

**Enzymatic Elucidation of Carbohydrate Utilization
in Marine *Bacteroidetes***

Inauguraldissertation

zur

Erlangung des akademischen Grades eines

Doktors der Naturwissenschaft (Dr. rer. Nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Universität Greifswald

vorgelegt von

Theresa Dutschei

Januar 2023

Dekan: Prof. Dr. Gerald Kerth

1. Gutachter*in: Prof Dr. Uwe T. Bornscheuer
2. Gutachter*in: Prof. Dr. Wolfgang Streit

Tag der Promotion: 27.03.23

Content

Content.....	III
Abbreviations.....	V
Scope and outline.....	VI
1 Background	1
1.1 Algae blooms and the marine carbon cycle	1
1.1.1 Polysaccharides connection plant-algae	2
1.2 Marine polysaccharides	2
1.2.1 Ulvan	2
1.2.3 Xylan.....	3
1.3 Polysaccharide utilization and CAZymes	4
1.3.1 Polysaccharide utilization of marine microbes.....	4
1.3.2 Carbohydrate active enzymes (CAZymes)	5
Glycoside hydrolases.....	5
Polysaccharide lyases	6
Carbohydrate esterases.....	7
Additional carbohydrate-associated enzymes	7
1.3.3 Enzymatic degradation pathways for marine polysaccharides.....	8
1.4 Application of marine polysaccharides and enzymatic cascades	9
1.4.1 Biotechnological application of algal polysaccharides	9
1.4.2 Algal biomass and its production.....	9
1.4.3 Biorefinery process: application of enzymatic cascades.....	9
2 Results	11
2.1 Expanding the ulvan polysaccharide degradation cascade (Articles I & II).....	11
2.1.1 Degradation of uronic acid-containing ulvan oligosaccharides	11
2.1.2 Exploration of a novel enzyme activity: ulvan oligosaccharide dehydratase	13
2.2 <i>Bacillus</i> meets ulvan: Application of the ulvan-enzymatic toolbox (Article III).....	14
2.3 Xylan degradation in marine <i>Bacteroidetes</i> (Article IV).....	17

2.3.1 Proteomic results reveal PUL target specialization.....	17
2.3.1 CAZyme activities underline the PULs target specialization	18
2.3.2 Contextualisation of marine xylan polysaccharide degradation	21
2.4 Microbial adaption mechanism in polysaccharide utilization (Articles V & VI)	21
2.3.1 The RuMP pathway is involved in the degradation of methylated sugars (Article V)	22
2.3.2 Role of novel alcohol dehydrogenases in polysaccharide degradation (Article VI)	23
3 Summary.....	25
4 References.....	26
Authors contributions	33
Article I	37
Article II	51
Article III	89
Article IV	127
Article V	181
Article VI	199
List of publications	237
Acknowledgments	239

Abbreviations

AA	Auxiliary activity
ADH	Alcohol dehydrogenase
<i>B. licheniformis</i> (Bli)	<i>Bacillus licheniformis</i>
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
BX	Beechwood xylan
CAZymes	Carbohydrate active enzymes
CE	Carbohydrate esterase
C-PAGE	Carbohydrate polyacrylamide gel electrophoresis
CPX	<i>Caulerpa prolifera</i> xylan
CYP	Cytochrome P450 monooxygenase
DNS	Dinitrosalicylic acid
<i>E. coli</i>	<i>Escherichia coli</i>
<i>F. agariphila</i>	<i>Formosa agariphila</i> KMM 3901 ^T
FACE	Fluorophore-assisted carbohydrate electrophoresis
<i>Flavimarina</i> sp.	<i>Flavimarina</i> Hel_I_48
GH	Glycoside hydrolase
Glc	D-Glucose
GlcA	D-Glucuronic acid
HPS	3-Hexulose-6-phosphate synthase (hxIA)
IdoA	L-Iduronic acid
MBTH	3-Methyl-2-benzothiazolinone-hydrazone-hydrochloride-hydrate
PHI	6-Phospho-3-hexulose isomerase (hxIB)
PL	Polysaccharide lyase
PPX	<i>Palmaria palmata</i> xylan
PUL	Polysaccharide utilization locus
RAX	Rye arabinoxylan
Rha	L-Rhamose
Rha3S	L-Rhamnose-3-sulfate
RuMP	Ribulose monophosphate pathway
UH	Ulvan hydrolysate
UL	Ulvan lyase
WAX	Wheat arabinoxylan
Xyl	D-Xylose
Xyl2S	D-Xylose-2-sulfate
<i>Z. galactanivorans</i>	<i>Zobellia galactanivorans</i> Dsji ^T

Scope and outline

Marine algae are essential for fixation of carbon dioxide, which they transform into complex polysaccharides. These carbohydrates are degraded e.g., by marine *Bacteroidetes* and the understanding of their decomposition mechanism can expand our knowledge how marine biomasses can be accessed. This understanding then gains insights into the marine carbon cycle.

This thesis summarizes the current knowledge of marine enzymatic polysaccharide degradation in review **Article I** and extends a previously discovered ulvan degradation pathway in **Article II** with the description of a novel dehydratase involved in the ulvan degradation pathway. This enlarged ulvan-degradation pathway can be used to generate fermentable sugars from the algal derived polysaccharide ulvan. A potential biorefinery process is proposed in **Article III**, where *B. licheniformis* was engineered to degrade ulvan, thus establishing the initial steps for a microbial cell factory development. In addition to ulvan, also plenty of other complex carbohydrate sources are present in the ocean. The enzymatic elucidation principles previously developed were thus adapted towards a new marine carbohydrate. In **Article IV** a xylan utilization pathway was elucidated, using enzymes present in *Flavimarina* Hel_I_48 as model bacterium. The *Flavimarina* genome contains two separated genome clusters which potentially targets xylose containing polymers reflecting the diversity and adaptations towards different marine xylan-like substrates. Besides, marine *Bacteroidetes* are adapted towards decomposition of methylated polysaccharide, e.g., porphyran, via demethylation catalyzed by cytochrome P450 monooxygenases. This reaction results in the formation of toxic formaldehyde and thus the marine *Bacteroidetes* require formaldehyde detoxification principles. The analysis of potential formaldehyde detoxification mechanisms revealed a marine RuMP pathway (**Article V**) and a novel auxiliary activity of an alcohol dehydrogenase of which the encoding gene is adjacent to the demethylase cluster (**Article VI**).

Article I: Marine polysaccharides: Occurrence, enzymatic degradation and utilization

M. Bäumgen*, [T. Dutschei*](#) and U. T. Bornscheuer, *ChemBioChem* **2021**, *22*, 2247-2256.

Marine algae polysaccharides are highly diverse in their structure and occur in macro- or microalgae. Therefore, they need a multitude of different enzymes classes working together in cascades to degrade carbohydrates. This review summarizes the current state of the knowledge in the enzymatic polysaccharide degradation of ulvan, laminarin, carrageenan and porphyran. It highlights the structural diversity and the opportunities of enzymatic degradation cascades in the biotechnology field.

Article II A new carbohydrate-active oligosaccharide dehydratase is involved in the degradation of ulvan

M. Bäumgen*, [T. Dutschei](#)*, D. Bartosik, C. Suster, L. Reisky, N. Gerlach, C. Stanetty, M. D. Mihovilovic, T. Schweder, J. H. Hehemann and U. T. Bornscheuer, *J. Chem. Biol.* **2021**, 297, 101210

Article II expands the ulvan degradation cascade of *Formosa agariphila* KMM 3901^T – discovered by us previously and summarized in **Article I** – for a uronic acid containing ulvan polysaccharide which is one of the most abundant subunits of ulvan. The additional elucidation of a novel enzyme activity of an oligosaccharide dehydratase, in frame of an alternative ulvan degradation cascade, enhances the ulvan degradation procedure and represents the first described oligosaccharide dehydratase so far. The alternative ulvan degradation pathway was additionally highlighted in **Article I**.

Article III Metabolic engineering enables *Bacillus licheniformis* to grow on the marine polysaccharide ulvan

[T. Dutschei](#), M. K. Zühlke, N. Welsch, T. Eisenack, M. Hilkmann, J. Krull, C. Stühle, S. Brott, A. Dürwald, L. Reisky, J.-H. Hehemann, D. Becher, T. Schweder and Uwe T. Bornscheuer, *Microb. Cell Fact.* **2022**, 21, 207

The ulvan degradation pathways described in **Articles I & II** can be used to generate fermentable sugars, serving as microbial feedstock. In this study, *B. licheniformis* was found to use ulvan derived sugars as carbon source. Furthermore, a microbial cell factory was developed, by introducing their genes into *B. licheniformis* which then can produce two ulvanolytic enzymes. This enables the strain to solely grow on the green algal polysaccharide ulvan. This study thus describes the principle to use ulvan polysaccharides in a biotechnological process.

Article IV Utilization of a diverse range of xylan structures by marine *Bacteroidetes*

[T. Dutschei](#), I. Beidler, D. Bartosik, J.-M. Seeßelberg, M. Teune, M. Bäumgen, S. Q. Ferreira, J. Heldmann, F. Nagel, J. Krull, L. Berndt, K. Methling, M. Hein, D. Becher, P. Langer, M. Delcea, M. Lalk, M. Lammers, M. Höhne, J. H. Hehemann, T. Schweder and U. T. Bornscheuer, *Environ. Microbiol.* **2022** (under revision)

In **Article IV** the concept described in **Article II** for ulvan degradation was adapted towards the polysaccharide xylan. The model organism *Flavimarina* sp. Hel_I_48 contains two putative xylan polysaccharide utilization loci (PULs). The corresponding genes were identified from a Helgoland spring algal bloom metagenomic dataset. The presence of two xylan targeting PULs and the multiple occurrences of corresponding transport proteins suggests that these PULs encode enzymes which may target a variety of different xylan-like targets in the ocean. This

was experimentally proven via a multidisciplinary approach including proteogenomic and biochemical characterization.

Article V Connecting algal polysaccharide degradation to formaldehyde detoxification

S. Brott, F. Thomas, M. Behrens, K. Methling, D. Bartosik, T. Dutschej, M. Lalk, G. Michel, T. Schweder, and U. T. Bornscheuer, *ChemBioChem* **2022**, *23*, e202200269

The breakdown of marine polysaccharides may result in the accumulation of either sugars or other metabolites, like formaldehyde. The degradation of porphyran described in review **Article I** includes marine cytochrome P450 monooxygenase, which demethylate 6-*O*-methyl galactose present in the porphyran and allows the organism to further metabolize this carbohydrate. The demethylation leads to the accumulation of formaldehyde, which is toxic for cells. Therefore, the organism *Z. galactanivorans* must have developed specific formaldehyde detoxification pathways. This publication shows that formaldehyde can be detoxified using the RuMP pathway, with 3-hexulose-6-phosphate synthase hxlA (HPS) and 6-phospho-3-hexulose isomerase hxlB (PHI). This niche formaldehyde detoxification pathway was biochemically characterized and its physiological role was proven.

Article VI A unique alcohol dehydrogenase involved in algal sugar utilization by marine bacteria

S. Brott, K. H. Nam, F. Thomas, T. Dutschej, L. Reisky, M. Behrens, H. C. Grimm, G. Michel, T. Schweder, and U. T. Bornscheuer *Appl. Microbiol. Biotechnol.* **2022** (*submitted*)

Building up on the study described in **Article V** demethylation clusters in marine *Bacteroidetes* were found to contain a conserved alcohol dehydrogenase (ADH) and an esterase. The potential impact in formaldehyde detoxification was investigated. Highlighting the structural and biochemical properties of two ADH, from the *Bacteroidetes* *F. agariphila* and *Z. galactanivorans*, respectively, revealed no *in vitro* formaldehyde detoxification but only a conversion of aromatic aldehydes. The knock-out of the ADH in *Z. galactanivorans* revealed that there is a need of this enzyme in metabolising methylated sugars. Therefore, the study showed the first insights in the role of ADHs in polysaccharide utilization.

*shared first author

1 Background

1.1 Algae blooms and the marine carbon cycle

The oceans cover about 70% of the earth surface making it the largest ecosystem on earth^[1]. Therefore, they play a major role in carbon sequestration, with most marine carbon being fixed in marine carbohydrate structures^[2] of algae (Fig. 1a). With 52 billion tons fixed carbon per year, algae contribute half of the global primary production besides terrestrial plants^[3]. Algae are eukaryotic photoautotroph organisms which can be divided in two subclasses: micro- and macroalgae. Additionally, macroalgae can be divided in three main classes: red algae (Rhodophyta), brown algae (Phaeophyta) and green algae (Chlorophyta) by their photosynthetic pigment, which allows them to absorb different wavelength of the sunlight, which is an adaption towards the marine ecosystem and seasons^[4,5].

Algae do not grow continuously, usually only in temporary blooms. Those blooms can trigger secondary blooms of planktonic bacteria (Fig. 1b), which then can feed on the algal biomass and release much of the stored carbon dioxide^[6,7]. The residual biomass sinks to the seafloor as carbonaceous particles, where particle associate bacteria can remineralize while they sink^[6,7]. It has been shown that the 2009 spring phytoplankton bloom in the North sea was reoccurring from 2010 to 2012 and showed low variation of the microbial community (Fig. 1b)^[6,8]. This suggests the importance of some genera of *Flavobacteriia* in the biomass degradation of the algae blooms, while further analysis of the degradation principles involved in the organisms provides further insight into the marine carbon cycle^[6,8].

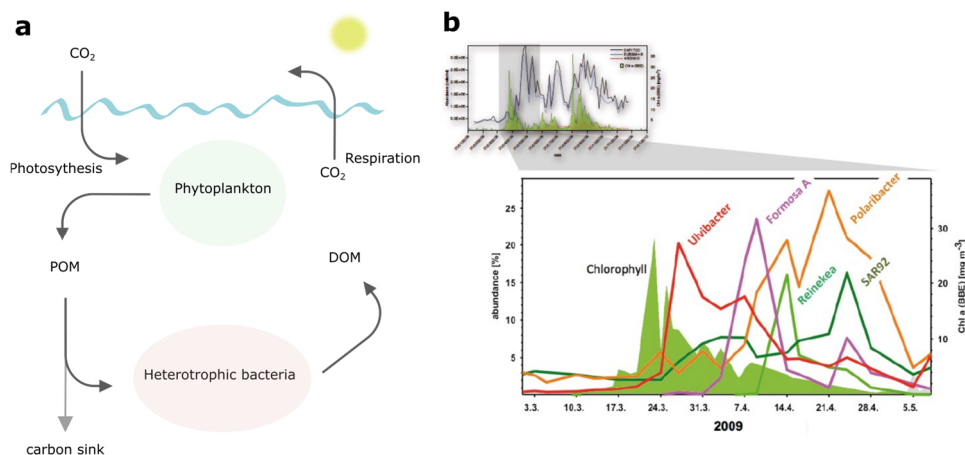


Figure 1: Microalgal bloom induced succession of bacterial clades. (a) simplified marine carbon cycle^[9] POM: particulate organic matter, DOM: dissolved organic matter (b) microalgal bloom in Helgoland 2009, figure adapted from Teeling *et al.*^[6,8]

1.1.1 Polysaccharides connection plant-algae

Plants and algae have a complex phylogenetic history, including genes responsible for carbohydrate synthesis and modification through a series of primary (leading to red algae, green algae, and land plants) and secondary (generating brown algae, diatoms, and dinoflagellates) endosymbiotic events^[10]. Evolutionary territorialization from algae to plants resulted in similar properties of polysaccharide compositions and structures^[10]. Especially, studies like Niklas *et al.* highlight the evolutionary connection of plant and algal cell wall^[11], based on similarities in polysaccharide metabolism^[12] and structural organisation^[13]. The similarities of the skeletal polysaccharides occur especially in cellulose, xylan, mannans^[13] and lignin-like structures^[14,15]. The structural modification of each marine polysaccharide makes it unique. Specific to the marine habitat is sulfation, which does not occur in the terrestrial origin^[10], as no sulfatase and carbohydrate sulfotransferases are found in plant genomes^[12]. The general occurrence of sulfatation patterns can therefore be seen as a marker of marine origin^[16].

1.2 Marine polysaccharides

The largest part of marine carbon is fixed in marine carbohydrate structures^[2]. The carbohydrates mainly occur in marine plants, macro- and microalgae, but also occur in marine fungi and marine invertebrates^[17–20]. Most organisms use their polysaccharides as carbon source and structural components^[13] while others secrete them as extracellular polysaccharide substances (EPS)^[21]. The carbohydrate composition, structure of the backbone and side chain modifications^[20] are highly individual between the different species. Additionally, the composition is highly location- and season-dependent^[20].

1.2.1 Ulvan

Ulvan is the main polysaccharide of the green algae *Ulva* spp. and serves as a structuring polysaccharide in the algal cell wall, resulting in strong and flexible properties of the algae^[22–24]. The heterogenous polysaccharide is an anionic, sulphated, water-soluble polysaccharide containing as constituent of ulvan are, L-rhamnose, D-xylose and D-glucuronic acid^[23]. In addition, D-glucose and L-iduronic acid as well as variable amounts of D-mannose and D-/L-galactose are found in ulvan^[23]. The main repeating unit of ulvan is a dimer called ulvanobiouronic acid A, which consists of D-glucuronic acid β -(1,4) linked to a L-rhamnose-3-sulfate unit (Fig. 2)^[23].

