbiota perturbation. *Nature microbiology* (2024).
1271 https://doi.org:10.1038/s41564-023-01581-x
1272 38 Copeland, R. A. Evaluation of enzyme inhibitors in drug discovery. A guide for
1273 medicinal chemists and pharmacologists. *Methods Biochem Anal* **46**, 1-265 (2005).
1274 39 Mateus, A. *et al.* Thermal proteome profiling in bacteria: probing protein state in
1275 vivo. *Mol Syst Biol* **14**, e8242 (2018). https://doi.org:10.15252/msb.20188242
1276 40 Martens, E. C. *et al.* Recognition and degradation of plant cell wall polysaccharides
1277 by two human gut symbionts. *PLoS biology* **9**, e1001221 (2011).
1278 https://doi.org:10.1371/journal.pbio.1001221
1279 41 Arjes, H. A. *et al.* Construction and characterization of a genome-scale ordered
1280 mutant collection of Bacteroides thetaiotaomicron. *BMC Biol* **20**, 285 (2022).
1281 https://doi.org:10.1186/s12915-022-01481-2
1282 42 Jaeger, T., Arsic, M. & Mayer, C. Scission of the lactyl ether bond of N-acetylmuramic
1283 acid by Escherichia coli "etherase". *The Journal of biological chemistry* **280**, 30100-
1284 30106 (2005). https://doi.org:10.1074/jbc.M502208200
1285 43 Jerga, A., Lu, Y. J., Schujman, G. E., de Mendoza, D. & Rock, C. O. Identification of a
1286 soluble diacylglycerol kinase required for lipoteichoic acid production in Bacillus
1287 subtilis. *The Journal of biological chemistry* **282**, 21738-21745 (2007).
1288 https://doi.org:10.1074/jbc.M703536200
1289 44 Miller, D. J., Jerga, A., Rock, C. O. & White, S. W. Analysis of the Staphylococcus
1290 aureus DgkB structure reveals a common catalytic mechanism for the soluble
1291 diacylglycerol kinases. *Structure* **16**, 1036-1046 (2008).
1292 https://doi.org:10.1016/j.str.2008.03.019
1293 45 Wang, Z. *et al.* Molecular basis of sphingosine kinase 1 substrate recognition and
1294 catalysis. *Structure* **21**, 798-809 (2013). https://doi.org:10.1016/j.str.2013.02.025
1295 46 Cartmell, A. *et al.* How members of the human gut microbiota overcome the
1296 sulfation problem posed by glycosaminoglycans. *Proceedings of the National*
1297 *Academy of Sciences of the United States of America* **114**, 7037-7042 (2017).
1298 https://doi.org:10.1073/pnas.1704367114
1299 47 Ndeh, D. *et al.* Metabolism of multiple glycosaminoglycans by Bacteroides
1300 thetaiotaomicron is orchestrated by a versatile core genetic locus. *Nature*
1301 *communications* **11**, 646 (2020). https://doi.org:10.1038/s41467-020-14509-4
1302 48 Badola, P. & Sanders, C. R., 2nd. Escherichia coli diacylglycerol kinase is an
1303 evolutionarily optimized membrane enzyme and catalyzes direct phosphoryl
1304 transfer. *The Journal of biological chemistry* **272**, 24176-24182 (1997).
1305 https://doi.org:10.1074/jbc.272.39.24176

49 An, D., Na, C., Bielawski, J., Hannun, Y. A. & Kasper, D. L. Membrane sphingolipids as essential molecular signals for Bacteroides survival in the intestine. *Proceedings of the National Academy of Sciences of the United States of America* **108 Suppl 1**, 4666-4671 (2011). https://doi.org:10.1073/pnas.1001501107

50 Moggridge, S., Sorensen, P. H., Morin, G. B. & Hughes, C. S. Extending the Compatibility of the SP3 Paramagnetic Bead Processing Approach for Proteomics. *J Proteome Res* **17**, 1730-1740 (2018). https://doi.org:10.1021/acs.jproteome.7b00913

51 Smits, A. H. *et al.* Biological plasticity rescues target activity in CRISPR knock outs. *Nat Methods* **16**, 1087-1093 (2019). https://doi.org:10.1038/s41592-019-0614-5

52 Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D. & Nesvizhskii, A. I. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat Methods* **14**, 513-520 (2017). https://doi.org:10.1038/nmeth.4256

53 da Veiga Leprevost, F. *et al.* Philosopher: a versatile toolkit for shotgun proteomics data analysis. *Nat Methods* **17**, 869-870 (2020). https://doi.org:10.1038/s41592-020-0912-y

54 Franken, H. *et al.* Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nat Protoc* **10**, 1567-1593 (2015). https://doi.org:10.1038/nprot.2015.101

55 Grosdidier, A., Zoete, V. & Michielin, O. Fast docking using the CHARMM force field with EADock DSS. *J Comput Chem* **32**, 2149-2159 (2011). https://doi.org:10.1002/jcc.21797

56 Lebedev, A. A. *et al.* JLigand: a graphical tool for the CCP4 template-restraint library. *Acta crystallographica. Section D, Biological crystallography* **68**, 431-440 (2012). https://doi.org:10.1107/S090744491200251X

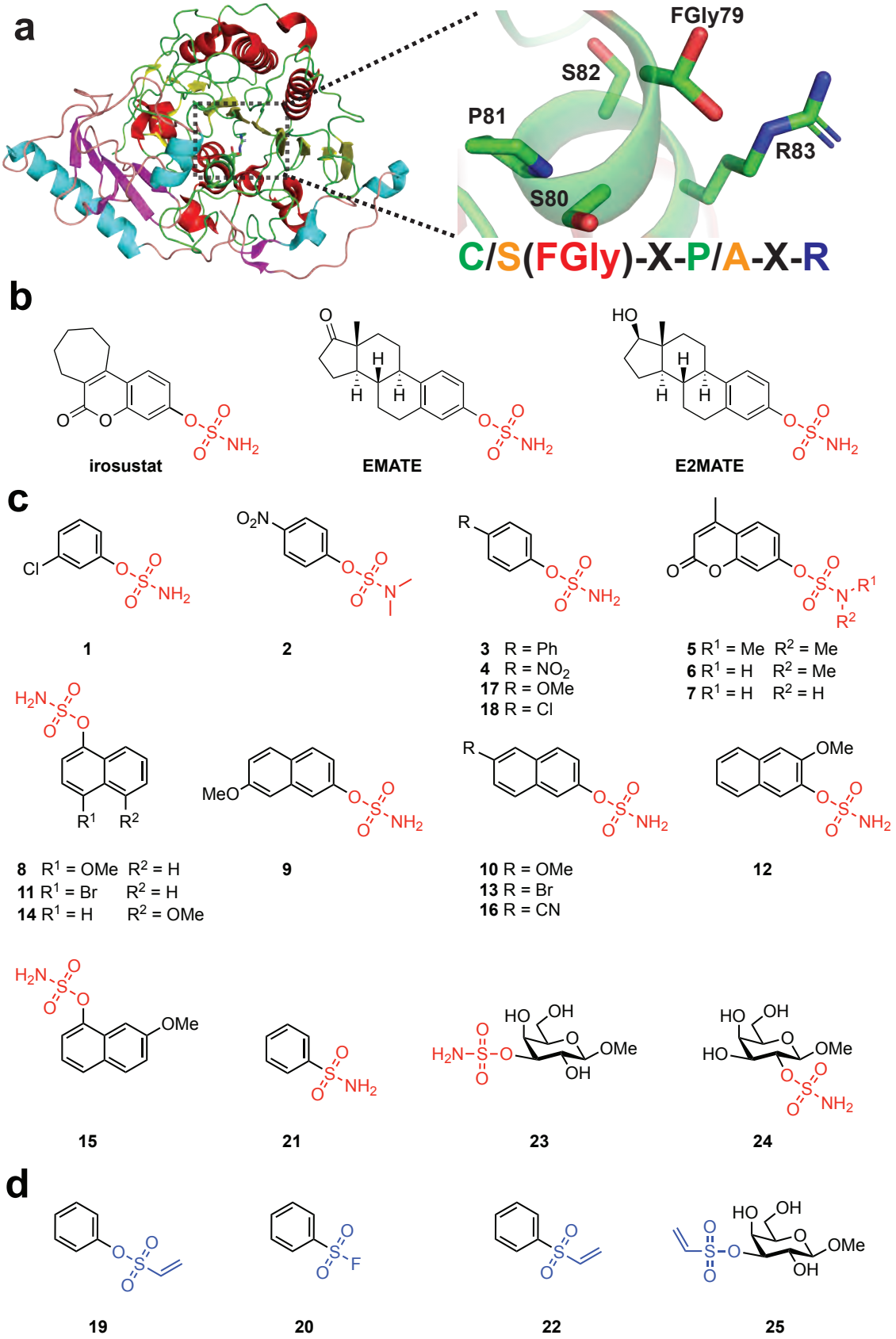


Figure 1 General example of the conserved S1 sulfatase fold and consensus active site sequence. **a.** The 3-D structure of the human lysosomal sulfatase GALNS (PDB:4FDJ), the N-terminal domain is coloured red and yellow for alpha helices and beta sheets, with the C-terminal subdomain coloured cyan and magenta for alpha helices and beta sheets; a zoom in of the consensus sequence of GALNS with a converted formylglycine, and with the general consensus pictured to the right; **b**. Structure of phase I and II arylsulfamate drugs developed to inhibit STS **c.** Chemical structures of aryl- and carbohydrate sulfamate inhibitors used in the study. **d**. Chemical structures of phenylvinyl sulfone, phenylsulfonyl fluoride, phenylvinyl sulfonate, and carbohydrate vinyl sulfone.