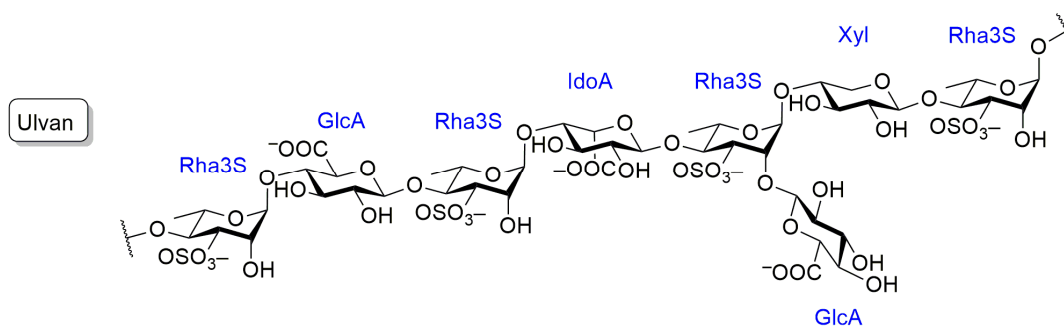


Figure 1: Structure of ulvan. It is composed of D-glucuronic acid (GlcA), L-iduronic acid (IdoA), D-xylose (Xyl) and L-rhamnose (Rha). The number and S combination represent the position of the sulfate group.

Ulvan extract is a thermoreversible gel, which can bind metal ions that support the action of gel formation^[23] so that this property can be regulated. These general physical-chemical properties and the content of pharmaceutical precursors (i.e, sulfated sugars) makes this polysaccharide an interesting starting material for new biopolymers for food technology and as resource for the pharmaceutical industry^[25].

1.2.3 Xylan

There are some polysaccharides in the marine ecosystem which share high similarity in their composition to their terrestrial relatives, like the hemicellulose xylan^[26]. This is known to occur in cell walls of woods and grains in form of arabinoxylan, galactoarabinoxylan and glucuronoxylan with a β -1,4-linked D-xylopyranose backbone (Fig. 3)^[27]. Marine xylan can be found in the cell wall of green algae (Chlorophyta/Charophyta) and red algae (Rhodophyta)^[10] in four different macroalgae families: Bryopsidaceae, Caularpaceae, Udotaceae and Dichotomosiphonaceae^[28]. Its backbone is composed of β -1,4- or β -1,3-linked D-xylopyranose, depending on the algal species and source. Substituted β -1,4-xylan was found in species of charophyte green algae^[26,29]. In chlorophyte green algae β -1,3-xylan is part of the cell wall^[30,31] and is reported to form triple helix microfibrils. Red algae contain mostly β -1,3-linked xylans^[32-34] and β -1,3-1,4-linked xylans, in which one β -1,3-linkage follows four β -1,4-linkages, like in the red algae *Palmaria palmata*^[31,35]. Furthermore, marine xylans can be sulfated or phosphorylated, further demonstrating the high variability in polysaccharide composition and adaption towards the marine environment^[35]. Besides, xylose-containing polysaccharides can also be found in microalgal biomasses^[21,36,37], as recently discovered in the diatoms *Thalassiosira weissflogii* and *Chaetoceros socialis*^[37]. They are suggested to contain β -1,4-linked xylan and potential arabinoxylan-like structures as determined by microarray assays with appropriate antibodies and controls^[21,37]. The broad spectrum of potential xylan-like structures in the ocean is discussed in **Article IV**.

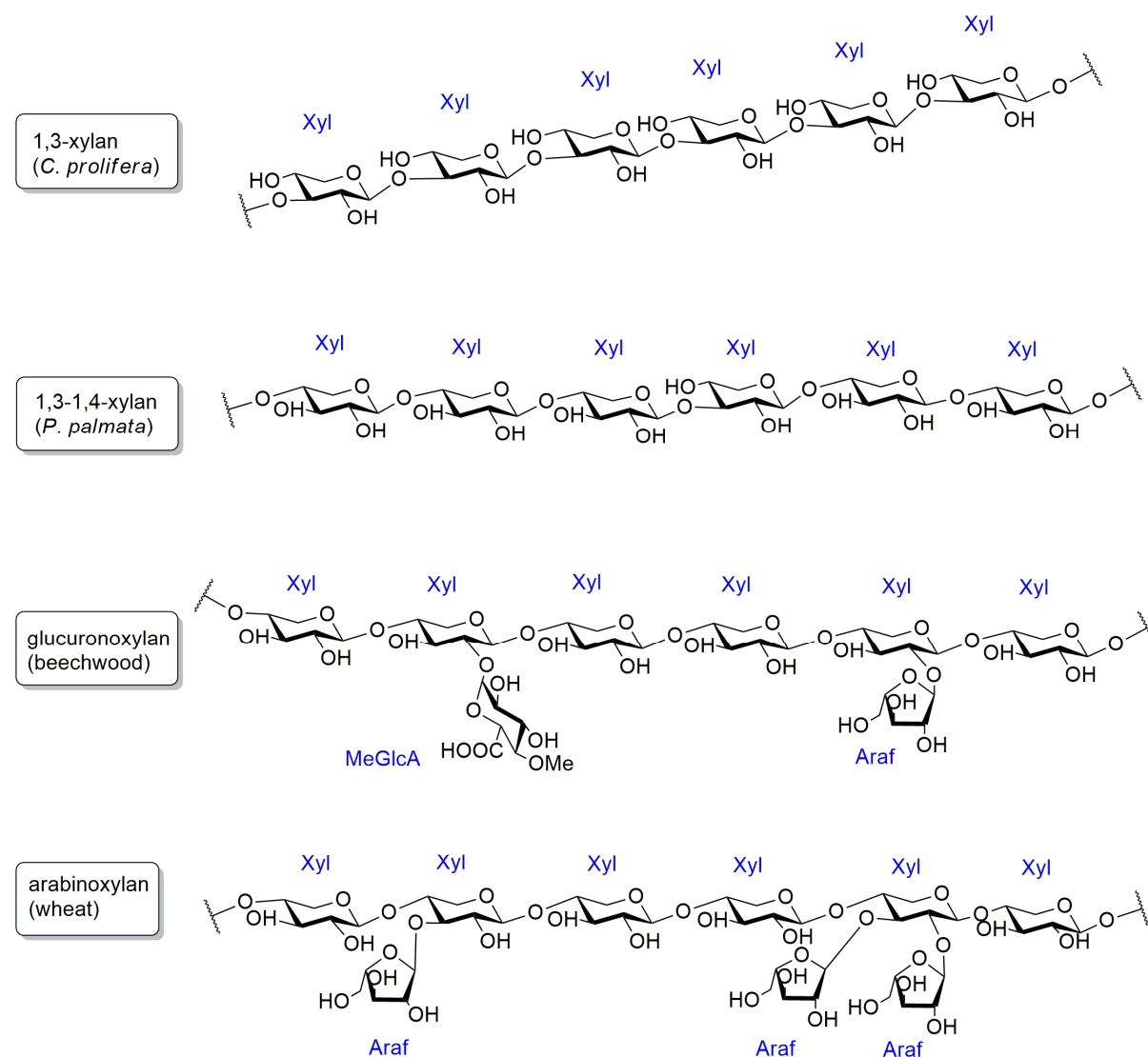


Figure 3: Xylan structures from marine and terrestrial origin. Structures are postulated by the common motifs and linkages known from the green algae *C. prolifera*, red algae *P. palmata*, beechwood and wheat hemicellulose^[26]. They contain moieties of D-xylose (Xyl), L-arabinofuranoside (Araf) and 4-O-methyl- α -D-glucopyranosyl uronate (MeGlcA).

1.3 Polysaccharide utilization and CAZymes

1.3.1 Polysaccharide utilization of marine microbes

The high diversity of the marine carbohydrate leads to the adaption of co-associate bacteria carrying genes encoding for the CAZymes that can degrade the corresponding polysaccharides. These CAZymes are widely distributed among bacterial and archaeal phyla^[38]. The phylum of the *Bacteroidetes* is considered a specialist in the degradation of high molecular weight organic matter, such as algal carbohydrates^[39–41]. The heterotrophic *Bacteroidetes* have evolved a unique degradation machinery, known as polysaccharide utilization loci (PULs), encoding for CAZymes, sugar recognition and uptake proteins (SusD

and SusC-like)^[38,42]. This pattern enables the recognition of PULs in metagenome datasets, due to their highly conserved structures which is specific for polysaccharides^[36,38,43].

The heterotrophic *Bacteroidetes* produces extracellular CAZymes which can predigest the polysaccharide to a size range that can be bound (SusD-like protein) and taken up by the sugar transporter (SusC-like protein) into the periplasm, where they are further degraded to the corresponding monosaccharides. These monomers are then subsequently up-taken across the cytoplasmic membrane and introduced into the sugar metabolism^[38,44–46]. This microbial polysaccharide utilization mechanism is considered as „selfish“-mechanism, with little to no extracellular hydrolysis products formed^[46]. Another principles of polysaccharide utilization is the "sharing"-mechanism, in which the organism have free or cell surface attached CAZymes degrading the polysaccharide to a suitable uptake size^[46], and the "scavenging" mechanism, in which the organism cannot produce CAZymes and only take up the pre-hydrolysed carbon source from another organism^[46] while "semi-selfish"-behaviour also occurs in when only some cleavage products are available for other bacteria^[47]. Overall, marine organisms exhibit social and cooperative behaviour^[45,47], forming a microbial community for polysaccharide degradation.

1.3.2 Carbohydrate active enzymes (CAZymes)

Polysaccharides are highly complex molecules which require different enzymes involved in their synthesis, degradation and modification. These are called carbohydrate active enzymes (CAZymes)^[48]. CAZymes can be divided into six different protein classes: glycoside hydrolases (GH), glycoside transferases (GT), polysaccharide lyases (PL), carbohydrate esterases (CE), non-catalytic carbohydrate binding modules (CBM) and the additional carbohydrate associated auxiliary activity (AA) enzymes e.g. LPMOs or P450s^[49]. In total, around 365 main protein families have already been described in the CAZY-database^[49]. Enzyme hydrolyzing the same substrate are sometimes found in different families; the International Union of Biochemistry (IUB) classification is different because enzymes are mostly classified according to their substrate specificity^[50].

Glycoside hydrolases

The mechanisms of glycoside hydrolases (GH) results in either in retention or inversion of the anomeric configuration^[51] defining the two major modes of action. The general catalysis involve two critical residues, a proton donor and a nucleophile/base^[50,51]. In the retaining hydrolysis, the nucleophilic base is close to the anomeric carbon and in the inverting enzymes the base is more distant which must accommodate a water molecule between the

catalytic residue and the sugar^[50,51]. Another classification of GHs is commonly made by their mode of action in terms of *exo* or *endo* activity, whether the enzyme cleaves of terminal moieties of a polysaccharide or in the main backbone of a polysaccharide^[50]. This is dependent on the structure of the catalytic sites, which can be a pocket, cleft or a tunnel, as a true *endo*-enzyme has an open cleft that allows a flexible binding and action within the polysaccharide backbone^[50]. Additionally, this activity distinction is highly dependent on the substrate of the enzyme, as GHs can perform multiple catalytic events coherently^[50]. For instance, cellobiohydrolases have a tunnel shaped active site, which helps to keep the substrate in the active site, enabling multiple hydrolysis events^[50]. The unsaturated unsaturated β -glucuronyl hydrolase (GH105) has a cavity for glycan binding. Its active site is specialized by the surrounded loops with a conserved active site pocket which accepts the unsaturated residue, like in the GH88 family^[52]. Both enzymes participate in the ulvan degradation cascade^[44] and the role of GH105 is further highlighted in **Article II**.

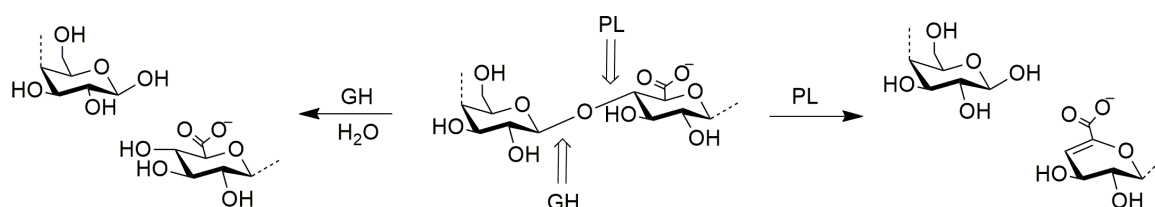


Figure 4: Comparison of the initial degrading enzymes GH and PL. GH cleaves the glycosidic bond by addition of water, leaving the 4-OH group at the non-reducing end of the cleavage product. PLs do not require water and generate a hexenuronic residue from the hydrolysis products at the non-reducing end^[53].

Polysaccharide lyases

In contrast to glycoside hydrolases, polysaccharide lyases can breakdown polysaccharides without the intervention of a water molecule (Fig. 4)^[53]. The polysaccharide lyases act in a β -elimination cleavage which only can occur when the sugar is substituted with an acidic group next to the carbon forming the glycosidic bond^[54,55]. The mechanism can be divided into three chemical steps: (I) the carboxyl group of the substrate is neutralized to reduce the pK_A of the C5 position, (II) an enolate anion intermediate is formed by proton abstraction at C5, and (III) an electron transferred from the carboxyl group to form a double bond in C4/C5 with a concomitant cleavage of the C4/C1 glycosidic bond^[55,56]. The neutralization step can be accomplished in two ways, either by metal-assisted neutralization of the acidic group of the sugar, with an arginine or lysine acting as the Brønsted base, or by neutralization of the acidic sugar by approaching an amino or acidic group that forces its protonation, tyrosine or tyrosine-histidine acting as the Brønsted base^[55].

The initial degradation of ulvan is performed by an ulvan lyase which can be found in PL24, PL25, PL28 and PL40 subclasses^[57–61]. Ulvan lyases cleave at the glycosidic bond next to GlcA and IdoA^[61] and lead to the formation of 5-dehydro-4-deoxy-D-glucuronate, which then can be cleaved-off by a GH105^[23]. Beside the polysaccharide lyases also an oligosaccharide dehydratase, targeted in **Article II**, can lead to the formation of an unsaturated uronic acid. In **Article II** it is postulated, that the elucidated oligosaccharide dehydratase follows a lyase-like elimination mechanism rather than oxidation mechanism as in known monosugar dehydratases^[62,63].

Carbohydrate esterases

Carbohydrate esterases (CE) catalyse *O*- or *N*-deac(et)ylations by removing ester decoration from carbohydrates^[64]. Removal may accelerate polymer degradation as this facilitates access for glycoside hydrolases^[65,66] and increases the solubility of the hemicellulose polymers by cleavage of acetyl groups or phenolic acids^[67] which was a target in **Article IV**. CEs span over 20 different families according to the CAZy-database^[64]. Most CEs follow a similar reaction mechanism as lipases and serine proteases with a catalytic triad composed of serine-histidine-aspartate, but some also follow a Zn²⁺-catalysed mechanism^[64].

The most extensively described carbohydrate esterase activity has been described for acetyl xylan esterases, which are found in CE families 1-7 and 16^[68]. These enzymes can deacetylate xylans which can be highly *O*-acetylated. They are mainly known from the decomposition of plant cell wall polysaccharides. Another well described subclass of CE is CE15 (4-*O*-methyl-glucuronyl methylesterase), an enzyme cleaving the interlinkage between hemicellulose and lignin moieties^[69,70]. Besides, the CE1 family encodes for ferulic acid esterases, which target phenolic groups e.g., bound to L-arabinose moieties of the polysaccharide in arabinoxylans.

Additional carbohydrate-associated enzymes

CAZymes often occur as multimodular proteins, either with other catalytic modules for coherent activity or with binding modules, which enable better access of the catalytic enzyme (part) to the polysaccharide. Those domains are called carbohydrate-binding modules (CBM) and they help to capture polysaccharide structures and can even recognise and accept amorphous structure^[49,64,71]. Another special enzyme class which plays a crucial role in marine carbohydrate degradation and often occurs as multimodular protein, are sulfatases, as sulfate groups are a common marine polysaccharide modification^[20].

Sulfatases are grouped into four classes based on sequence, homology, structure and mechanism. The main known class are the type I sulfatases which are formylglycine-dependent enzymes^[16,72]. They require a co-translational modification for their catalytic activity, in which a cysteine or serine is transformed to a formylglycine amino acid residue^[16]. In aerobic bacteria, the oxidation of the amino acid is catalysed by the formylglycine-generating enzyme (FGE), which can recognize the consensus sequence C/S-X-P-X-R as the amino acid motif^[16]. Besides, in anaerobic bacteria, the amino acid residues can be modified by an enzyme called sulfatase-maturing enzyme (anSME)^[73].