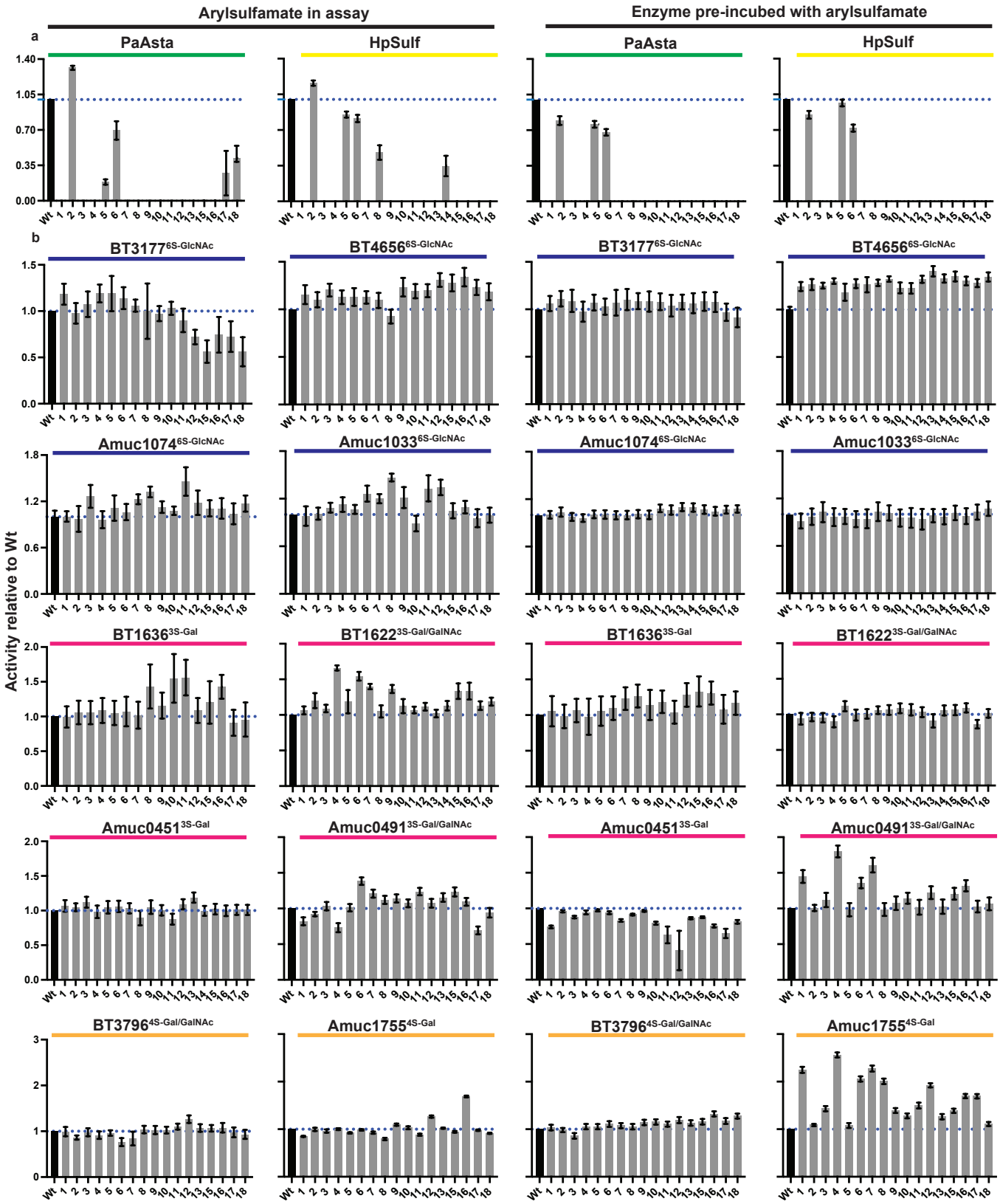


Figure 2 The effectiveness of arylsulfamate inhibitors against S1 steroid and carbohydrate sulfatases. **a.** The enzymes PaAsta and HpSulf, belonging to S1_4 (green bar) and S1_2 (yellow bar), respectively, assayed against the panel of arylsulfamates (1-19 listed in Figure 2). **b**. The S1_11 sulfatases (BT31776S-GlcNAc, BT46566S-GlcNAc, Amuc_10336S-GlcNAc, Amuc_10746S-GlcNAc; blue bar above), the S1_20 sulfatases (BT16223S-Gal/GalNAc, BT16363S-Gal, Amuc_04513S-Gal, Amuc_04913S-Gal; pink bar above), and the S1_16 enzymes (BT37964S-Gal/GalNAc and BT17554S-Gal; orange bar above) against the panel of arylsulfamates (1-19 listed in Figure 2). For both a and b the left two panels display relative activity to wildtype protein with the arylsulfamate compound in the assay, whilst the left two panels are overnight preincubation of the enzyme with selected arylsulfamates then dilution of this preincubation mix into the reaction. The numbers indicate which arylsulfamates from Figure 2 were used and in all instances 1 mM of compound was deployed. Wt indicates reaction with enzyme and no arylsulfamate. Kinetic curves from which the data are derived are shown in supplemental figures S1-S24.

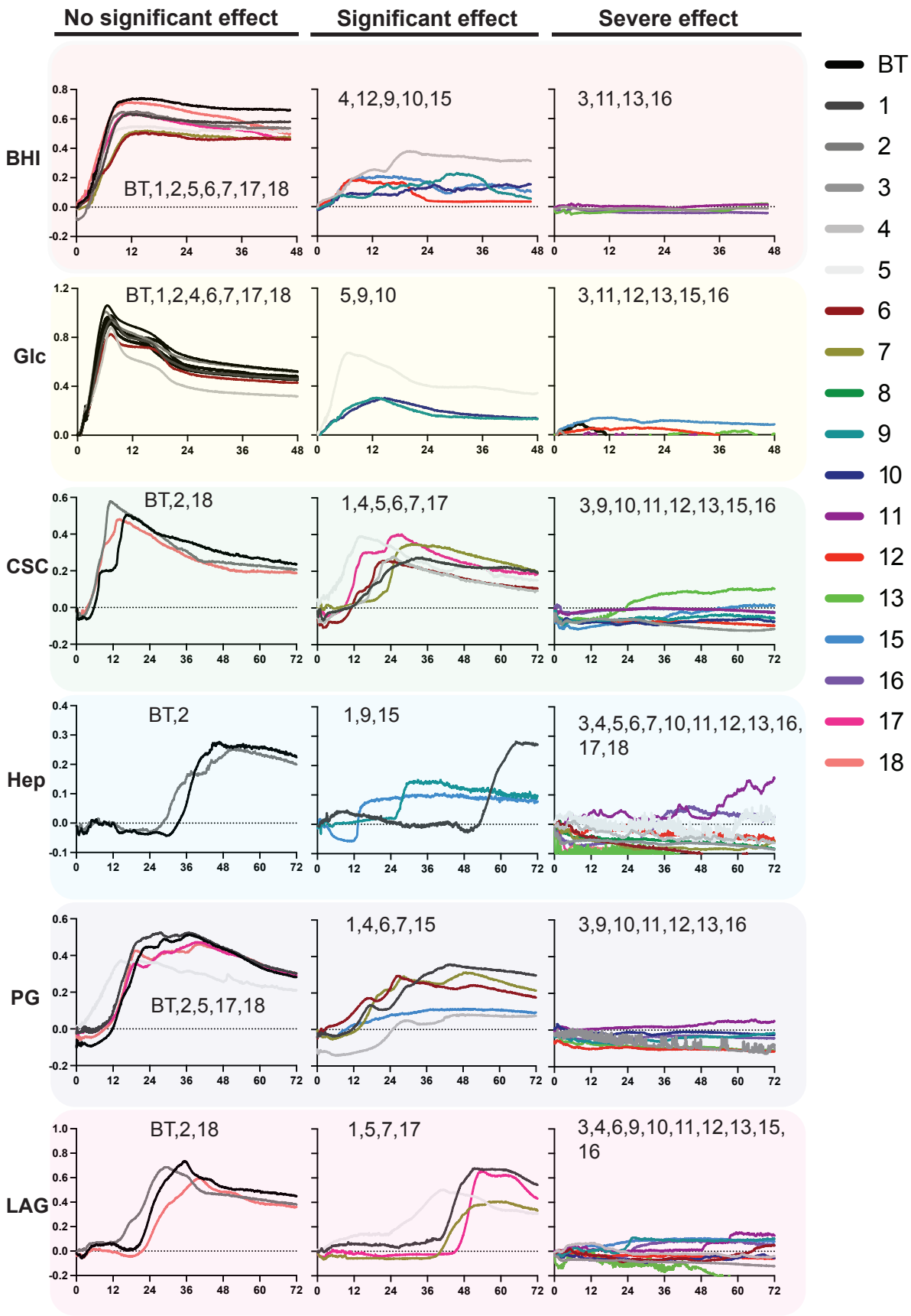


Figure 3 Growth of Bacteroides thetaiotaomicron on varied carbon sources in the presence of arylsulfamate inhibitors

Bacteroides thetaiotaomicron (B.theta) grown on varied carbon sources; two rich carbon sources, brain heart infusion media (BHI) and glucose (Glc); two sulfated host glycan carbon sources, chondroitin sulfate c (CSC) and Heparin (Hep), and two plant glycan carbon sources potato galactan (PG) and Larch arabinogalactan (LAG), in the presence of arylsulfamates inhibitors, labelled 1-19 (see Figure 2 for structures). Complex glycans and arylsulfamates inhibitors were used at a concentration of 5 mg/ml and 1 mM, respectively. All carbon source were utilised in a prepared minimal media except BHI which was purchased from sigma and simply dissolved in water. The traces are the mean of 3 independent growths but errors bars have been omitted for clarity; individual graphs with error bars are presented in supplemental figures S25-S30.

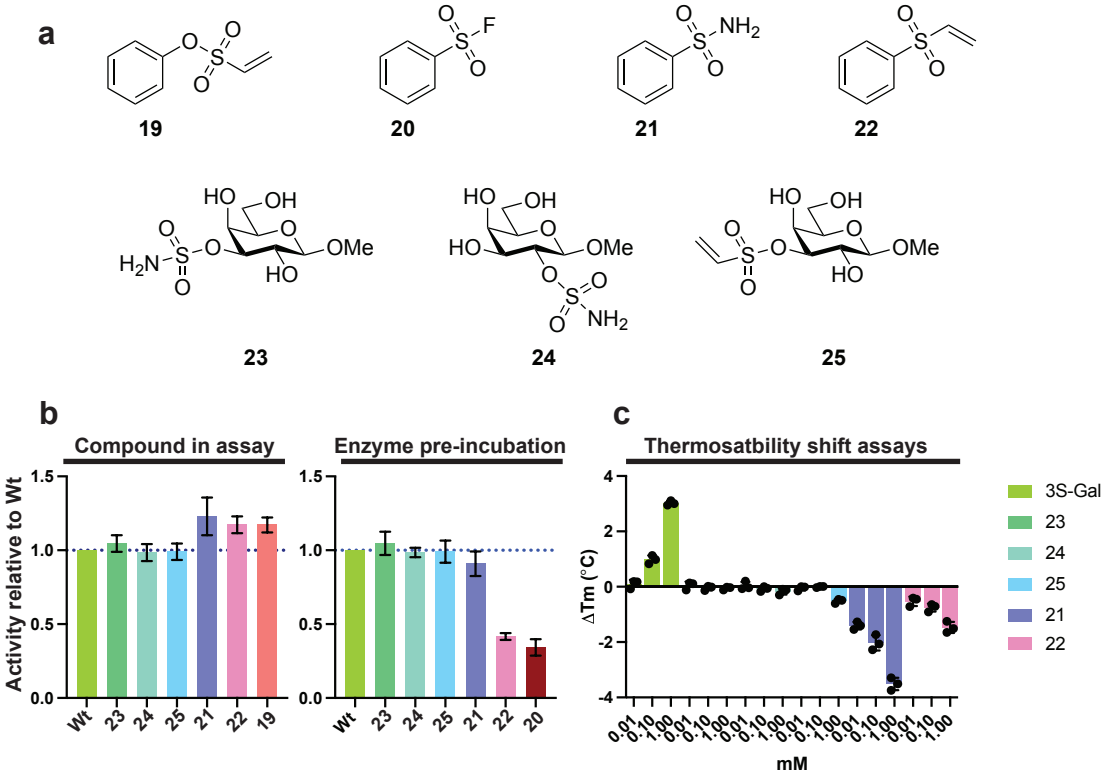


Figure 4. The effectiveness of carbohydrate sulfamate substrate mimics against the S1 carbohydrate sulfatase BT1636$^{3S\text{-}Gal}$.

a. Structure of the carbohydrate sulfonate/sulfamates and their aryl equivalents. **b.** The graphs show the effectiveness of carbohydrate sulfamates, carbohydrate sulfovinyl, and arylsulfonates against BT16363S-Gal; the left graph shows relative activity to wildtype with the compounds included in the assay, whilst the right graph show the relative activity after the enzyme had been preincubated with the compounds overnight then diluted into the reaction mixture;
c. Thermostability shift assays (TSA) using inactive BT1636$^{3S\text{-}Gal}$ against the substrate 3S-Gal, and the putative inhibitor structures carbohydrate sulfamates, carbohydrate sulfovinyl, and arylsulfonates. All compounds in b were used at a concentration of 1 mM.

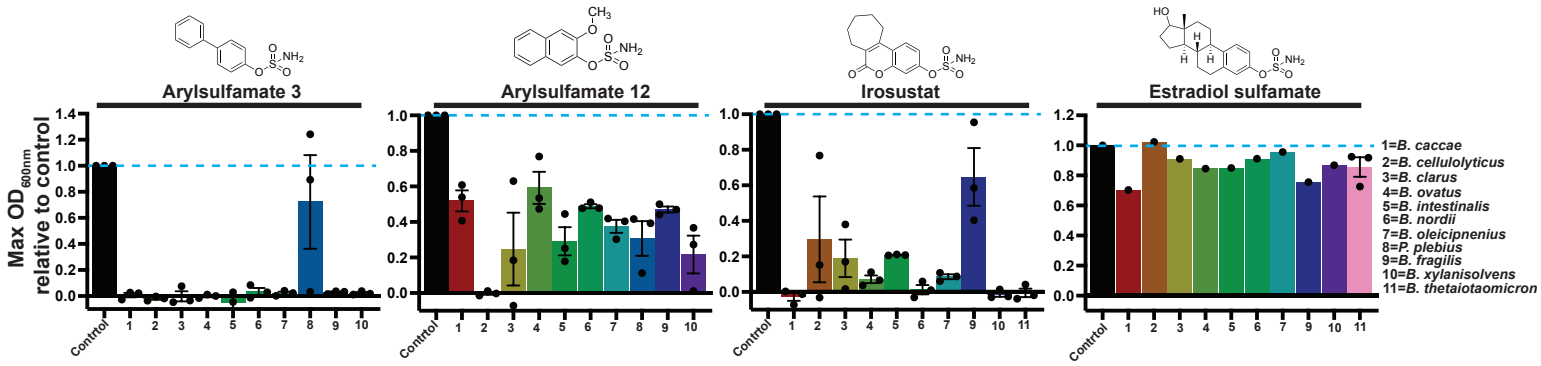


Figure 5. Arylsulfamates broadly affect Bacteroides spp. of the HGM
Bacteroidota species were grown in BHI media in the presence, or absence, of 1 mM of the listed arylsulfamate. The data plotted are the maximal OD_{600nm} observed for arylsulfamate treated growths relative to the untreated controls which were both ran in triplicate and randomly paired then averages derived. For all the estradiol growths a single experiment was conducted except for *B. theta* which was ran in triplicate; this was due to limited material.