1.3.3 Enzymatic degradation pathways for marine polysaccharides

In general, the marine polysaccharide decomposition is an interplay of the individual enzyme activities. Fucoidan, the complex brown algae polysaccharide, requires almost a hundred different enzymes to be degraded by the bacterial phylum *Verrucomicrobiota*^[74], which is due to high complexity of the marine polysaccharide composition and its variations in season and location. Besides, the phylum *Bacteroidetes* is considered to be a specialist in the degradation of marine carbohydrates^[39]. In the marine *Bacteroidetes Formosa agariphila* KM3901^T 27 genes are encoded in the ulvan utilization locus. In the study of Reisky *et al.* 14 of those genes were functionally characterized and were mostly involved in the degradation of the xylo-ulvan-oligosaccharide degradation. With further characterization of CAZymes, the degradation of D-glucuronic acid-containing oligosaccharides was elucidated as summarized in **Article II**. The current state of enzymatic cascade reactions of the most abundant marine polysaccharides, carrageenan, porphyran, laminarin and ulvan have been summarized in the review **Article I**.

The heterotrophic bacterium *Zobellia galactanivorans* also uses a cluster of genes for the decomposition of the red macroalgal carbohydrate carrageenan. In addition to glycoside hydrolases, sulfatases also play a crucial role in the organisms catabolism^[75]. Further, important enzyme with different activities are involved in maturing the polysaccharide for decomposition, like a newly found marine P450 monooxygenase^[76]. This enzyme demethylates the galactan porphyran which can then be further degraded by glycoside hydrolases^[76,77]. The enzymatic reaction results in the accumulation of formaldehyde which the organism need to cope due to its cell toxicity. Therefore, marine *Bacteroidetes* like *Z. galactanivorans* evolved specific additional formaldehyde detoxification mechanisms, like the RuMP-pathway, via 3-hexulose-6-phosphate synthase (HPS) and a 6-phospho-3-hexuloisomerase (PHI)^[78]. The knockout of the relevant enzymes of the pathway leads to cell death caused by a deficiency of formaldehyde tolerance^[78]. While this detoxification

mechanism is known for the terrestrial organism this is the first study describing it in a marine organism and its connection with the marine carbohydrate degradation (**Article V**).

1.4 Application of marine polysaccharides and enzymatic cascades

1.4.1 Biotechnological application of algal polysaccharides

The knowledge of enzymatic saccharification processes and specific CAZyme activities opens a new field of accessibility for algal derived compounds (**Article II**) and as general alternative microbial feedstock (**Article III**). Therewith, the worthwhile generation of defined bioactive substances and rare sugars are now possible^[22,44,79]. Algae contain valuable compounds beyond their polysaccharides, like polyphenols, proteins and lipids^[79] that are already used in industry due to their antimicrobial, antioxidative, antiviral and anticancerogenic activity^[25]. Therefore macro- and microalgae and their components find broad applications in the field of cosmetics^[80], pharmaceuticals^[81–84], biochemical production^[85], generation of biobased polyesters^[86] and as biofuels^[87].

1.4.2 Algal biomass and its production

Algae biomasses accumulate in coastal areas^[88,89] which are highly undesirable due to the negative impacts on tourism and beach ecology. Additionally, the disposal of algal waste is very cost intensive^[89,90], thus the use of the abundant biomass for having also beneficial applications is of interest. The challenging part in using wild harvest algal biomass for bioprocesses is that there are high variations of the biomass composition depending on seasons and species^[22,91]. Additionally, the biomass is highly heterogeneous in availability due to contaminations with sand, animals and molluscs^[88]. Alternatively, growing algae in offshore farms in the sea is possible. Marine farms produce more biomass per hectare sustainably compared to their terrestrial counterparts without using arable land^[92]. Additionally, different algal biomass cultivation techniques like aquacultures (e.g. open ponds, integrated multitrophic aquaculture) or photoreactors can be used^[88]. These enable much more controlled cultivation conditions which are relevant for the application in the pharmaceutical, cosmetic and food industry^[92]. The biomass resource is therefore crucial for the further establishment of bioprocesses.

1.4.3 Biorefinery process: application of enzymatic cascades

Macroalgae can be a feedstock for biorefinery^[93], due to their high polysaccharide content. They have a high potential to serve as a new renewable feedstock replacing fossil oil or

terrestrial biomasses in e.g. biofuel production^[92,94]. Some examples already exist like the production of bioethanol by fermenting macroalgal biomass^[95,96]. Beyond the macroalgal utilization, knowledge of the enzymatic degradation process can enable, metabolic engineering approaches that allowing the integration of biomasses into already established biotechnological processes (**Article III**). This expands the possibilities of marine carbohydrate utilization. Such an approach was already demonstrated by Wargacki *et al.*, where alginate lyases were expressed in *Escherichia coli*, leading to the ability to degrade, uptake and metabolize alginate under bioethanol production^[95].

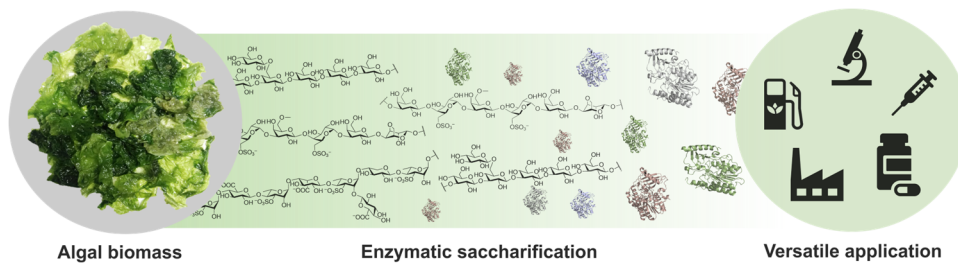


Figure 5: Biotechnological application of algae carbohydrates. Marine polysaccharides can be harvested from algae biomass. Enzymatic saccharification with CAZymes can lead to valuable products for biofuel, pharmaceutical and biotechnological industry. This figure is derived from Article I.

2 Results

2.1 Expanding the ulvan polysaccharide degradation cascade (**Articles I & II**)

In the previous study of Reisky *et al.* the ulvan degradation pathway from *F. agariphila* KMM3901^T was revealed^[44]. There, 14 from 27 catalytically active enzymes of polysaccharide utilization loci (PUL) H were found and characterized, while focusing on the degradation of xylose-containing ulvan oligosaccharides, like the Δ -Rha3S[2GlcA]-Xyl-Rha3S oligosaccharide (Fig. 6). This study did not further evaluate the degradation of glucuronic acid containing-ulvan oligosaccharide which can accumulate due to a product inhibition of the initial ulvan degrading enzyme P30_PL28^[60]. As GlcA-Rha3S is regarded as the main repetition unit of ulvan^[23], it was necessary to further explore the degradation pathway of this residual glucuronic acid-containing ulvan oligosaccharides and investigate further ulvanolytic enzymes from *F. agariphila* KMM3901^T PUL H. The additional oligosaccharide Δ -Rha3S-GlcA/IdoA-Rha3S was chosen an example and basis for the analysis of the degradation pathway of this uronic acid-containing ulvan oligosaccharide, as described in **Article II**. The herewith expanded ulvan degradation pathway was additionally highlighted in the review **Article I** (Fig. 6).

2.1.1 Degradation of uronic acid-containing ulvan oligosaccharides

Ulvan lyases can digest the ulvan polymer up to a dimeric level, but can be product inhibited^[60]. This can lead to an incomplete digestion of uronic acid-containing ulvan oligosaccharides, a so-called alternative degradation route for uronic acid-containing ulvan, was therefore the target of the investigations which resulted in **Article II**.

The Δ -Rha3S-GlcA/IdoA-Rha3S derives from the digestion of the polymer ulvan by the enzyme P30_PL28. Mechanistically, this is the result of an elimination mechanism leading to an uronic acid at the non-reducing end. This oligomer served as model substrate to study the incomplete digestion catalyzed by the enzyme P30_PL28. This uronic acid can be cleaved off using the unsaturated glucuronyl hydrolase P33_GH105 releasing an α -keto acid (5-dehydro-4-deoxy-D-glucuronate). The resulting Rha3S-GlcA/IdoA-Rha3S is then further processed by the P36 multimodular enzyme. The crystal structure of the P36 enzyme revealed a multimodularity, containing beside an sulfatase S1_25 domain an additional GH78 rhamnosidase domain. The P36_S1_25 was already known as non-reducing end sulfatase acting on the Rha3S-Xyl-Rha3S oligomer, which enables further

degradation by the rhamnosidase P20_GH78 as described by Reisky *et al.* (Fig. 6) or by the P36_GH78 domain. This proves the general known concept that multimodular proteins can have sequential enzyme activities^[16]. The resulting GlcA/IdoA-Rha3S oligomer can then be either degraded by the P34_GH3 enzyme releasing the GlcA/IdoA residue from the non-reducing end or can be dehydrated by the P29_PDnc enzyme. This enzyme function as an oligomer dehydratase in carbohydrate degradation was described for the first time in **Article II**. The enzyme activity results in an unsaturated uronic acid which then can be further cleaved off by the unsaturated glucuronyl hydrolase P33_GH105 resulting in Rha3S. The ulvan degradation pathways elucidated in Reisky *et al.* and in **Article II** are summarized as an overall pathway as shown in Figure 6 and as highlighted in **Article I**.

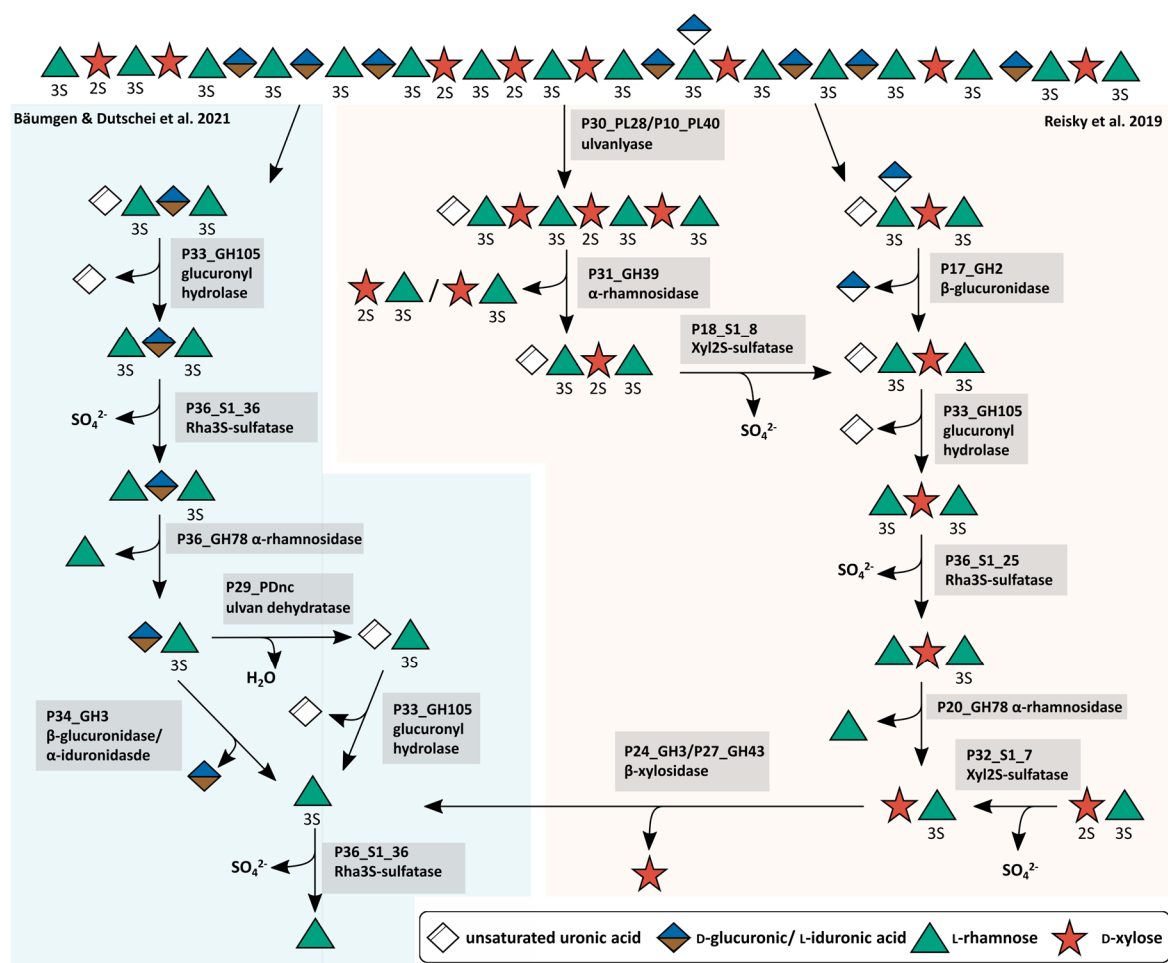


Figure 6: Metabolic ulvan degradation pathway by CAZymes[62,111]. The oligosaccharide on the top represents a section of a larger polysaccharide chain. The alternative ulvan pathway described in Article II is framed in blue, the ulvan degradation pathway published by Reisky *et al.* is framed in orange. This figure was adapted from Article I.

2.1.2 Exploration of a novel enzyme activity: ulvan oligosaccharide dehydratase

The enzyme P29_PDnc was previously characterized by Konasani *et al.* and it was claimed that it has a classic ulvan lyase activity similar to the P30_PL28 enzyme from *F. agariphila* KMM3901^T on polymeric ulvan^[97]. This enzyme activity could not be reproduced in our laboratory, with the enzyme bearing a N-terminal His-tag. To exclude differences in the construct designs, the plasmid described in the study of Konasani *et al.* was recreated with a C-terminal His-tag. The protein was produced according to the Konasani *et al.* study^[97]. Both enzyme constructs were compared with the P30_PL28 ulvan lyase using three different methods: (I) visualizing the degradation products of the hydrolysis using carbohydrate polyacrylamide gel electrophoresis (C-PAGE), (II) measuring the increase of reducing ends (MBTH-assay) and (III) measuring the formation of the lyase products spectrophotometrically over time (ulvan lyase assay) (Fig. 7A-C). In all three assays no hydrolysis products from the polymer were detectable indicating no activity on polymeric ulvan substrates. Excluding possible activity differences due to seasonal and regional variation of the algae and corresponding polysaccharides, a wide range of ulvans from different origin were investigated, but no hydrolysis products were detectable at all. Therefore, the activity of the enzyme on polymeric ulvan could be excluded and the general activity of the enzyme was still unknown. The investigation of a potential lyase activity led to analyses of the enzymatic activity regarding GlcA/IdoA-containing fragments as the carboxylic acid residue is crucial for a lyase-type mechanism. The screening of different GlcA/IdoA containing ulvan oligosaccharides resulted in an activity of the P29_PDnc on the dimer GlcA/IdoA-Rha3S, as a shift of the oligo band was seen in AMAC-FACE analysis. The formed product was purified and investigated via NMR analysis, showing that a Δ -Rha3S was formed. This reaction must be due to a release of a water molecule from the dimer GlcA-Rha3S, which was proven via mass spectrometry. Additionally, the unsaturated glucuronyl hydrolase GH105 was able to remove an α -ketoacid from the PD29_nc measured via the thiobarbituric acid assay (Fig. 7D). Overall, all chosen methods pointed to an oligosaccharide dehydratase mechanism.

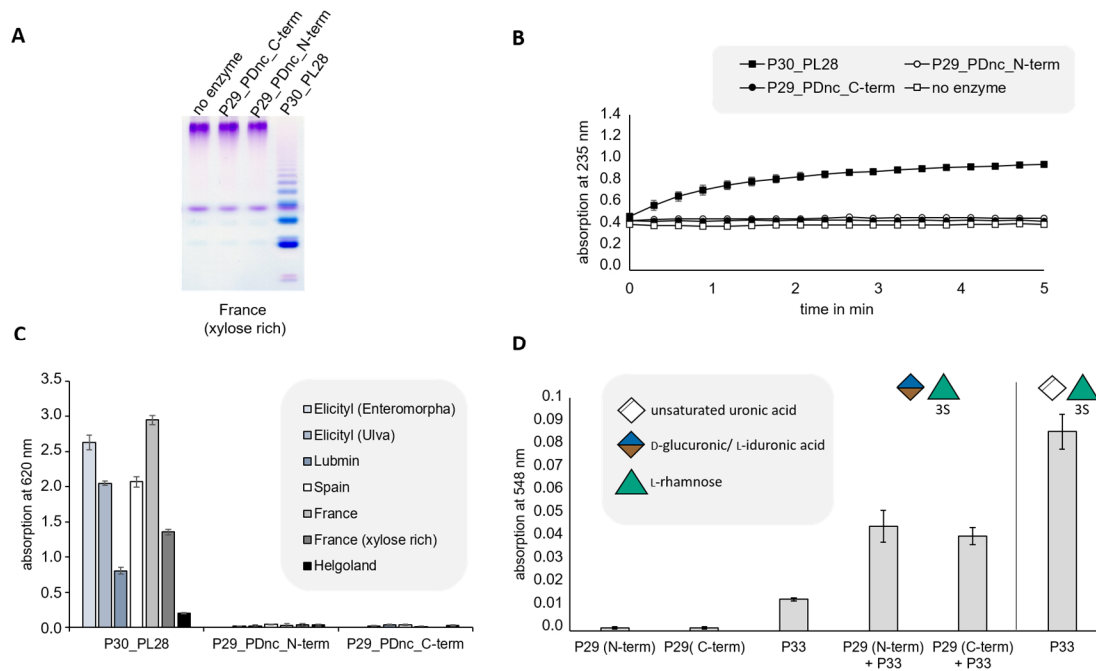


Figure 7: P29_PDnc is a dehydrate not an ulvan polymer lyase. Comparing the ulvan lyase P30_PL28 with the P29_PDnc construct with a His-tag using (A) C-PAGE (B) the ulvan lyase assay and (C) the MBTH assay. The activity on the oligosaccharide GlcA-Rha3S (C) was investigated with the thiobarbituric acid assay, measuring the formation of an unsaturated uronic acid via the release of an α -ketoacid by the enzyme P33_GH105.