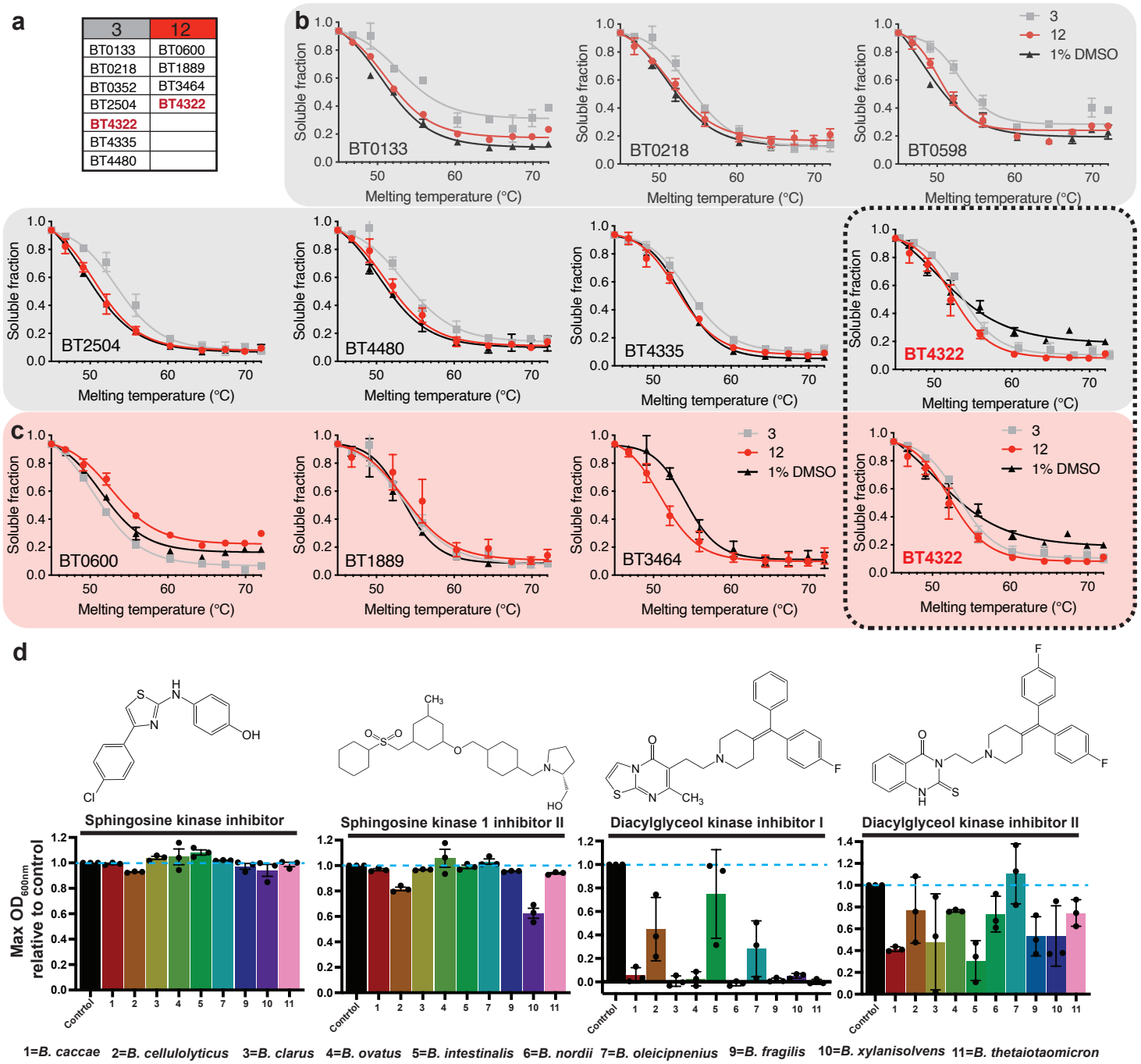
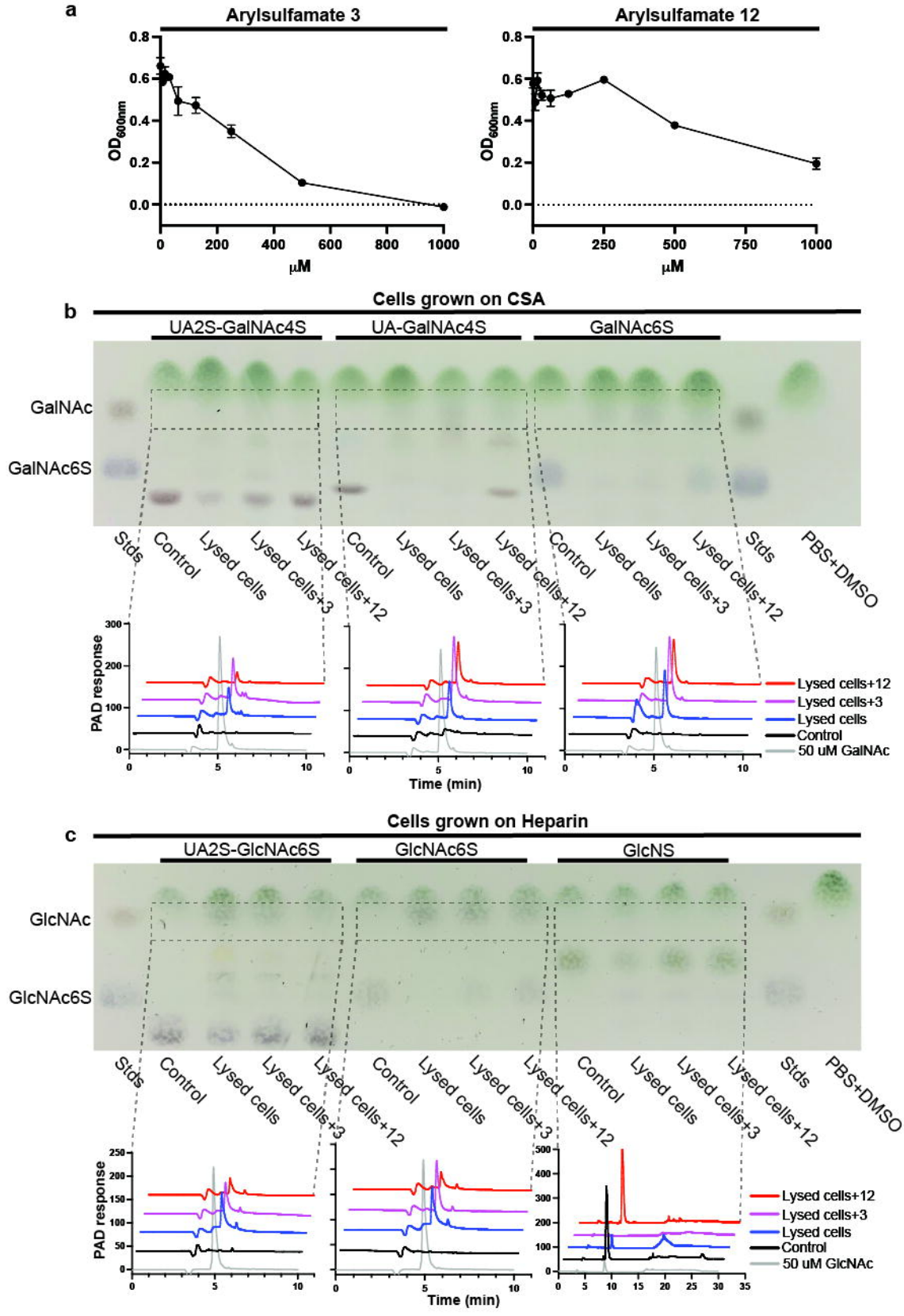