2.2 *Bacillus* meets ulvan: Application of the ulvan-enzymatic toolbox (**Article III**)

An enzymatic cascade enabled the hydrolysis of the marine polysaccharide ulvan from the green algae *Ulva* sp. On the one hand the cascade can be used to isolate rare sugars and on the other hand the sugar-rich hydrolysate can be used as carbon source for biotechnology relevant strains and can form the basis of a new ulvan based biorefinery process. The aim of this project was to find a suitable microbial strain which can consume ulvanoligosaccharides and to further enable this strain to self-sufficiently grow on the polymeric ulvan, resulting in the first application of the ulvan-enzymatic toolbox originating from *F. agariphila* KMM3901^T.

Initially 11 different microorganisms, e.g. *Bacillus subtilis*, *Escherichia coli* or *Pseudomonas putida*, were analyzed for their capability to grow on the ulvan hydrolysate and the ulvan containing monosaccharides. The hydrolysates resulted from the digestion of ulvan using ulvanolytic cascade enzymes which were characterized in the Reisky *et al.* study and reported in **Article II**. The growth in M9-mineral medium with a variation of carbon source revealed *Bacillus licheniformis* DSM13 to be a suitable candidate since the strain was growing on the hydrolysate and accepted the ulvan containing monosaccharides as sole carbon source. Further investigations of the oligosaccharide consumption revealed that

predigestion of ulvan by P30_PL28 and P33_GH105, two enzymes from *F. agariphila*, already suits as a sufficient carbon source for this microbial strain. In addition, the general impact of the strain to degrade ulvan and its oligosaccharide was analyzed via a proteogenomic approach, which indicated putative xylosidase activity regarding the xylose-containing ulvan oligosaccharides.

B. licheniformis was selected to be further developed to a self-sufficient strain. Therefore, the expression and secretion of the two initial ulvanolytic enzymes P30_PL28 and P33_GH105 were investigated in *Bacillus* sp. The host-vector toolboxes of Krüger *et al.*^[98] for *Bacillus* sp. were used (Fig. 8a). We investigated *B. subtilis* JK136 and *B. licheniformis* MW3 as expression hosts. In the initial experiments, *Bacillus subtilis* JK136 actively produced the enzymes using the high-copy vector pMSE3, but they were not present in the *Bacillus licheniformis* MW3 strain. Finally, switching to the protease deficient *B. licheniformis* strain SH006 led to a functional expression of both enzymes. Once the expression of the individual enzymes was established, a co-expression system was designed. The co-expression cassette pBE-S_PL28-GH105 was created, and its protein expression was compared with the single pMSE3_PL28 and pMSE3_GH105 plasmid constructs in *B. licheniformis* SH006, while grown in simulated fed-batch conditions (EnpressoB). The expression was verified via the ulvan lyase assay, the thiobarbituric acid assay and using C-PAGE.

In the final experiment, the single and co-expression constructs were cultivated in M9-mineral medium with D-glucose as growth control, without a carbon source as negative control and with ulvan as sole carbon source investigating the potential of the self-sufficient strain. In comparison with the single construct strains, in which the CAZymes were functionally expressed (Fig. 8c) but did not enable further growth, the co-expression strain *B. licheniformis* SH006 pBE-S_PL28-GH105 was able to grow on ulvan as sole carbon source (Fig. 8b). This was complementary to the growth analysis with the prehydrolyzed ulvan, showing once more that the analogous of the *F. agariphila* enzymes PL28 and GH105 are missing in the *B. licheniformis* strain.

Overall, we successfully generated a self-sufficient *B. licheniformis* strain, which is the first step for development of a new ulvan based biorefinery process. Additionally, we highlighted a suitable metabolic engineering strategy for future strain development in *Bacillus* sp.

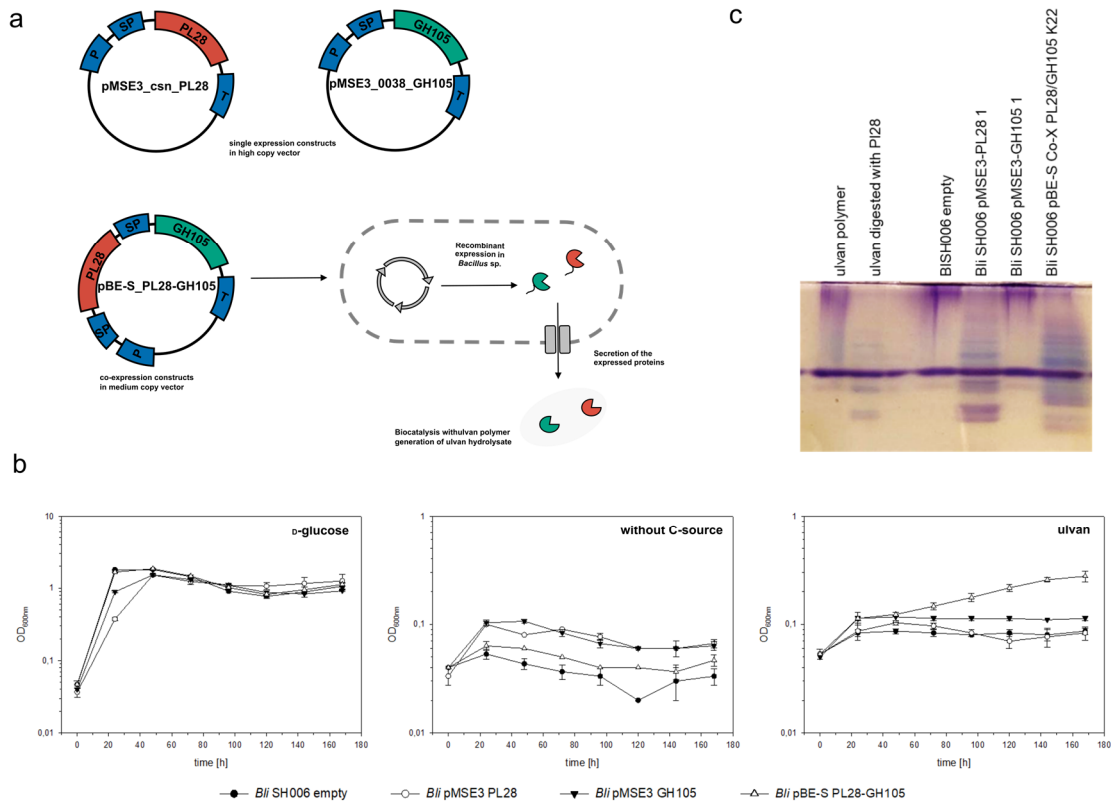


Figure 8: Summary of the self-sufficient *Bacillus* species strain. (a) General concept of CAZyme expression and vector design of the single constructs in the high-copy vectors pMSE3_csn_PL28 and pMSE3_0038_GH105 and the co-expression construct in a medium-copy vector pBE-S_PL28-GH105 (b) Growth of the *B. licheniformis* (Bli) SH006 strain with the three constructed vectors against an empty vector strain (*Bli* SH006 empty). The strain with the empty vector was cultivated in M9-mineral medium without carbon source (C-source), D-glucose and ulvan (c) C-PAGE from the culture supernatant of *B. licheniformis* cultivated with ulvan. The ulvan polysaccharide served as a negative control and ulvan digested with P30_PL28 as a positive control. The figure was adapted from **Article III**.

2.3 Xylan degradation in marine *Bacteroidetes* (Article IV)

Metagenomes of *Bacteroidetes* derived from the spring algae bloom in the year 2016, *Bacteroidetes* were analyzed bioinformatically and the polysaccharide utilization loci were mapped towards their polysaccharide target compounds^[43]. One of the abundant carbon sources in the bloom are xylose containing polysaccharide^[21]. Out of these metagenome datasets, the *Flavimarina* sp. Hel_I_48 was chosen to be the model organism for the investigation of the marine *Bacteroidetes* xylan degradation. The organism contains 15 PULs and encodes two xylan-degradation gene clusters with multiple SusC/D-like-pairs, PUL I (P162_RS02310-RS02390) and PUL II (P162_RS04015-RS04080). The main target of this study was to investigate the specific xylan PUL targets and the putative role of multiple transport proteins. A multidisciplinary approach of proteomic, comparative genomic and biochemical characterization of proteins, enabled to illuminate the marine xylan degradation in *Bacteroidetes*.

2.3.1 Proteomic results reveal PUL target specialization

To identify the xylan structures targeted by the individual PULs, *Flavimarina* sp. was cultivated on different available xylan sources, covering common motifs of arabino-, glucuronyl-, and acetate modifications of 1,4-linked xylan known from terrestrial xylans (beechwood xylan, BX; rye arabinoxylan, RAX; wheat arabinoxylan medium viscosity, WMX) and homoxylans from marine macroalgae which are either 1,3-linked (*Caulerpa prolifera*, CPX), or 1,3/1,4-linked (*Palmaria palmata*, PPX). The resulting expression pattern can indicate the proteins involved in the degradation of the corresponding polysaccharide (Fig. 9). The expression of SusC/D like proteins can be seen as an indicator for substrate acceptance, which is due to its role to recognize and uptake polysaccharides. Overall, *Flavimarina* sp. contains 6 SusC/D-like pairs, where 4 of them are encoded in PUL I and 2 in PUL II. The SusD_I_1 was expressed at higher protein levels with all xylans, this can indicate the general recognition of this SusD-like protein regarding the xylan backbone. On the other hand, the other SusD proteins of PUL I SusD_I_2 and SusD_I_3 were especially upregulated by glucuronoxylan (BX). Additionally, the interaction of the SusD_I_2 was shown via affinity gel electrophoresis, as the presence of this substrates lead to a retention in the gel matrix. This leads to the putative target specialization regarding glucuronoxylan alike polysaccharide. The PUL II SusD-like proteins, SusD_II_1 and SusD_II_2, showed higher expression level with arabinoxylans indicating a substrate specificity regarding arabinoxylan like structures.

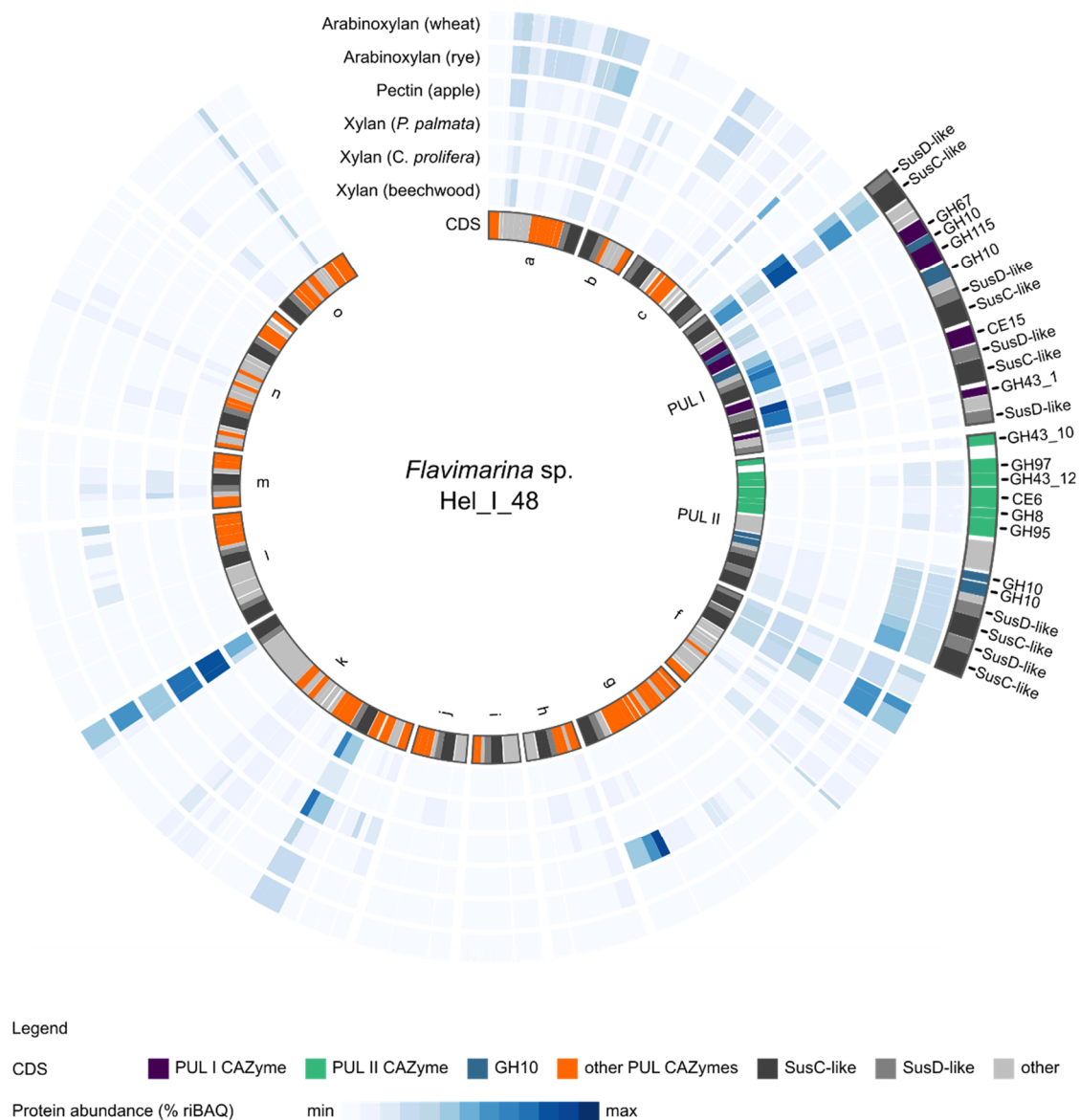


Figure 9: Proteomic profile of *Flavimarina* sp. Hel_I_48. Results of the proteomic analysis of *Flavimarina* sp. Hel_I_48 grown with different xylose-containing substrates as sole carbon sources; as control it was grown on pectin (apple). Every ring stands for a different substrate and the colour indicates the abundances (%riBAQ) of all PUL-encoded proteins, with a darker colour indicating a higher abundance. The two PULs investigated in this study are described outside the ring. The figure is derived from **Article IV**.

2.3.1 CAZyme activities underline the PULs target specialization

The enzymes (Table 1) of the two xylan PULs were recombinantly expressed in *E. coli* BL21(DE3) and purified to study their biochemical activity. Methodologies described in **Article II** were adapted to the xylan polysaccharides enabling the **Article IV** investigations. The enzyme activities were then analysed via ANTS-FACE, reducing end assay (DNS-assay), artificial substrates or by HPLC analysis. The annotated genome functions and the experimental proven functional annotations are summarized in Table 1.

Table 1: List of the enzymes analysed in the study.

Name	Gene Locus Taq	Annotation and modularity	Functional annotation
FI1_GH67	P162_RS02330	GH67	n.d.
FI2_GH10	P162_RS02335	GH10	<i>Exo</i> -1,4-xylanase
FI3_GH115	P162_RS02340	GH115 GH115	Alpha-(4- <i>O</i> -methyl)-glucuronidase (EC3.2.1.)
FI4_GH10	P162_RS02345	CBM4 GH10	<i>Endo</i> -xylanase
FI5_hyp	P162_RS02350	hyp	n.d.
FI6_CE15	P162_RS02365	CE15 CBM9	4- <i>O</i> -Methyl-glucuronyl methylesterase
FI7_GH43_1	P162_RS02380	GH43_1	n.d.
FI8A_CEnc*	P162_RS02385	Putative CE6	Acetyl-xylan esterase
FI8B_CEnc*	P162_RS02385	Putative CE6	Acetyl-xylan esterase
FII1A_CEnc	P162_RS04015	CE3 GH43_10	Acetyl-xylan esterase/xylosidase
FII1B_GH43_10	P162_RS04015	CE3 GH43_10	n. d.
FII2_GH97	P162_RS04020	GH97	α -D-galactosidase
FII3_GH43_12	P162_RS04025	GH43	α -L-arabinofuranosidase
FII4_CE6	P162_RS04030	CE6 CEnc CEnc (CEnc putative CE1)	Feruloyl xylan esterase /acetyl xylan esterase
FII5_GH8	P162_RS04035	GH8	<i>Exo</i> -xylanase
FII6_GH95	P162_RS04040	GH95	n.d.
FII7_GH10	P162_RS04050	GH10	n.d.
FII8_GH10	P162_RS04055	CBM4 GH10	<i>Endo</i> -1,4-xylanase
FII9_hyp	P162_RS04060	Hyp (DUF1735)	n.d.

The enzymatic activities show that the PULs contain on the one hand promiscuous and on the other hand highly adapted CAZymes regarding polysaccharide side chain modifications. The extracellular GH10 of the PULs (FI2_GH10, FI4_GH10, FII8_GH10) revealed a xylanase activity towards β -1,4-linked xylan without limitation by the side chain moieties (Fig. 10a). In contrast, the intracellular glycoside hydrolases are more specialized towards glucuronoxylan-like or arabinoxylan-like structures, e.g., the PUL I encodes for glucuronidases, FI3_GH115, and the PUL II for arabinases (FII3_GH43_12) (Fig. 10b). This complements the result from the proteomic approach defining glucuronoxylan-like substrates for PUL I and arabinoxylan-like substrates for PUL II. Beyond this, the absence of glycoside hydrolase activities and the activity of the galactosidase FII2_GH97 against an artificial substrate show that the actual marine target compound is much more complex than the structure of the chosen model substrates. In addition, the encoded carbohydrate esterases play a crucial role in removing acetate or ferulic acid esters enabling further degradation of the undecorated xylooligosaccharides by glycoside hydrolases e.g.,

2 Results

xylosidase FII1B_GH43_10 (Fig. 10a and c). Overall, the marine *Bacteroidetes* harbour a variety of CAZymes which are adapted towards different xylan structures, indicating a high diversity of xylan structures in the marine ecosystem.

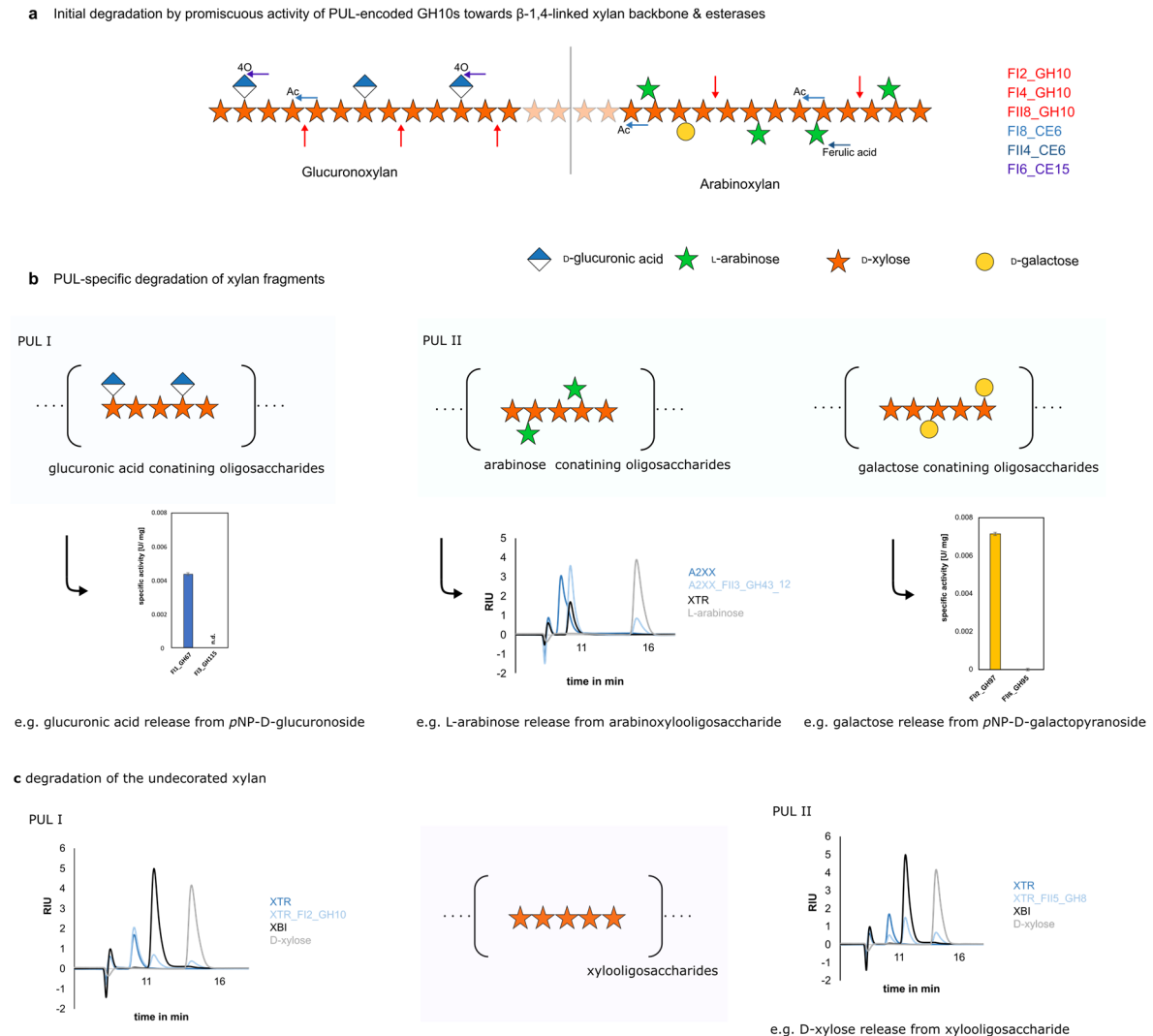


Figure 10: Schematic representation of the observed enzymatic reactions and the specific xylan degradation targets of PUL I and PUL II. The carbohydrate structure is assumed to be similar to glucuronoxylan and arabinoxylan known from grains^[99]. Summary of the initial degradation activities (a) from the GH10 xylanases (F12_GH10, F14_GH10 and F18_GH10) and the esterases F18_CE6 cleaving off acetate moieties, F14_CE6 cleaving acetate and ferulic acid modifications and F16_CE15 cleaving off 4-O-methylglucuronyl methylesters. The resulting xylooligosaccharides can be further degraded by PULs specific enzyme activities (b) glucuronidases for PUL I and arabinases and galactosidases for PUL II. The undecorated xylooligosaccharides can then be further decomposed by xylosidases (c). An exemplary enzymatic proof of each degradation step is either shown via HPLC results or via hydrolytic activity on p-NP-sugar substrates together with the schematic sugar backbone representation in (b) and (c).

2.3.2 Contextualisation of marine xylan polysaccharide degradation

The proteomic and enzymatic elucidation revealed a potential target specialization of the genome clusters regarding xylan-like structures, which share the 1,4-backbone specific GH10 CAZyme family. The GH10 shows activity against a multitude of different xylan substrates, which can be seen as a marker for xylan degradation. In addition, the *Flavimarina* sp. PUL CAZyme modularity was compared in the marine, terrestrial and marine database to elucidate the abundance of the degradation pathways. While a lot of organisms have xylan PULs like *Flavimarina* sp., containing the basic set of CAZymes GH43, GH67 and GH10 for xylan degradation, further CAZymes can be found in separate PUL structures or mixed PUL organisations, showing that there are individual genomic structures and niche adaptations of the strains regarding xylose-containing polysaccharides in the ocean. While they can be mostly found in the marine habitat, only a lower number of gene clusters were identified with a similar genomic repertoire as in the *Flavimarina* sp. PULs. This leads to the assumption that the cluster organisation of the xylan degradation clusters is a rather unique adaptation towards a variety of different xylose-containing polysaccharides in the ocean.

The main characterization of the xylan utilization was performed with terrestrial polysaccharides and marine macroalgal polysaccharides. As *Flavimarina* sp. was harvested during a diatom bloom it is likely that diatoms contain the actual target polysaccharides. The precise structure of the polysaccharides from diatom origin has not been elucidated so far, but e.g., *Thallessiosira weissflogii* contains a high amount of xylose, suggesting the occurrence of xylan. Microarray assays also showed that *T. weissflogii* contains 1,4-linked xylans. Therefore, the chosen biomass is sufficient to represent different types of xylan as previous studies shown the occurrence of 1,4-linked xylan in diatom blooms ^[21]. The polysaccharide arabinogalactan has been detected in algae blooms^[100]. *Flavimarina* harbours the genetic potential to degrade a wide range of xylose containing polymers.

2.4 Microbial adaption mechanism in polysaccharide utilization (**Articles V & VI**)

Marine carbohydrates are highly diverse in their structure and modifications, which was highlighted in review **Article I**. The marine polysaccharide porphyran from red algae is a galactan which is highly methylated. Marine organisms adapted their genetic repertoire for the degradation of complex polysaccharides, like those containing methylated sugars. In the study of Reisky *et al* the first marine cytochrome P450 monooxygenases involved in the demethylation of sugars were described. During the demethylation reaction formaldehyde

is accumulating which is toxic for the cells due to its high reactivity as an electrophile^[101]. Beside the adaptation of the organisms towards degradation of marine polysaccharides they also needed to develop formaldehyde detoxification mechanisms. The elucidation of the role of different formaldehyde detoxification principles and putatively involved enzymes has been the target of **Article V** and **Article VI**.

2.3.1 The RuMP pathway is involved in the degradation of methylated sugars (**Article V**)

The model organisms *Zobellia galactanivorans* Dsij^T and *Formosa agariphila* KMM3901^T contain both cytochrome P450 monooxygenase demethylation enzymes ^[76,102], associated with CAZymes (**Article VI**). The aim of **Article V** was to elucidate how the organism copes with the associated formaldehyde accumulation. Both organisms were grown in the presence of different concentrations of formaldehyde. Showing that *F. agariphila* is tolerating less formaldehyde (~100 µM) than *Z. galactanivorans* (<500 µM), leading towards a greater adaptation of the *Z. galactanivorans* strain regarding formaldehyde.

The genetic comparison of different known formaldehyde detoxification principles revealed that they share the serine and tetrahydrofolate pathway, but that *Z. galactanivorans* contains a ribulose monophosphate (RuMP)-pathway which *F. agariphila* lacks. In the RuMP pathway the formaldehyde is inverted to ribulose-5-phosphate to produce fructose-6-phosphate, catalyzed by a 3-hexulose-6-phosphate synthase (HPS) and a 6-phospho-3-hexuloisomerase (PHI) (Fig. 11b). The biochemical investigation of the recombinantly produced *Z. galactanivorans* HPS and PHI proved the impact in formaldehyde detoxification, while the generated knock-out of HPS and PHI genes led to a lack of tolerance of formaldehyde, which again proved its importance and physiological role for *Z. galactanivorans*. Additionally, comparative genomic analysis revealed that the RuMP pathway is rather an adaptation to specific bacteria living on multicellular algae and does not occur in every marine *Bacteroidetes* which contain marine cytochrome P450 monooxygenases.

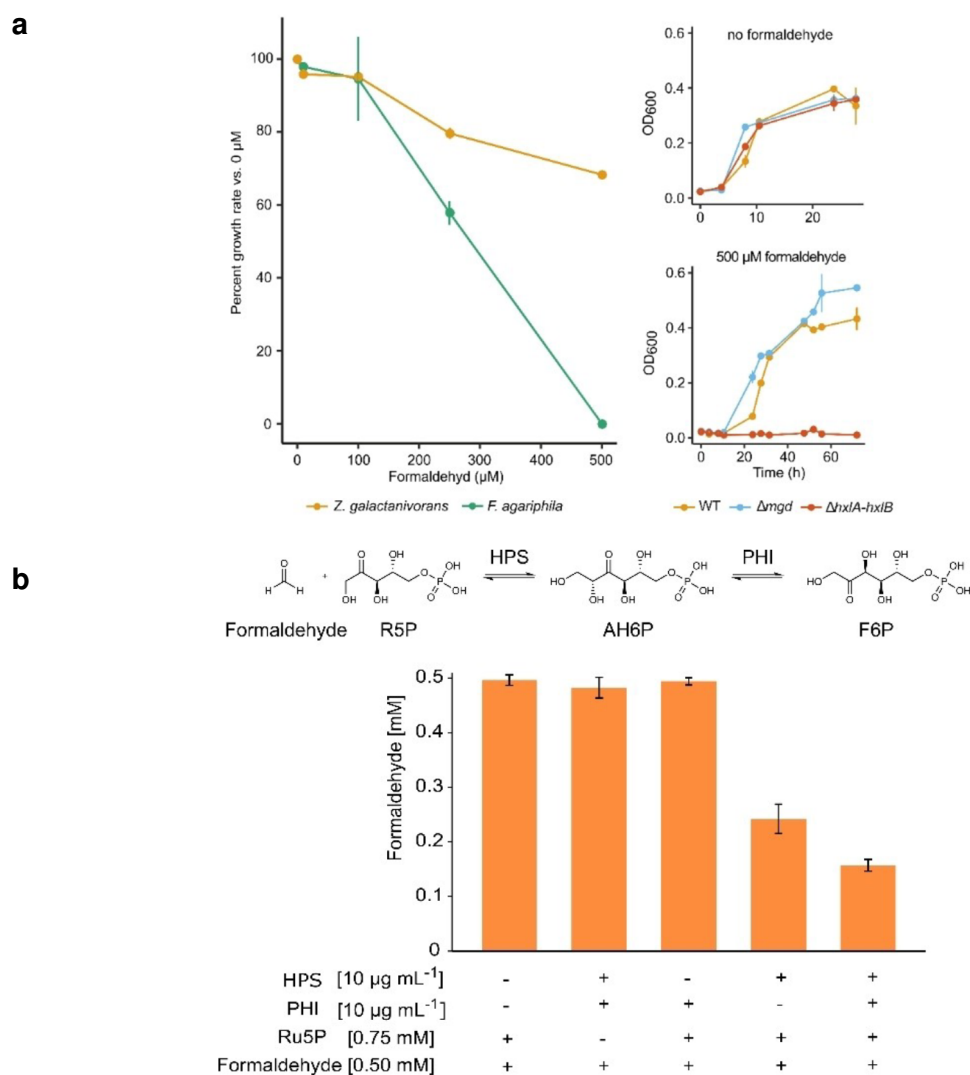


Figure 11: RuMP-pathway elucidation in marine *Bacteroidetes*. (a) *F. agariphila* and *Z. galactanivorans* were grown in the presence of formaldehyde. Knock-out of the genes encoding the RuMP-pathway 3-hexulose-6-phosphate synthase *hxIA* (HPS) and 6-phospho-3-hexulose isomerase *hxIB* (PHI) in *Z. galactanivorans* were created and these were grown with formaldehyde in comparison with the cytochrome P450 monooxygenase knock-out strain (Δmgd) and the wild-type strain (WT). (b) Shows the RuMP pathway reaction and the biochemical proven formaldehyde detoxification of the recombinant expressed proteins in combination. The formaldehyde concentration was measured via the Nash reagent. The figure was adapted from **Article V**.

2.3.2 Role of novel alcohol dehydrogenases in polysaccharide degradation (**Article VI**)

Beside the discovery of the RuMP pathway in **Article V**, comparative genomic studies revealed the co-occurrence of an alcohol dehydrogenase (ADH) and an esterase in the demethylation cluster in marine *Bacteroidetes* (Fig. 12a). The genetic proximity might indicate a successive physiological role. The initial hypothesis was that those two enzymes are involved in a glutathione cofactor-dependent type of formaldehyde detoxification,

decomposing the formaldehyde from the oxidative demethylation^[103]. Due to that, the aim of **Article VI** was to find out the biochemical characteristics of the ADH (from the model organism *Z. galactanivorans* and *F. agariphila*) and to investigate its role in polysaccharide utilization and oxidative demethylation. Therefore, the crystal structure of the ZoADH and FoADH enzymes were solved and biochemical properties were determined. The activity profile was elucidated regarding different formaldehyde substrates, revealing an activity towards aromatic aldehydes, while showing a narrow active site. Unfortunately, no activity with formaldehyde was found. Additionally, possible known co-factors were investigated due to the potential similarities regarding glutathione-dependent formaldehyde detoxification mechanism. In the marine and terrestrial realm different potential co-factors are known, like bacillithiol or mycothiol, but both showed no impact on the activity of the ADH. Sequence similarity analyses and corresponding genome neighbourhood analyses revealed that the ZoADH and FoADH do not cluster with the majority of cofactor-dependent ADHs, indicating a co-factor independency. Even though the impact in formaldehyde detoxification could not be proven, the role in polysaccharide utilization was clearly stated in **Article VI**. The generation of knock-out strains lacking the ZoADH gene in *Z. galactanivorans* revealed that without the ZoADH enzyme the strain is incapable of growing on methylated sugars (Me6Gal). The ADH thus must be involved in the metabolism of the methylated sugars, which states the ZoADH and FoADH are the first ADHs involved in polysaccharide utilization as described in **Article VI**, and this claims a potential novel auxiliary activity enzyme.