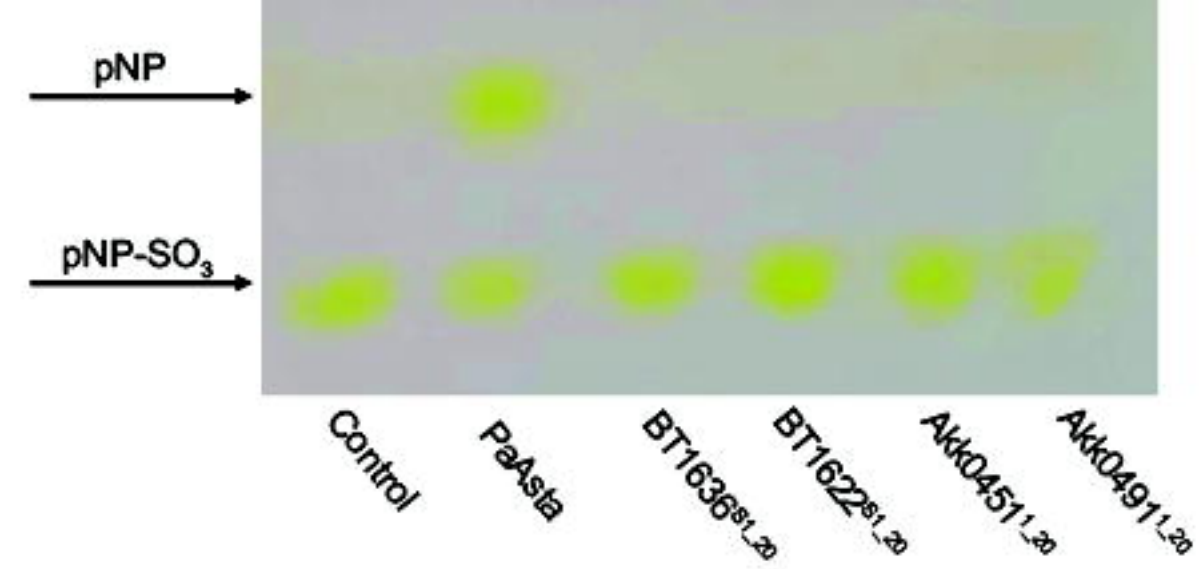
Figure 6. Global protein thermal stability reveals off target interations of arysulfamate inhibitors 3 and 12. **a.** Table of proteins which displayed an altered Tm in the presence of arylsulfamates 3 and 12. **b.** Proteins identified with a significance change in melting point upon inhibitor 3 treatment (n=2). **c.** Proteins identified with a significance change in melting point upon inhibitor 12 treatment (n=2). A dashed box indicates BT4322 is the only protein common across both groups. **d.** Bacteroides species were grown in BHI media in the presence, or absence, of 0.125 mM or 0.0625 mM diacylglycerol kinase inhibitor I and II, respectively. The data plotted are the maximal OD_{600nm} observed for inhibitor treated growths relative to the untreated controls which were both ran in triplicate and randomly paired then averages derived.

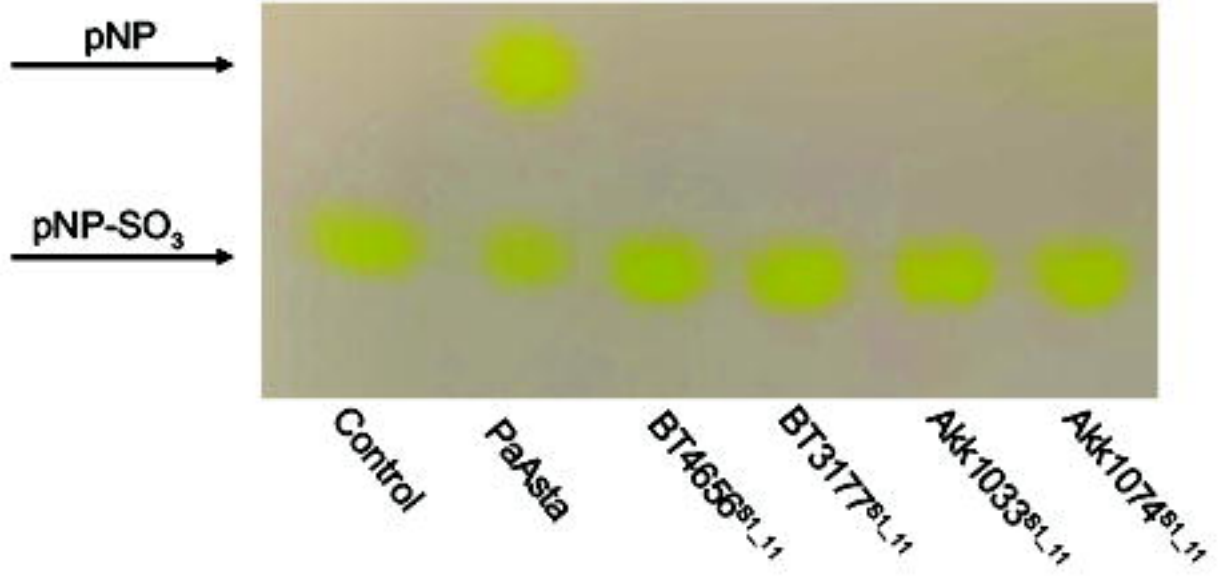


Extended data 1 Cell lysates of *Bacteroides thetaiotaomicron* incubated with two arylsulfamates inhibitors

a. The effects of varied concentrations of arylsulfamate 3 and 12 on the growth of *B. theta* in BHI media.

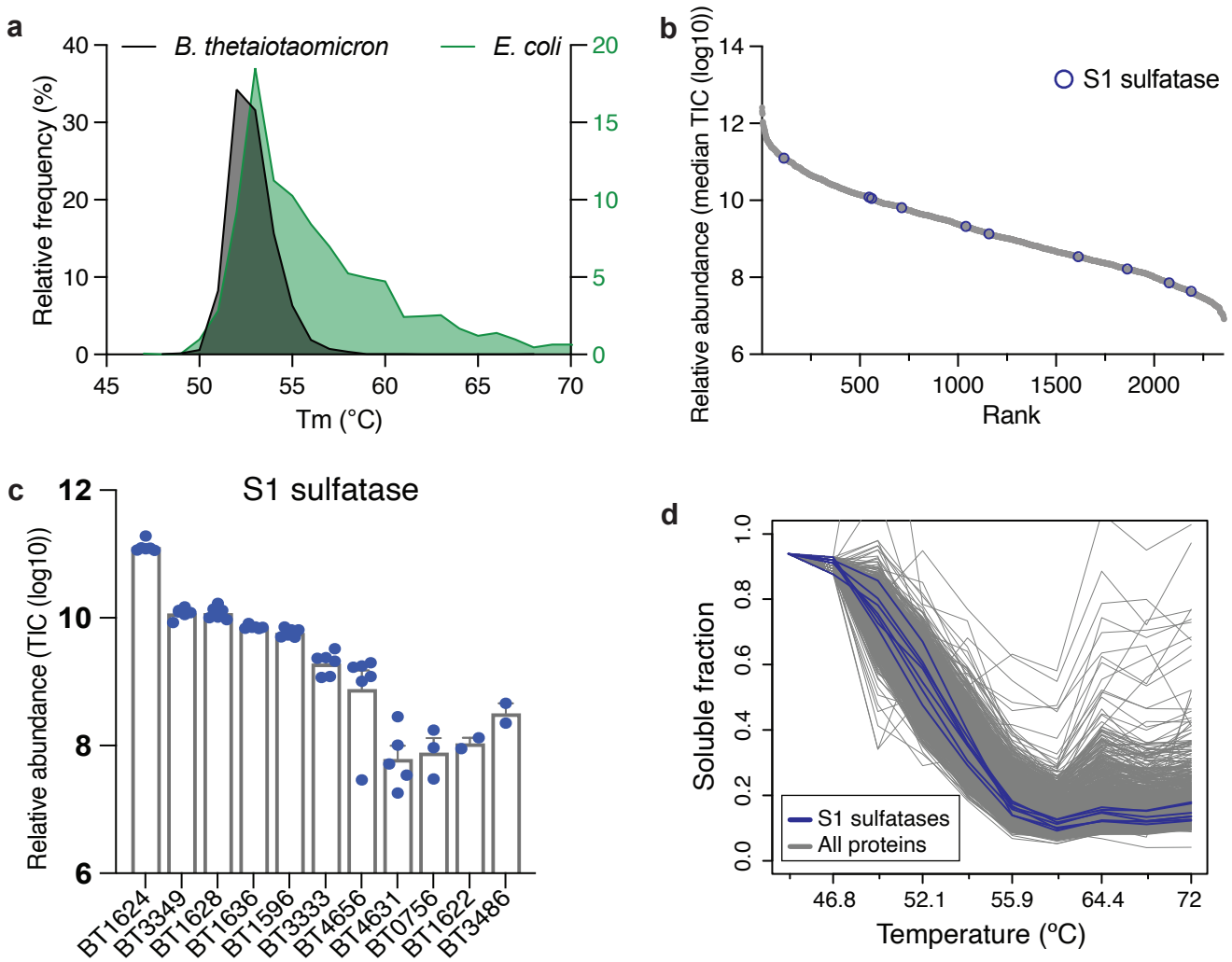
b. Thin layer chromatography (TLC) analysis of cell lysates from *B.theta* cells grown to mid-exponential phase on CSA as a carbon source, and incubated with arylsulfamates 4 and 13 (Figure 2), mixed 1:1 with the sulfated substrates: O2 sulfated uronic acid β1,3 linked to O4 sulfated N-acetylgalactosmaine (UA2S-GalNAc4S), uronic acid β1,3 linked to O4 sulfated N-acetylgalactosmaine (UA-GalNAc4S), and O6 sulfated N-acetylgalactosmaine (GalNAc6S). Below are high performance anion exchange chromatography (HPAEC) traces showing the production of GalNAc in all instances except the negative control; **c**. Thin layer chromatography (TLC) analysis of cell lysates from *B.theta* cells grown to mid-exponential phase on Heparin as a carbon source, and incubated with arylsulfamates 4 and 13 (Figure 2), mixed 1:1 with the sulfated substrates: O2 sulfated uronic acid α1,4 linked to O4 sulfated N-acetylgalactosmaine (UA2S-GalNAc4S), O6 sulfated N-acetylglucosmaine (GlcNAc6S), and N-sulfated glucosamine (GlcNS). Below are high performance anion exchange chromatography (HPAEC) traces showing the production of GlcNAc in all instances except the negative control. All reactions were in PBS, with 5 % DMSO, and incubated at 37ºC overnight.