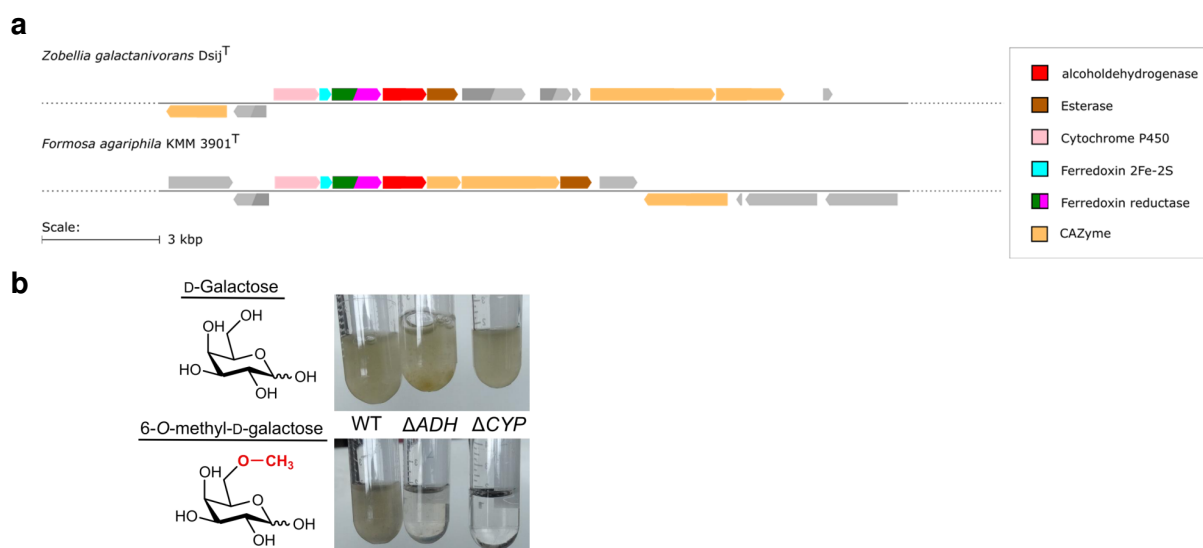


Figure 12: Marine *Bacteroidetes* demethylation cluster and *Z. galactanivorans* growth analysis. (a) Genome neighbourhood diagram of the marine P450 of *Z. galactanivorans* and *F. agariphila*^[104]. (b) The knockout of the ADH and P450(CYP) respectively, lead to a loss of growth on methylated sugar indicating clear association with the carbohydrate utilization of 6-O-D-methyl-galactose of the ADH.

3 Summary

Algae play an essential role in the overall carbon cycle. They fix CO₂ into polysaccharides which can either serve as algae storage compound, carbon sink, exopolysaccharide matrix or can be converted by heterotrophic bacteria. The heterotrophic bacteria from the phylum *Bacteroidetes* are known to be specialists in carbohydrate degradation. They encode carbohydrate-active enzymes in a polysaccharide utilization gene cluster (PUL). The genes encoded in these PULs work successively, while being strongly adapted towards their target polysaccharide. The polysaccharide structure and composition of each macro- and microalgae hardly vary in species and seasons. This leads towards different adaptation mechanisms of a given strain regarding their structure.

The current state of the art in the enzymatic marine polysaccharide degradation and the diversity of marine polysaccharides have been summarized in the review **Article I**. Therewith this review summarize the marine polysaccharide degradation principles relevant for **Articles II** to **VI** and highlights the elucidated ulvan degradation cascade from **Article II**. New complex enzymatic cascades include novel enzymatic functions, which assist in the enzymatic degradation or in the metabolism of the sugars (**Articles V, VI**). The now enlarged ulvanolytic toolbox and the discovery of novel enzyme activities, like the oligosaccharide dehydratase or the discovery of ADHs in polysaccharide utilization, expands the current knowledge how marine *Bacteroidetes* can metabolize marine polysaccharides. These principles gained enhance the knowledge about the marine carbon cycle and enables the access of novel biomasses and its transformation into valuable compounds for pharmaceutical and biotechnological applications. **Article III** provides an example of potential applications of the enzymatic toolboxes described in **Article II**. This algae biomass can thus serve as biotechnological microbial feedstock. In **Article III B**, *licheniformis* was found be able to use ulvan hydrolysate as sole carbon source. Later, this bacterium was developed into a self-sufficient strain, capable to produce ulvanolytic enzyme and to grow on the ulvan polymer biomass.

Overall, this thesis focused on the polysaccharide degradation principles for ulvan and xylans from marine *Bacteroidetes*, also in terms of their use, showing the first application process from the ulvan-degradation cascade. The articles highlight the diversity of the marine polysaccharides in the ocean and discuss microbial adaptation mechanisms including new enzyme classes.

4 References

- [1] S. Das, P. S. Lyla, S. A. Khan, *Curr. Sci.* **2006**, 90, 1325–1335.
- [2] W. M. Post, T.-H. Peng, W. R. Emanuel, A. W. King, V. H. Dale, D. L. DeAngelis, *Am. Sci.* **1990**, 78, 310–326.
- [3] C. B. Field, *Science* **1998**, 281, 237–240.
- [4] V. Evangelista, A. M. Frassanito, V. Passarelli, L. Barsanti, P. Gualtieri, *Photochem. Photobiol.* **2006**, 82, 1039–1046.
- [5] M. T. Cesário, M. M. R. da Fonseca, M. M. Marques, M. C. M. D. de Almeida, *Biotechnol. Adv.* **2018**, 36, 798–817.
- [6] H. Teeling, B. M. Fuchs, C. M. Bennke, K. Krüger, M. Chafee, L. Kappelmann, G. Reintjes, J. Waldmann, C. Quast, F. O. Glöckner, J. Lucas, A. Wichels, G. Gerds, K. H. Wiltshire, R. I. Amann, *eLife* **2016**, 5, e11888.
- [7] S. Emil Ruff, D. Probandt, A.-C. Zinkann, M. H. Iversen, C. Klaas, L. Würzberg, N. Krombholz, D. Wolf-Gladrow, R. Amann, K. Knittel, *Deep Sea Res. Part II Top. Stud. Oceanogr.* **2014**, 108, 6–16.
- [8] H. Teeling, B. M. Fuchs, D. Becher, C. Klockow, A. Gardebrecht, C. M. Bennke, M. Kassabgy, S. Huang, A. J. Mann, J. Waldmann, M. Weber, A. Klindworth, A. Otto, J. Lange, J. Bernhardt, C. Reinsch, M. Hecker, J. Peplies, F. D. Bockelmann, U. Callies, G. Gerds, A. Wichels, K. H. Wiltshire, F. O. Glöckner, T. Schweder, R. Amann, *Science* **2012**, 336, 608–611.
- [9] D. K. Steinberg, M. R. Landry, *Annu. Rev. Mar. Sci.* **2017**, 9, 413–444.
- [10] Z. A. Popper, G. Michel, C. Hervé, D. S. Domozych, W. G. T. Willats, M. G. Tuohy, B. Kloareg, D. B. Stengel, *Annu. Rev. Plant Biol.* **2011**, 62, 567–590.
- [11] K. J. Niklas, *BioScience* **2004**, 54, 831–841.
- [12] G. Michel, T. Tonon, D. Scornet, J. M. Cock, B. Kloareg, *New Phytol.* **2010**, 188, 82–97.
- [13] B. Kloareg, R. S. Quatrano, *Oceanogr. Mar. Biol. Annu. Rev.* **1988**, 26, 259–315.
- [14] S. B. Kroken, L. E. Graham, M. E. Cook, *Am. J. Bot.* **1996**, 83, 1241–1254.
- [15] P. T. Martone, J. M. Estevez, F. Lu, K. Ruel, M. W. Denny, C. Somerville, J. Ralph, *Curr. Biol.* **2009**, 19, 169–175.
- [16] W. Helbert, *Front. Mar. Sci.* **2017**, 4, 1–10.
- [17] B. Subhadra, M. Edwards, *Energy Policy* **2010**, 38, 4897–4902.

-
- [18] R. Taylor, R. L. Fletcher, J. A. Raven, *Bot. Mar.* **2001**, 44, 327–336.
- [19] R. A. Cohen, P. Fong, *Ecol. Appl.* **2006**, 16, 1405–1420.
- [20] M. Bäumgen, T. Dutschei, U. T. Bornscheuer, *ChemBioChem* **2021**, 22, 2247–2256.
- [21] S. Vidal-Melgosa, A. Sichert, T. B. Francis, D. Bartosik, J. Niggemann, A. Wichels, W. G. T. Willats, B. M. Fuchs, H. Teeling, D. Becher, T. Schweder, R. Amann, J.-H. Hehemann, *Nat. Commun.* **2021**, 12, 1150.
- [22] J. T. Kidgell, M. Magnusson, R. de Nys, C. R. K. Glasson, *Algal Res.* **2019**, 39, 101422.
- [23] M. Lahaye, A. Robic, *Biomacromolecules* **2007**, 8, 1765–1774.
- [24] A. Robic, C. Gaillard, J.-F. Sassi, Y. Lerat, M. Lahaye, *Biopolymers* **2009**, 91, 652–664.
- [25] S. M. Cardoso, L. G. Carvalho, P. J. Silva, M. S. Rodrigues, O. R. P. and L. Pereira, *Curr. Org. Chem.* **2014**, 18, 896–917.
- [26] Y. Hsieh, P. J. Harris, *Polymers* **2019**, 11, 354.
- [27] A. Ebringerová, *Macromol. Symp.* **2005**, 232, 1–12.
- [28] Frei Eva, Preston Reginald Dawson, *Proc. R. Soc. Lond. B Biol. Sci.* **1964**, 160, 293–313.
- [29] J. K. Jensen, M. Busse-Wicher, C. P. Poulsen, J. U. Fangel, P. J. Smith, J.-Y. Yang, M.-J. Peña, M. H. Dinesen, H. J. Martens, M. Melkonian, G. K.-S. Wong, K. W. Moremen, C. G. Wilkerson, H. V. Scheller, P. Dupree, P. Ulvskov, B. R. Urbanowicz, J. Harholt, *New Phytol.* **2018**, 218, 1049–1060.
- [30] I. M. Mackie, E. Percival, *J. Chem. Soc. Resumed* **1959**, 1, 1151–1156.
- [31] M. Lahaye, C. Rondeau-Mouro, E. Deniaud, A. Buléon, *Carbohydr. Res.* **2003**, 338, 1559–1569.
- [32] E. G. V. Percival, S. K. Chanda, *Nature* **1950**, 166, 787–788.
- [33] J. R. Turvey, E. L. Williams, *Phytochemistry* **1970**, 9, 2383–2388.
- [34] A. S. Cerezo, *Carbohydr. Res.* **1972**, 22, 209–211.
- [35] E. Deniaud, B. Quemener, J. Fleurence, M. Lahaye, *Int. J. Biol. Macromol.* **2003**, 33, 9–18.
- [36] T. B. Francis, D. Bartosik, T. Sura, A. Sichert, J.-H. Hehemann, S. Markert, T. Schweder, B. M. Fuchs, H. Teeling, R. I. Amann, D. Becher, *ISME J.* **2021**, 15, 2336–2350.
-

- [37] G. Huang, S. Vidal-Melgosa, A. Sichert, S. Becker, Y. Fang, J. Niggemann, M. H. Iversen, Y. Cao, J. Hehemann, *Limnol. Oceanogr.* **2021**, 66, 3768–3782.
- [38] K. Krüger, M. Chafee, T. B. Francis, T. G. del Rio, D. Becher, T. Schweder, R. I. Amann, H. Teeling, *ISME J.* **2019**, 13, 2800–2816.
- [39] F. Thomas, J.-H. Hehemann, E. Rebuffet, M. Czjzek, G. Michel, *Front. Microbiol.* **2011**, 2, DOI 10.3389/fmicb.2011.00093.
- [40] J. M. Grondin, K. Tamura, G. Déjean, D. W. Abbott, H. Brumer, *J. Bacteriol.* **2017**, 199, e00860-16.
- [41] R. Munoz, R. Rosselló-Móra, R. Amann, *Syst. Appl. Microbiol.* **2016**, 39, 281–296.
- [42] M. K. Bjursell, E. C. Martens, J. I. Gordon, *J. Biol. Chem.* **2006**, 281, 36269–36279.
- [43] L. Kappelmann, K. Krüger, J.-H. Hehemann, J. Harder, S. Markert, F. Unfried, D. Becher, N. Shapiro, T. Schweder, R. I. Amann, H. Teeling, *ISME J.* **2019**, 13, 76–91.
- [44] L. Reisky, A. Précoux, M. K. Zühlke, M. Bäumgen, C. S. Robb, N. Gerlach, T. Roret, C. Stanetty, R. Larocque, G. Michel, S. Tao, S. Markert, F. Unfried, M. D. Mihovilovic, A. Trautwein-Schulz, D. Becher, T. Schweder, U. T. Bornscheuer, J.-H. Hehemann, *Nat Chem Biol* **2019**, 15, 803–812.
- [45] G. Reintjes, C. Arnosti, B. M. Fuchs, R. Amann, *ISME J.* **2017**, 11, 1640–1650.
- [46] G. Reintjes, C. Arnosti, B. Fuchs, R. Amann, *ISME J.* **2019**, 13, 1119–1132.
- [47] S. Rakoff-Nahoum, M. J. Coyne, L. E. Comstock, *Curr. Biol.* **2014**, 24, 40–49.
- [48] Y. Yin, X. Mao, J. Yang, X. Chen, F. Mao, Y. Xu, *Nucleic Acids Res.* **2012**, 40, W445–W451.
- [49] V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho, B. Henrissat, *Nucleic Acids Res.* **2014**, 42, D490-495.
- [50] G. Davies, B. Henrissat, *Structure* **1995**, 3, 853–859.
- [51] D. E. Koshland Jr., *Biol. Rev.* **1953**, 28, 416–436.
- [52] P. N. Collén, A. Jeudy, J.-F. Sassi, A. Groisillier, M. Czjzek, P. M. Coutinho, W. Helbert, *J. Biol. Chem.* **2014**, 289, 6199–6211.
- [53] V. Lombard, T. Bernard, C. Rancurel, H. Brumer, P. M. Coutinho, B. Henrissat, *Biochem. J.* **2010**, 432, 437–444.

-
- [54] P. Gacesa, FEBS Lett. **1987**, 212, 4.
- [55] M.-L. Garron, M. Cygler, Glycobiology **2010**, 20, 1547–1573.
- [56] V. L. Y. Yip, S. G. Withers, Curr. Opin. Chem. Biol. **2006**, 10, 147–155.
- [57] T. Ulaganathan, M. T. Boniecki, E. Foran, V. Buravenkov, N. Mizrachi, E. Banin, W. Helbert, M. Cygler, ACS Chem. Biol. **2017**, 12, 1269–1280.
- [58] T. Ulaganathan, W. Helbert, M. Kopel, E. Banin, M. Cygler, J. Biol. Chem. **2018**, 293, 4026–4036.
- [59] T. Ulaganathan, E. Banin, W. Helbert, M. Cygler, J. Biol. Chem. **2018**, 293, 11564–11573.
- [60] L. Reisky, C. Stanetty, M. D. Mihovilovic, T. Schweder, J.-H. Hehemann, U. T. Bornscheuer, Appl. Microbiol. Biotechnol. **2018**, 102, 6987–6996.
- [61] M. Lahaye, Carbohydr. Res. **1998**, 314, 1–12.
- [62] J. R. Somoza, S. Menon, H. Schmidt, D. Joseph-McCarthy, A. Dessen, M. L. Stahl, W. S. Somers, F. X. Sullivan, Structure **2000**, 8, 123–135.
- [63] S. T. M. Allard, K. Beis, M.-F. Giraud, A. D. Hegeman, J. W. Gross, R. C. Wilmouth, C. Whitfield, M. Graninger, P. Messner, A. G. Allen, D. J. Maskell, J. H. Naismith, Structure **2002**, 10, 81–92.
- [64] B. L. Cantarel, P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, Nucleic Acids Res. **2009**, 37, D233–D238.
- [65] L. P. Christov, B. A. Prior, Enzyme Microb. Technol. **1993**, 15, 460–475.
- [66] G. V. Pereira, A. M. Abdel-Hamid, S. Dutta, C. N. D'Alessandro-Gabazza, D. Wefers, J. A. Farris, S. Bajaj, Z. Wawrzak, H. Atomi, R. I. Mackie, E. C. Gabazza, D. Shukla, N. M. Koropatkin, I. Cann, Nat. Commun. **2021**, 12, 459.
- [67] A. M. Nakamura, A. S. Nascimento, I. Polikarpov, Biotechnol. Res. Innov. **2017**, 1, 35–51.
- [68] P. Biely, Biotechnol. Adv. **2012**, 30, 1575–1588.
- [69] C. De Santi, O. A. Gani, R. Helland, A. Williamson, Sci. Rep. **2017**, 7, 17278.
- [70] M. D. Charavgi, M. Dimarogona, E. Topakas, P. Christakopoulos, E. D. Chrysina, Acta Crystallogr. D Biol. Crystallogr. **2013**, 69, 63–73.
- [71] A. B. Boraston, A. L. Creagh, Md. M. Alam, J. M. Kormos, P. Tomme, C. A. Haynes, R. A. J. Warren, D. G. Kilburn, Biochemistry **2001**, 40, 6240–6247.
- [72] S. R. Handson, M. D. Best, C.-H. Wong, Angew. Chem. Int. Ed. **2004**, 43, 5736–5763.
- [73] O. Berteau, B. Mulloy, Glycobiology **2003**, 13, 29R-40R.
-

- [74] A. Sichert, C. H. Corzett, M. S. Schechter, F. Unfried, S. Markert, D. Becher, A. Fernandez-Guerra, M. Liebeke, T. Schweder, M. F. Polz, J.-H. Hehemann, *Nat. Microbiol.* **2020**, 5, 1026–1039.
- [75] E. Ficko-Blean, A. Préchoux, F. Thomas, T. RoCHAT, R. Larocque, Y. Zhu, M. Stam, S. Génicot, M. Jam, A. Calteau, B. Viart, D. Ropartz, D. Pérez-Pascual, G. Correc, M. Matard-Mann, K. A. Stubbs, H. Rogniaux, A. Jeudy, T. Barbeyron, C. Médigue, M. Czjzek, D. Vallenet, M. J. McBride, E. Duchaud, G. Michel, *Nat. Commun.* **2017**, 8, 1685.
- [76] L. Reisky, H. C. Büchsenschütz, J. Engel, T. Song, T. Schweder, J.-H. Hehemann, U. T. Bornscheuer, *Nat. Chem. Biol.* **2018**, 14, 342–344.
- [77] G. Correc, J.-H. Hehemann, M. Czjzek, W. Helbert, *Carbohydr. Polym.* **2011**, 83, 277–283.
- [78] S. Brott, F. Thomas, M. Behrens, K. Methling, D. Bartosik, T. Dutschei, M. Lalk, G. Michel, T. Schweder, U. T. Bornscheuer, *ChemBioChem* **2022**, e202200269.
- [79] C. Filote, S. C. R. Santos, V. I. Popa, C. M. S. Botelho, I. Volf, *Environ. Chem. Lett.* **2020**, 969–1000.
- [80] N. V. Thomas, S.-K. Kim, *Mar. Drugs* **2013**, 11, 146–164.
- [81] A. J. Smit, *J. Appl. Phycol.* **2004**, 16, 245–262.
- [82] A. Silva, S. A. Silva, M. Carpena, P. Garcia-Oliveira, P. Gullón, M. F. Barroso, M. A. Prieto, J. Simal-Gandara, *Antibiotics* **2020**, 9, 642.
- [83] T.-S. Vo, D.-H. Ngo, S.-K. Kim, *Process Biochem.* **2012**, 47, 386–394.
- [84] M. F. De Jesus Raposo, A. M. B. De Moraes, R. M. S. C. De Moraes, *Mar. Drugs* **2015**, 13, 2967–3028.
- [85] M. Alvarado-Morales, I. B. Gunnarsson, I. A. Fotidis, E. Vasilakou, G. Lyberatos, I. Angelidaki, *Algal Res.* **2015**, 9, 126–132.
- [86] A. Noreen, K. M. Zia, M. Zuber, M. Ali, M. Mujahid, *Int. J. Biol. Macromol.* **2016**, 86, 937–949.
- [87] L. M. L. Laurens, J. Markham, D. W. Templeton, E. D. Christensen, S. V. Wychen, E. W. Vadellius, M. Chen-Glasser, T. Dong, R. Davis, P. T. Pienkos, *Energy Environ. Sci.* **2017**, 10, 1716–1738.
- [88] R. Araújo, F. Vázquez Calderón, J. Sánchez López, I. C. Azevedo, A. Bruhn, S. Fluch, M. Garcia Tasende, F. Ghaderiardakani, T. Ilmjärv, M. Laurans, M.

- Mac Monagail, S. Mangini, C. Peteiro, C. Rebours, T. Stefansson, J. Ullmann, *Front. Mar. Sci.* **2021**, 7.
- [89] V. Smetacek, A. Zingone, *Nature* **2013**, 504, 84–88.
- [90] R. H. Charlier, P. Morand, C. W. Finkl, *Int. J. Environ. Stud.* **2008**, 65, 191–208.
- [91] P. Schiener, K. D. Black, M. S. Stanley, D. H. Green, *J. Appl. Phycol.* **2015**, 27, 363–373.
- [92] L. M. L. Laurens, M. Lane, R. S. Nelson, *Trends Biotechnol.* **2020**, 38, 1232–1244.
- [93] K. A. Jung, S.-R. Lim, Y. Kim, J. M. Park, *Bioresour. Technol.* **2013**, 135, 182–190.
- [94] F. Fernand, A. Israel, J. Skjermo, T. Wichard, K. R. Timmermans, A. Golberg, *Renew. Sustain. Energy Rev.* **2017**, 75, 35–45.
- [95] A. J. Wargacki, E. Leonard, M. N. Win, D. D. Regitsky, C. N. S. Santos, P. B. Kim, S. R. Cooper, R. M. Raisner, A. Herman, A. B. Sivitz, A. Lakshmanaswamy, Y. Kashiyama, D. Baker, Y. Yoshikuni, *Science* **2012**, 335, 308–313.
- [96] M. Enquist-Newman, A. M. E. Faust, D. D. Bravo, C. N. S. Santos, R. M. Raisner, A. Hanel, P. Sarvabhowman, C. Le, D. D. Regitsky, S. R. Cooper, L. Peereboom, A. Clark, Y. Martinez, J. Goldsmith, M. Y. Cho, P. D. Donohoue, L. Luo, B. Lamberson, P. Tamrakar, E. J. Kim, J. L. Villari, A. Gill, S. A. Tripathi, P. Karamchedu, C. J. Paredes, V. Rajgarhia, H. K. Kotlar, R. B. Bailey, D. J. Miller, N. L. Ohler, C. Swimmer, Y. Yoshikuni, *Nature* **2014**, 505, 239–243.
- [97] V. R. Konasani, C. Jin, N. G. Karlsson, E. Albers, *Sci. Rep.* **2018**, 8, 14713.
- [98] A. Krüger, N. Welsch, A. Dürwald, H. Brundiek, R. Wardenga, H. Piascheck, H. G. Mengers, J. Krabbe, S. Beyer, J. F. Kabisch, L. Popper, T. Hübel, G. Antranikian, T. Schweder, *Appl. Microbiol. Biotechnol.* **2022**, 106, 5137–5151.
- [99] A. Rogowski, J. A. Briggs, J. C. Mortimer, T. Tryfona, N. Terrapon, E. C. Lowe, A. Baslé, C. Morland, A. M. Day, H. Zheng, T. E. Rogers, P. Thompson, A. R. Hawkins, M. P. Yadav, B. Henrissat, E. C. Martens, P. Dupree, H. J. Gilbert, D. N. Bolam, *Nat. Commun.* **2015**, 6, 7481.
- [100] M. R. Cases, C. A. Stortz, A. S. Cerezo, *Int. J. Biol. Macromol.* **1994**, 16, 93–97.

- [101] N. H. Chen, K. Y. Djoko, F. J. Veyrier, A. G. McEwan, *Front. Microbiol.* **2016**, 7.
- [102] C. S. Robb, L. Reisky, U. T. Bornscheuer, J.-H. Hehemann, *Biochem. J.* **2018**, 475, 3875–3886.
- [103] H. Achkor, M. Díaz, M. R. Fernández, J. A. Biosca, X. Parés, M. C. Martínez, *Plant Physiol.* **2003**, 132, 2248–2255.
- [104] R. Zallot, N. Oberg, J. A. Gerlt, *Biochemistry* **2019**, 58, 4169–4182.

Authors contributions

Article I: Marine polysaccharides: Occurrence, enzymatic degradation and utilization

M. Bäumgen*, T. Dutschei* and U. T. Bornscheuer, *ChemBioChem* **2021**, 22, 2247-2256.

UTB initiated the review and outlined the manuscript. MB summarized the individual enzymatic cascades and TD the ecological role, marine polysaccharide diversity and biotechnical applications. All authors have read and approved the final manuscript.

Article II A new carbohydrate-active oligosaccharide dehydratase is involved in the degradation of ulvan

M. Bäumgen*, T. Dutschei*, D. Bartosik, C. Suster, L. Reisky, N. Gerlach, C. Stanetty, M. D. Mihovilovic, T. Schweder, J. H. Hehemann and U. T. Bornscheuer, *J. Chem. Biol.* **2021**, 297, 101210

T. S., J.-H. H., and U. T. B. conceptualization; D. B., Christoph Suster, L. R., Christian Stanetty, and U. T. B. formal analysis; T. S., J.-H. H., and U. T. B. funding acquisition; M. B., T. D., Christoph Suster, L. R., N. G., Christian Stanetty, J.-H. H., and U. T. B. investigation; M. B., T. D., Christoph Suster, N. G., Christian Stanetty, and J.-H. H. methodology; D. B. software; M. D. M., J.-H. H., and U. T. B. supervision; J.-H. H. validation; D. B. visualization; M. B. and T. D. writing—original draft; M. B., T. D., D. B., Christoph Suster, L. R., N. G., Christian Stanetty, M. D. M., T. S., J.-H. H., and U. T. B. writing—review and editing.

Article III Metabolic engineering enables *Bacillus licheniformis* to grow on the marine polysaccharide ulvan

T. Dutschei, M. K. Zühlke, N. Welsch, T. Eisenack, M. Hilkmann, J. Krull, C. Stühle, S. Brott, A. Dürwald, L. Reisky, J.-H. Hehemann, D. Becher, T. Schweder and Uwe T. Bornscheuer, *Microb. Cell Fact.* **2022** (under revision)

UTB, LR and TS designed the study, supervised its execution and co-wrote the manuscript. TD wrote the main manuscript with the support of MKZ and NW. TD performed the initial screening of organism and growth and activity assays, with TE and CS. The proteome analysis was performed by TE and MKZ. DB coordinated MS measurements. The computational analysis was performed by MKZ, AD and SB. JK performed the sugar

monosaccharide composition analysis in the lab of JHH. SB expressed and purified the putative sulfatases and performed the activity assays. NW developed the *Bacillus* host-vector system and the strain *Bacillus* design with MH. All authors have read and approved the final manuscript.

Article IV Utilization of a diverse range of xylan structures from marine *Bacteroidetes*

T. Dutschei, I. Beidler, D. Bartosik, J.-M. Seeßelberg, M. Teune, M. Bäumgen, S. Q. Ferreira, J. Heldmann, F. Nagel, J. Krull, L. Berndt, K. Methling, M. Hein, D. Becher, P. Langer, M. Delcea, M. Lalk, M. Lammers, M. Höhne, J. H. Hehemann, T. Schweder and U. T. Bornscheuer, *ISME J.* **2022** (*under revision*)

The study was designed by UTB, TS, MB and TD. The main biochemical characterizations were performed by MB, SQF, JH and TD with support by JMS, MT, LB, KM in supervision of ML, MH and ML. The proteome was analyzed by IB for which DBe provided resources. DBa did the computational analysis. Additional analyses of the polysaccharides were performed by FN, JK in the labs of MD and JHH. TD wrote the main manuscript with support by DB, IB, JMS and TS. All authors have read and approved the final manuscript.

Article V Connecting algal polysaccharide degradation to formaldehyde detoxification

S. Brott, F. Thomas, M. Behrens, K. Methling, D. Bartosik, T. Dutschei, M. Lalk, G. Michel, T. Schweder, and U. T. Bornscheuer, *ChemBioChem* **2022** 23, e202200269

MG, TS and UTB initiated the study and directed the project. FT conducted the growth studies and created the knockout strains. SB and MB expressed and purified the enzymes and performed the biocatalysis. DB performed the computational analysis. SB prepared the main manuscript, which was revised by FT, MK, ML, TD, MG, TS and UTB and approved by all authors.

Article VI A unique alcohol dehydrogenase involved in algal sugar utilization by marine bacteria

S. Brott, K. H. Nam, F. Thomas, T. Dutschei, L. Reisky, M. Behrens, H. C. Grimm, G. Michel, T. Schweder, and U. T. Bornscheuer, **2022** *Appl. Microb. Biotechnol.* (submitted)

M.G., T.S. and U.T.B. initiated the study and directed the project. F.T. conducted the growth studies and created the knock-out strain. K.H.N performed the crystallization and structural analyses. T.D. performed the computational analysis. L. R. and H.C.G. performed the cloning and initial experiments on enzyme function of FoADH. S.B. and M.B. expressed and purified the enzymes and performed further experiments on enzyme function and characterization. S.B. and K.H.N prepared the main manuscript, which was revised by F.T., T.D, H.C.G., L. R., M.G, T.S. and U.T.B. and was approved by all authors.

Theresa Dutschei

Prof. Dr. Uwe. T. Bornscheuer

Article I

Accepted Article

Title: Marine Polysaccharides: Occurrence, Enzymatic Degradation and Utilization

Authors: Marcus Bäumgen, Theresa Deutschei, and Uwe Bornscheuer

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

To be cited as: *ChemBioChem* 10.1002/cbic.202100078

Link to VoR: <https://doi.org/10.1002/cbic.202100078>

Marine Polysaccharides: Occurrence, Enzymatic Degradation and Utilization

Marcus Bäumgen^{+[a]}, Theresa Dutschei^{+[a]} and Uwe T. Bornscheuer^{*[a]}

[a] Dr. M. Bäumgen, M.Sc., T. Dutschei, M.Sc., Prof. Dr. U. T. Bornscheuer
Department of Biotechnology & Enzyme Catalysis, Institute of Biochemistry
University of Greifswald
17487 Greifswald, Germany
E-mail: uwe.bornscheuer@uni-greifswald.de

[*] These authors contributed equally to this work

Abstract: Macroalgae species are fast growing and their polysaccharides are already used as food ingredient due to their properties as hydrocolloids or they have potential high value bioactivity. The degradation of these valuable polysaccharides to access the sugar components remained mostly unexplored so far. One reason is the high structural complexity of algal polysaccharides, but also the need for suitable enzyme cocktails to obtain oligo- and monosaccharides. Among them, there are several rare sugars with high value. Recently, considerable progress was made in the discovery of highly specific carbohydrate-active enzymes able to decompose complex marine carbohydrates such as carrageenan, laminarin, agar, porphyran and ulvan. This minireview summarizes these achievements and highlights potential applications of the now accessible abundant renewable resource of marine polysaccharides.

degradation pathway of alginate was illustrated^[12] which is complemented by a recent summary of the characteristics and applications of alginate lyases^[13] and new insights into fungal alginate lyases from *Paradendryphiella salina*.^[14] However, a detailed article dealing with the enzymatic degradation of other complex marine polysaccharides to access rare sugars is missing. This minireview therefore focuses on the current status of the microbial decomposition of the marine polysaccharide carrageenan, laminarin, agar, porphyran and ulvan (Scheme 1). We aim to provide an overview of the complexity of marine polysaccharides and the ubiquitous potential of this carbon source in biotechnological applications.