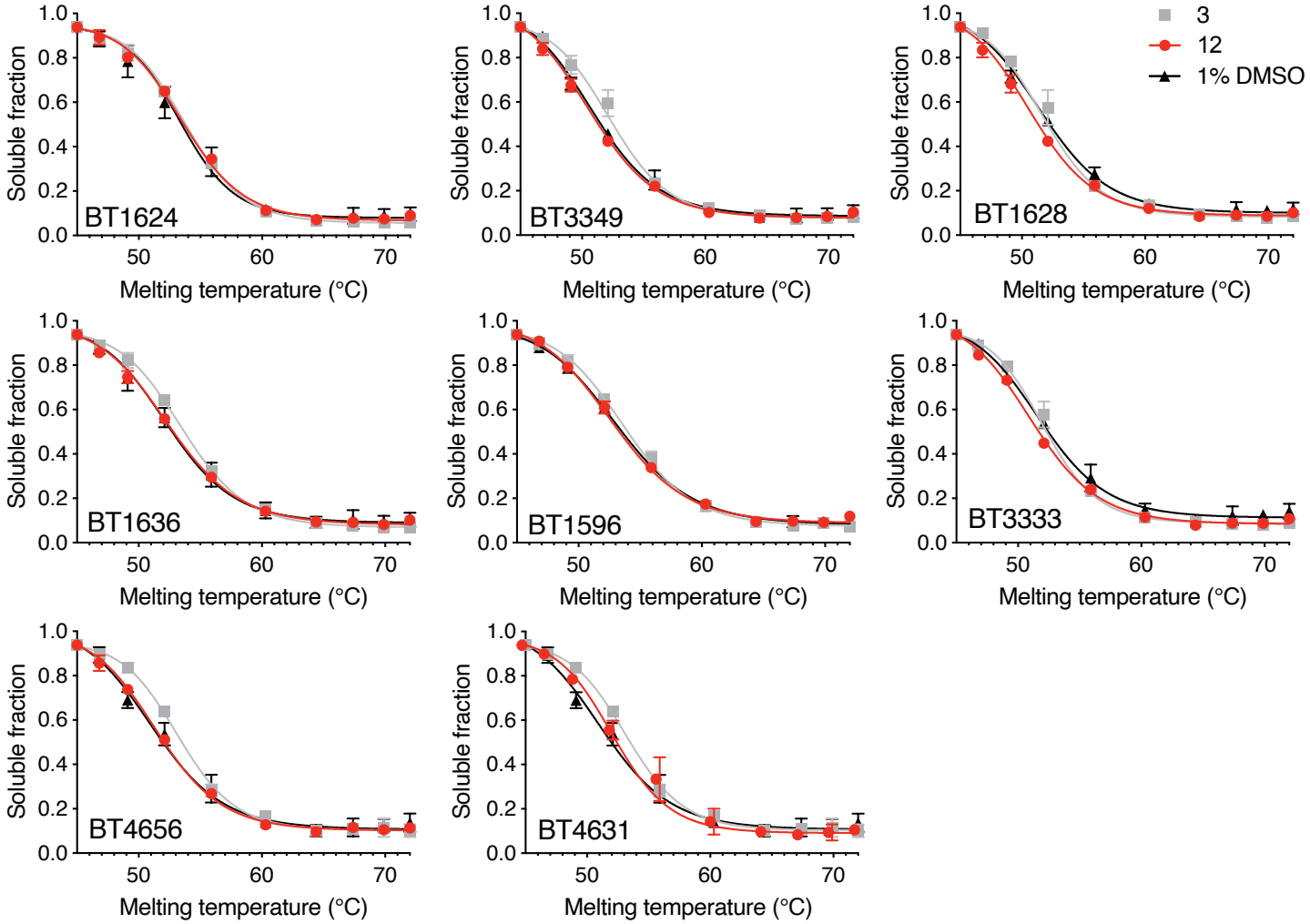


Extended data 2 Activity of carbohydrate sulfatases on para-nitrophenol sulfate.
Sulfatases were incubated with 5 mM para-nitrophenol sulfate over night in 100 mM BTP pH 7.0 with 150 mM NaCl and 5 mM CaCl2. Reactions were then ran on silica based thin layer chromatography using a resolving buffer of butanol:acetic acid:water (2:1:1) and develop by soaking the plates with 1 M NaOH and heating with a heat gun. It can be seen only the steroid sulfatase PaAsta shows the production of a para-nitrolphenol product.

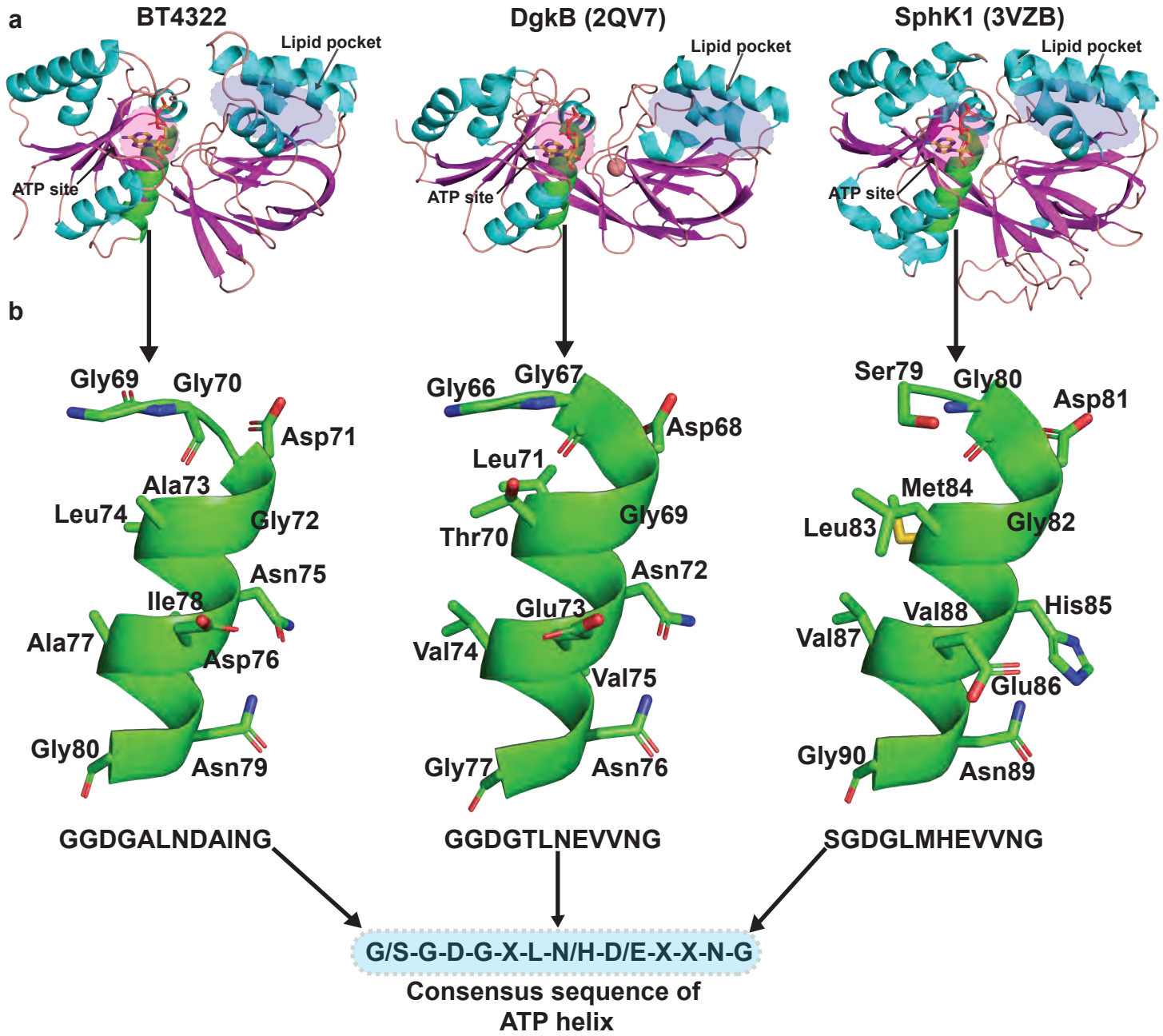
Extended data 3



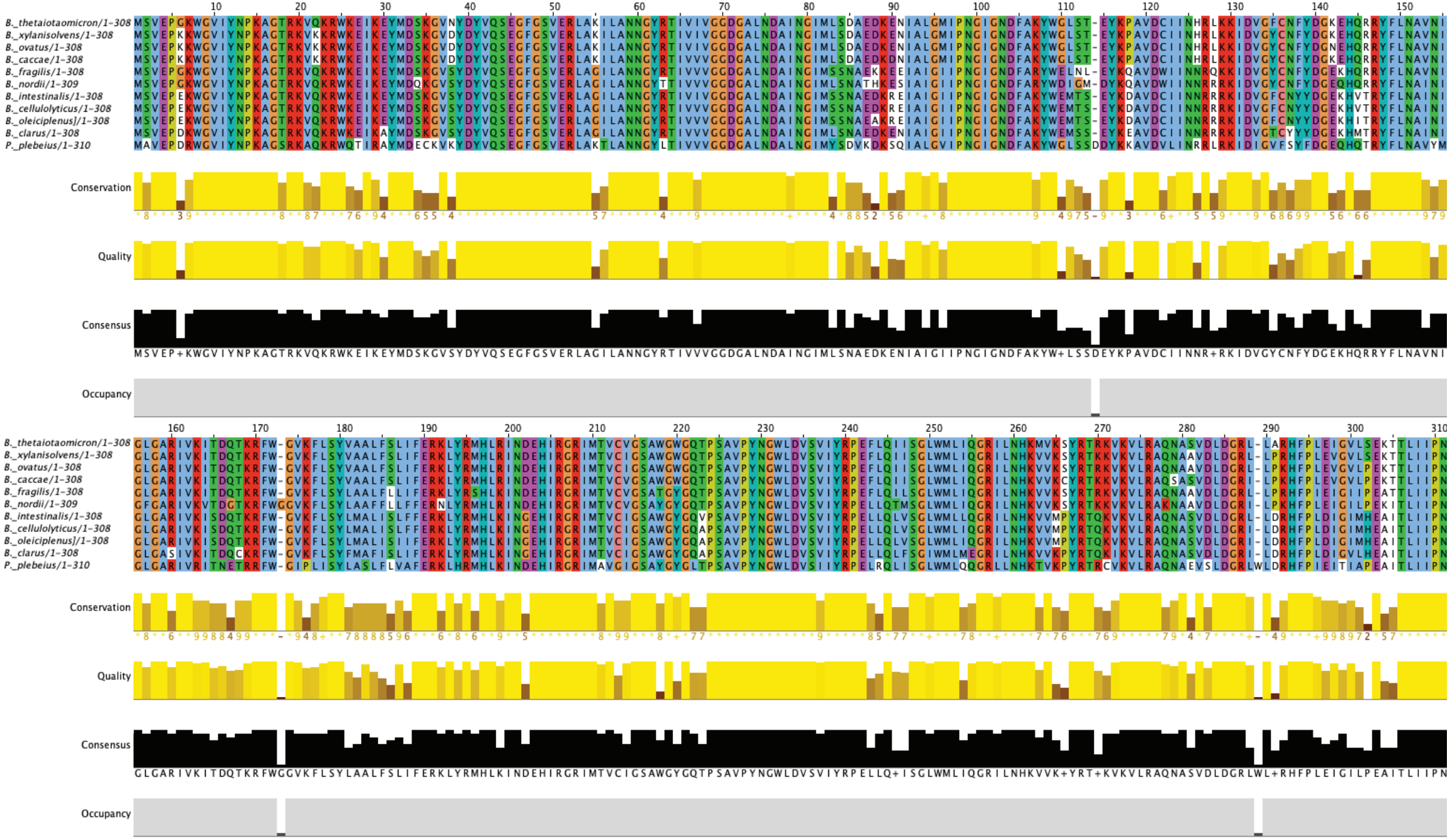
Extended data 3 Global features of the thermal proteome profiling data
a. Distribution of the melting temperature (Tm) of *Bacteroides thetaiotaomicron* VPI-5482 (*B. theta*) lysate proteome and the whole cell *E.coli* proteome. **b.** Ranked protein abundance plot of all identified proteins in the 1%DMSO treated controls, S1 sulfatase are highlighted. **c.** Relative abundance of 11 S1 sulfatases identified by mass spectrometry for *B. theta* cultured in the presence of chondroitin sulfate A (n=6). (b) Melting curve profiles for all identified *B. theta* proteins in control (1%DMSO) treated lysate. The thermal stability profiles for each individual S1 sulfatases are highlighted in blue. **d.** Melting curve profiles for all identified *B. theta* proteins in control (1%DMSO) treated lysate. The thermal stability profiles for each individual S1 sulfatases are highlighted in blue.



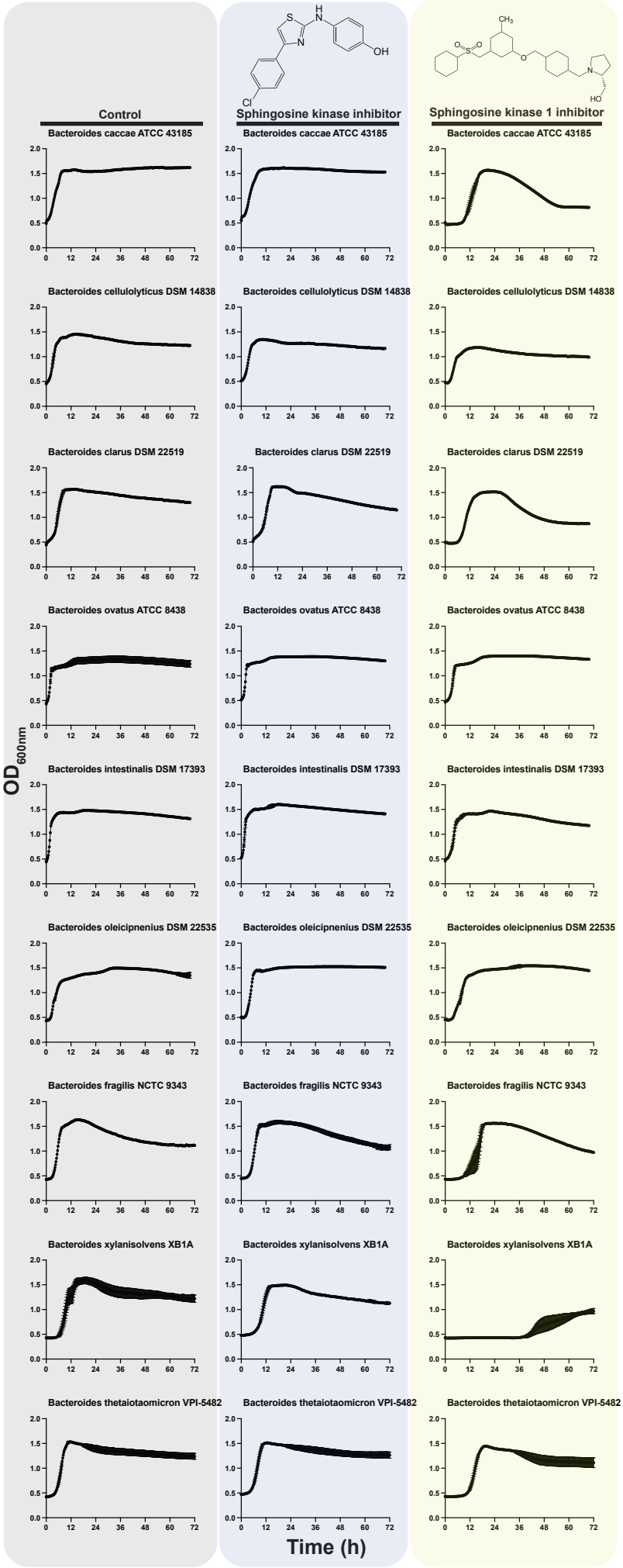
Extended data 4. Global protein thermal stability analysis confirms that arysulfamate inhibitors 3 and 12 do not directly interact with S1 sulfatases.
Individual melt curve for selected S1 sulfatase for control and arysulfamate inhibitors 3 and 12 treated lysates (n=2 per treatment group).



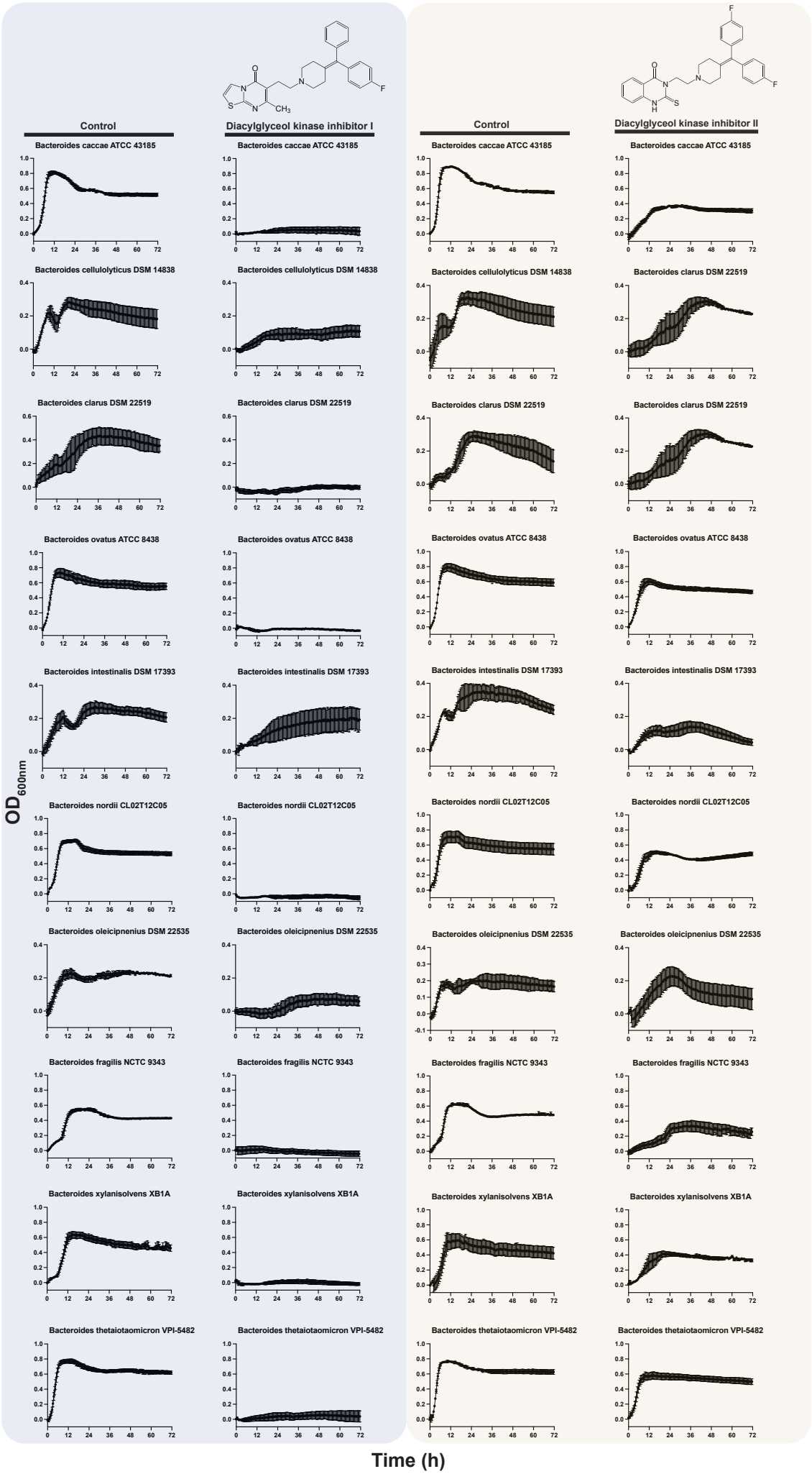
Extended data 5 Structure of BT4322, SaDgKB, and SphK1 and comparison of the ATP helix
a. Tertiary structures of BT4322, SaDgKB, and SphK1 shown as with a helices, b sheets, and loops coloured cyan, magenta, and pink. The 'ATP helix' is coloured in green. **b.** An expanded view of the ATP helix and its sequence.



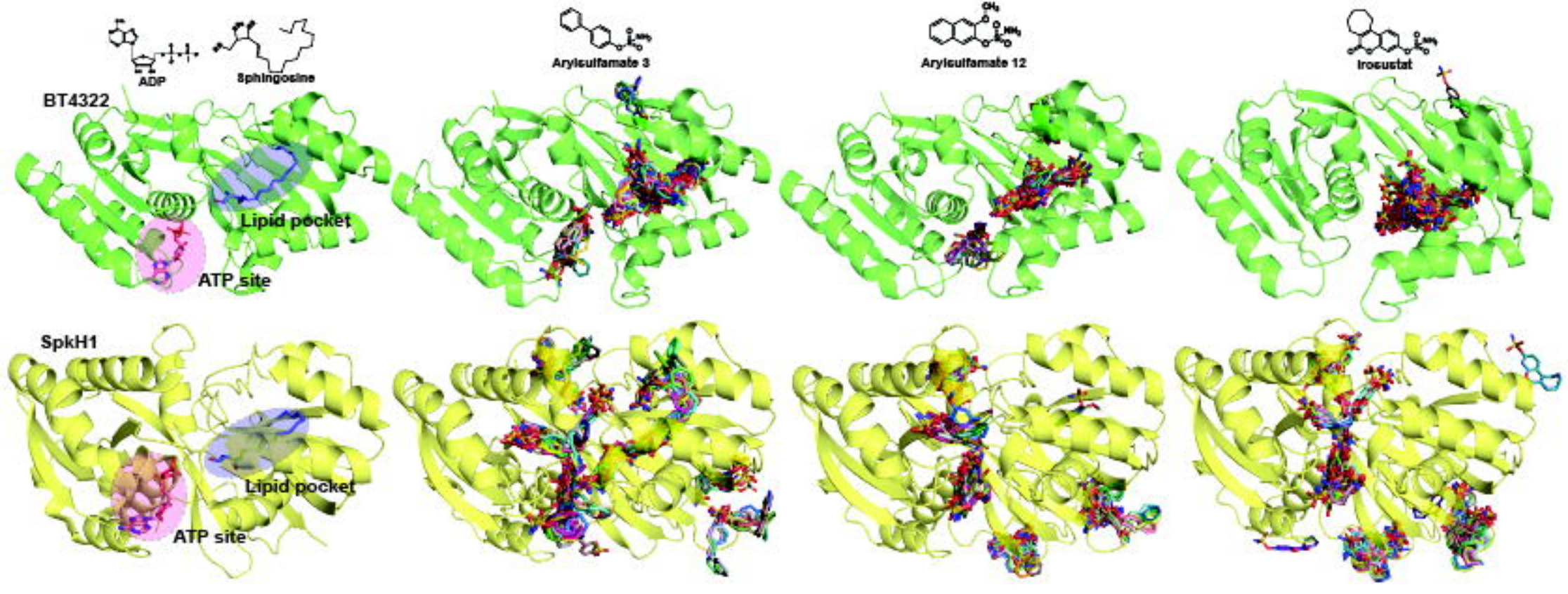
Extended data 6 Sequence alignment of putative DAGKs from select Bacteroides species.
Sequences were aligned using the MAFFT online server and visualised using Jalview.



Extended data 7. The effects of sphigosine kinase inhibitors on the growth of HGM Bacteroidota species
Bacteroidota species were grown in BHI, 1 % DMSO, with 0.1 mM of sphingosine kinase inhibitor (SKI) and 0.5 mM sphingosine kinase 1 inhibitor II and the effects on growth observed. Data are technical triplicates with the standard error of the mean.



Extended data 8. The effects of diacylglycerol kinase inhibitors on the growth of HGM Bacteroidota species
Bacteroidota species were grown in BHI, 1 % DMSO, with 0.125 mM of diacylglycerol kinase inhibitor I (DAGKI-i) and 0.0625 mM diacylglycerol kinase inhibitor II (DAGKI-ii) and the effects on growth observed. Data are technical triplicates with the standard error of the mean.



Extended data 9 Blind docking of arylsulfamates onto BT4322 and SpkH1.
The results of blind docking (no site specified) arylsulfamate inhibitors into the alphafold model of BT4322 (Q89ZQ4) and the crystal structure of SpKH1 (3VZB). It can be observed all arylsulfamates preferentially cluster in, or near, the lipid binding pocket. For SpKH1 there appears to be no single preferred clustering site.