1. Introduction

The marine realm covers 70% of the earth's surface making the oceans the largest ecosystem on earth^[1], which may contain over 80% of world's plant and animal species.^[2] In particular, the marine systems have great influence on the atmospheric CO₂ concentration as the oceans contain the largest carbon pool in the carbon cycle.^[3] The increased level of atmospheric carbon dioxide, however, leads to a higher absorption rate by the world's oceans, resulting in a decreased pH-value.^[4] Consequently, the carbonate concentration in the surface water is reduced, making the ocean acidification a disturbing effect for the aquatic carbonate chemistry, which is of great importance for marine calcifying organisms like molluscs, crustaceans, echinoderms, corals, large calcareous algae, foraminifera and some phytoplankton.^[5] Besides the increasing CO₂ concentration on earth, the eutrophication of the oceans has a huge impact on the marine ecosystem. The increasing nutrient supply can cause an immense proliferation of algae, so called 'algae blooms' like the 'Golden tides', which are formed by the genus *Sargassum* in the Atlantic ocean or the 'Green tides', which are formed by the genus *Ulva* and occur worldwide.^[6,7] Beside the harmful environmental effects and high disposal cost of algal waste, the rising occurrence of algal biomass from these blooms also has a huge potential for biotechnological applications. One bottleneck for its use is access to the valuable chemical compounds within the algae, which has been described in recent reviews for the lipids and protein fractions.^[8–10] For marine polysaccharides, Trincone provided an overview about carbohydrate-active enzymes (CAZymes) involved in the degradation of macroalgal polysaccharides^[11] and Filote *et al.* covered aspects of potential biorefinery processes utilizing marine sugars.^[8] In the review by Ertesvåg, the enzymatic

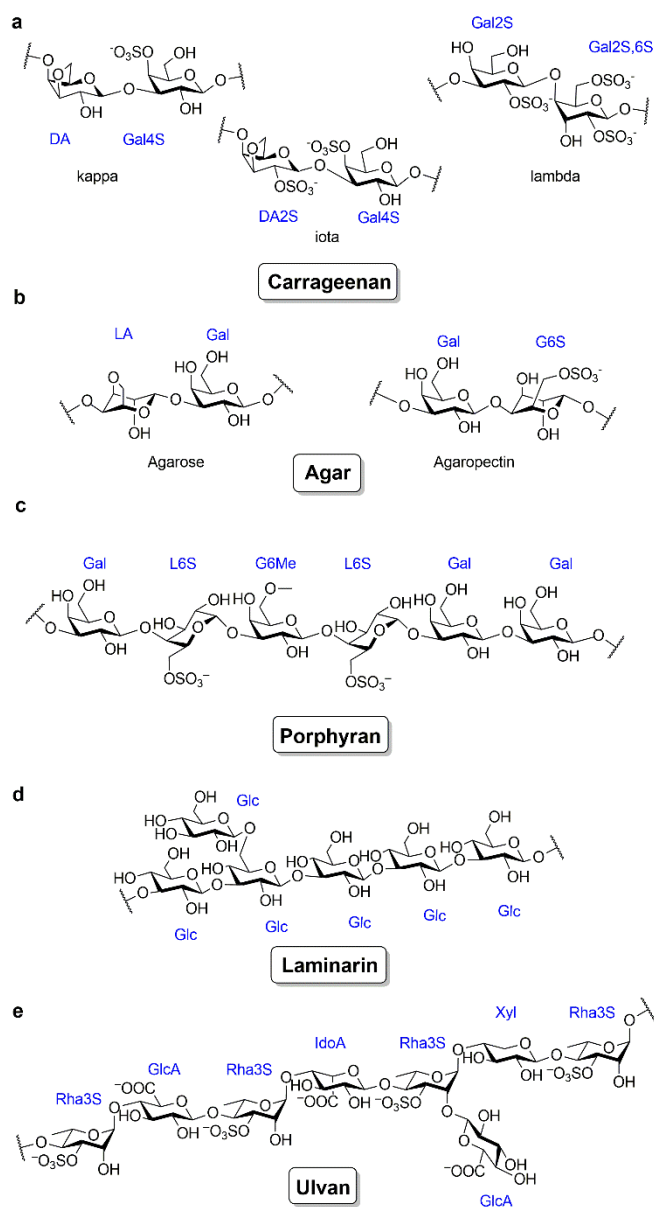
2. Diversity of marine carbohydrate structures

Carbohydrates represent the largest proportion of marine biomass. They mainly occur in marine plants, macro- and microalgae^[15–17] and can represent more than 50% of the algal dry weight.^[18–20] Many organisms use polysaccharides as intracellular energy storage compounds as well as structural cell wall components^[21] or secrete them as extracellular polymeric substances (EPS) with various functions.^[22] The polysaccharide composition varies substantially depending on the type of algae. Red algae mainly produce sulfated galactans, which are generally divided into agarans and carrageenans. While ulvan is the main polysaccharide in green algae, brown algae are known for the production of fucans.^[23] The polysaccharides of diatoms contain sulfated glucuronomannans and laminarin.^[24–27] The differences between terrestrial and marine carbohydrates originate in the variety of carbohydrate structures in their backbone as well as various modifications (Table 1). This is believed to be an adaptation to the marine environment.^[26,28,29] In comparison to freshwater and soil, the oceans contain a higher concentration of sulfate^[30,31] allowing for distinct sulfation patterns of the carbohydrates.^[26,28,29] Due to the anionic properties of marine polysaccharides, especially through sulfation, algae presumably are resistant to desiccation,^[32] osmotic stress and heavy metal toxicity^[33] as well as more extreme temperature and pH values.^[34] The side group modifications and decorations of the carbohydrates further increase the algae's recalcitrance to degradation by enzymatic attack. This drives the adaptation of marine organism, especially bacteria, to develop specific enzymes which can remove these modifications from the carbohydrate backbone and then use common CAZymes to hydrolyse the glycosidic sugar bonds. An overview of the diversity of selected carbohydrates from marine algal and their monosaccharide composition is given in Table 1.

Table 1. Overview of marine algae carbohydrates and organisms of marine origin containing characterized CAZymes. The different marine polysaccharides are listed with their monosaccharide composition, methylation- and sulfation-patterns. Furthermore, their main chain linkages and the occurrence of the corresponding polysaccharides in marine habitats are summarized. Marine organisms with characterized CAZymes for the degradation of the corresponding polysaccharide are also listed.

	Sugar composition ^[a]	–CH ₃ ^[b]	–OSO ₃ ^[b]	Marine occurrence	Major CAZyme ^[c]	Marine polysaccharide degrader ^[d]
Agar^[e]	β-1,4-D-Galactose α-1,3-3,6-Anhydro-L-galactose α-1,3-D-Galactose	+	+	Red algae	GH16, GH117, GH50, α-agarase EC. 3.2.1.158 β-agarase EC 3.2.1.81	<i>Zobellia galactanivorans</i> , ^[35] <i>Saccharophagus degradans</i> , ^[36] <i>Alterococcus agarolyticus</i> , ^[37] <i>Flammeovirga</i> sp. SJP92, ^[38]
Alginate	β-1,4-D-Mannuronic acid α-1,4-L-Guluronic acid	+	–	Brown algae	PL7 Mannuronate lyase EC 4.2.2.3 Guluronate lyase EC 4.2.2.11	<i>Sphingomonas</i> sp. MJ-3, ^[39] <i>Microbulbifer</i> sp. ALW1, ^[40] <i>Flavobacterium</i> sp. UMI-01 ^[41]
Carrageenan	β-1,4-D-Galactose α-1,3-3,6-Anhydro-D-galactose	+	+	Red algae	GH16 Carrageenase EC 3.2.1.83	<i>Pseudoalteromonas atlantica</i> , ^[42] <i>Zobellia galactinivorans</i> , ^[43] <i>Pseudoalteromonas carrageenovora</i> 9T, ^[44-46]
Cellulose	β-1,4-D-Glucose β-1,6-D-Glucose	–	–	Green and brown algae	GH48, GH17, GH16, GH9 Endoglucanase EC 3.2.1.6 Exoglucanase EC 3.2.1.74	<i>Glaciecola</i> sp. 4H-3-7+YE-5 ^[47] , <i>Actinoalloteichus</i> sp. MHA15 ^[48] , <i>Exiguobacterium</i> sp. Alg-S5 ^[49]
Fucoidan	α-1,3-L-Fucose, α-1,2-L-Fucose α-1,2-D-Glucuronic acid	–	+	Brown algae	GH29, GH107, GH168 α-L-Fucosidase EC 3.2.1.51 α-1,3-1,4-L-Fucosidase EC.3.2.1.111 endo-Fucoidanase EC 3.2.1.212	<i>Luteolibacter algae</i> H18, ^[50] <i>Wenyngzhuangia fucanilytica</i> , ^[51] <i>Lamellidens corrianus</i> , ^[52] <i>Vibrio</i> sp. EJY3 ^[53]
Laminarin	β-1,3-D-Glucose β-1,6-D-Glucose	–	–	Brown algae and diatoms	GH5 β-1,3-glucanase EC 3.2.1.6	<i>Formosa agariphila</i> GH17A, ^[54] <i>Formosa</i> sp. nov strain He1_33_131, ^[54] <i>Pseudocardium sachalinensis</i> , ^[54] <i>Vibrio campbellii</i> ^[55]
Mannan	β-1,4-D-Mannose α-1,4-D-Mannose	–	–	Red and Green algae	GH5 β-mannanase EC 3.2.1.78	<i>Streptomyces</i> sp. Alg-S25 ^[56]
Pectin	α-1,4-D-Galacturonic acid, α-1,6-D-Galactose, β-1,4-D-Xylose α-1,5-L-Arabinose α-1,2-D-Apiose α-1,2-L-Rhamnose	+	–	Green algae, diatoms	PL1, PL2, PL3 Pectin lyase EC 4.2.2.10	<i>Pseudoalteromonas</i> sp. PS47, ^[57] <i>Pseudoalteromonas haloplanktis</i> ANT/505 ^[58]
Porphyran	β-1,4-D-Galactose α-1,3-L-Galactose	+	+	Red algae	GH16, GH86 β-Porphyranase EC. 3.2.1.178	<i>Z. galactanivorans</i> , ^[59] <i>Bacteroides plebeius</i> ^[60]
Ulvan	β-1,4-D-Xylose α-1,2-L-Iduronic acid, β-1,2-D-Glucuronic acid, α-1,4-L-Rhamnose	+	+	Green algae	PL24, PL25, PL28 Ulvan lyase EC 4.2.2.-	<i>Formosa agariphila</i> ^[61,62]
Xylan	β-1,4-D-Xylose ^[f] β-1,3-D-Xylose ^[f]	+	+	Red and green algae	GH10, GH11, GH30 Endo-1,4-beta xylanase EC 3.2.1.8	<i>Paraglaciecola mesophile</i> KMM241, ^[63] <i>Vibrio</i> sp. XY-214, ^[64] <i>Alcaligenes</i> sp. XY-234, ^[65] <i>Glaciecola</i> sp. 4H-3-7+YE-5, ^[47] <i>Psychrobacter</i> sp. Strain 2-7, ^[66]

[a] The most prominent monosaccharides are listed. [b] Methylation (–CH₃) or sulfation (–OSO₃) patterns of the polysaccharides are indicated. The potential occurrence of these monosaccharide-decorations is marked with + or in their absence with -. [c] CAZyme families only represents the enzyme for initial depolymerisation of the polysaccharide [d] Characterized CAZymes from marine organism refer mostly to examples published between 2016–2020.^[11] [e] Agar is composed of agarose and agaropectin. [f] Red algae xylan consist of mixed linked type β-1,4-D-Xylose and β-1,3-D-Xylose while green algae xylan contains mostly β-1,3-D-Xylose.^[67]



Scheme 1. Structures of the marine polysaccharides carrageenan (a), agar (b), porphyran (c) laminarin (d) and ulvan (e). Carrageenan is composed of 3,6-anhydro-D-galactose (DA) and D-galactose (Gal). Agar divides in agar and agaropectin which contain Gal and 3,6-anhydro-L-galactose (LA). Laminarin contains D-glucose. Porphyran is composed of D-galactose, L-galactose (L) and D-glucuronic acid (GlcA). Ulvan is composed of D-glucuronic acid, L-iduronic acid (IdoA), D-xylose (Xyl) and L-rhamnose (Rha). A number in combination with an 'S' attached to a sugar represents the position of sulfate groups. A number in combination with a 'Me' attached to a sugar represents the position of methyl groups.

3. Enzymatic degradation of marine polysaccharides

The ability to compose and decompose polysaccharides is crucial for the global carbon cycle. To use them as an energy source, heterotrophic organisms require a suitable set of CAZymes in order to degrade them to monosaccharides, which can be further converted through the central sugar metabolism. Marine Bacteroidetes are specialized to use complex algal polysaccharides of different origins as a nutrient and therefore have developed surprisingly complex and dedicated enzyme toolboxes. This is also reflected by the observation that recurrent patterns of dominant bacterial groups outgrow during phytoplankton blooms in the North.^[68] Gene clusters encoding a set of enzymes and further proteins (i.e., for sugar transport) required to decompose algal polysaccharides are organized in

Bacteroidetes in so-called polysaccharide utilization loci (PULs). These encode a broad variety of CAZymes to decompose the complex polysaccharides.^[62,69] Without detailed knowledge on relevant enzyme functions, the guided degradation of marine polysaccharides *in vitro* is rather difficult.

The CAZy database (www.CAZy.org)^[70,71] lists CAZymes grouped by their enzyme class and genetic relationship. This presently includes 163 classes of glycoside hydrolases (GHs), 111 classes of glycosyl transferases (GTs), 40 classes of polysaccharide lyases (PLs), 18 classes of carbohydrate esterases (CEs) and 16 classes of enzymes with auxiliary activity (AAs). For the depolymerization of carbohydrates many different enzyme functions are necessary. There are *endo*-active CAZymes, which cleave within the polysaccharide chain and *exo*-enzymes, which remove saccharide fragments from the ends. Glycoside hydrolases are the most diverse family of CAZymes. They catalyse the hydrolysis of glycosidic bonds.^[72] In polysaccharides that contain uronic acid residues, like alginate or ulvan, polysaccharide lyases catalyse the non-hydrolytic cleavage of the chain at an uronic acid residue via a β elimination mechanism.^[73] Several side groups increase the resistance against backbone-cleaving enzymes. Besides further GHs that cleave off various monosaccharide side chains, other enzymes are required for the deprotection of the polysaccharide backbone. Polysaccharide sulfatases remove sulfate ester groups,^[26] while carbohydrate esterases catalyse the cleavage of O- and N-acetyl groups from carbohydrates.^[74] In contrast to the CEs, the sulfatases are not implemented in the CAZy database but are listed in the SulfAtlas database instead.^[75] The class of 'auxiliary activities' includes redox enzymes that act in conjunction with other CAZymes.^[76] This includes lytic polysaccharide monoxygenases (LPMOs) and enzymes known to be involved in lignin degradation. Another example for enzymes with auxiliary activities is the recently discovered cytochrome P450 monoxygenases from the marine bacteria *Formosa agariphila* and *Zobellia galactanivorans*. It was shown that they specifically catalyse the demethylation of 6-O-methyl-D-galactose present in the algal polysaccharides agarose and porphyran. Only after enzymatic hydroxylation and subsequent decomposition step – yielding the free hydroxyl group of D-galactose and formaldehyde – further degradation can occur.^[77] In the following, these complex pathways are highlighted for selected algal carbohydrates.

3.1 Carrageenan

Besides agars, carrageenans are the main cell wall polysaccharides of red macroalgae.^[78] Their structure is very complex and depend on the algal species. In general, they consist of sulfate esters of α -1,3-linked 3,6-anhydro-D-galactose and β -1,4-linked D-galactose (Table 1).^[11] The most prominent carrageenans for commercial applications are κ -, ι - and λ -carrageenan (Scheme 1).^[11,32] The decomposition of such complex polysaccharides to the monomeric level requires many different enzyme functions.

The main carrageenan degrading enzymes are κ -carrageenases (EC 3.2.1.83), ι -carrageenases (3.2.1.157) and λ -carrageenases (3.2.1.162). They cleave the β -1,4-linkages of polymeric carrageenans under the production of oligomeric neocarrabiose.^[79] The first ι -carrageenase-activity was reported for enzymes from *Alteromonas fortis* and *Z. galactanivorans* leading to the generations of GH family 82 which differs from κ -carrageenases.^[79] These enzyme groups and their mode of action were reviewed a few years ago.^[80] As carrageenans are highly sulfated polysaccharides, the removal of sulfate groups is required for achieving complete decomposition (Figure 1). The

first carrageenan sulfatase from *Pseudoalteromonas carrageenovora* was a 4-O- κ -carrabiose sulfatase.^[81] The synergistic degradation of sulfated carrageenans by GHs and sulfatases was demonstrated for *Z. galactanivorans*.^[78] Here, ι - and κ -carrageenan required a desulfation of the C4 sulfate group of D-galactose by two specialized sulfatases resulting in α - or β -carrageenan. A third sulfatase converts α -carrageenan into desulfated β -carrageenan by removing the C2 sulfate group from anhydro-galactose. Without these desulfations further degradation steps by GHs were blocked.^[78] In detail, the first step of ι -carrageenan degradation is the cleavage of the polysaccharide chain into smaller oligosaccharides by ι -carragenases of family GH82. In κ -carrageenan this first step is carried out by a κ -carrageenase of family GH16. Oligomeric ι -carrageenan requires a desulfation of the D-galactose residues by an ι -carrageenan G4S-sulfatase from family S1_19 resulting in oligomeric α -carrageenan. The same steps occur in κ -carrageenan. Here, a κ -carrageenan G4S-sulfatase hydrolyses

the sulfate ester at D-galactose residues, leading to unsulfated β -carrageenan. To convert α -carrageenan into the unsulfated β -carrageenan a desulfation of the remaining 3,6-anhydro-D-galactose residues by an α -carrageenan DA2S-sulfatase from family S1_7 is required. Finally, the unsulfated β -carrageenan can be successively degraded from the non-reducing end by 3,6-anhydro-D-galactosidases from family GH127 or GH129 proteins and β -galactosidases from family GH2 (Figure 1).^[78] Recently the carrageenan decomposition was investigated in several *Pseudoalteromonas* species.^[82] Two GHs from family GH16 with high identity to a previously described GH16 family κ -carrageenase from *P. carrageenovora* 9T^[44–46] and a β -carrageenan-specific *endo*-hydrolase from *Paraglaciecola hydrolytica* SS66T^[83] were able to degrade κ -carrageenan into even numbered κ -neocarrageenan oligosaccharides. The synergistic activity between previously described S1_19 sulfatase^[84] and these GH16 enzymes were revealed to resemble the previous results for *Z. galactanivorans*.^[78]

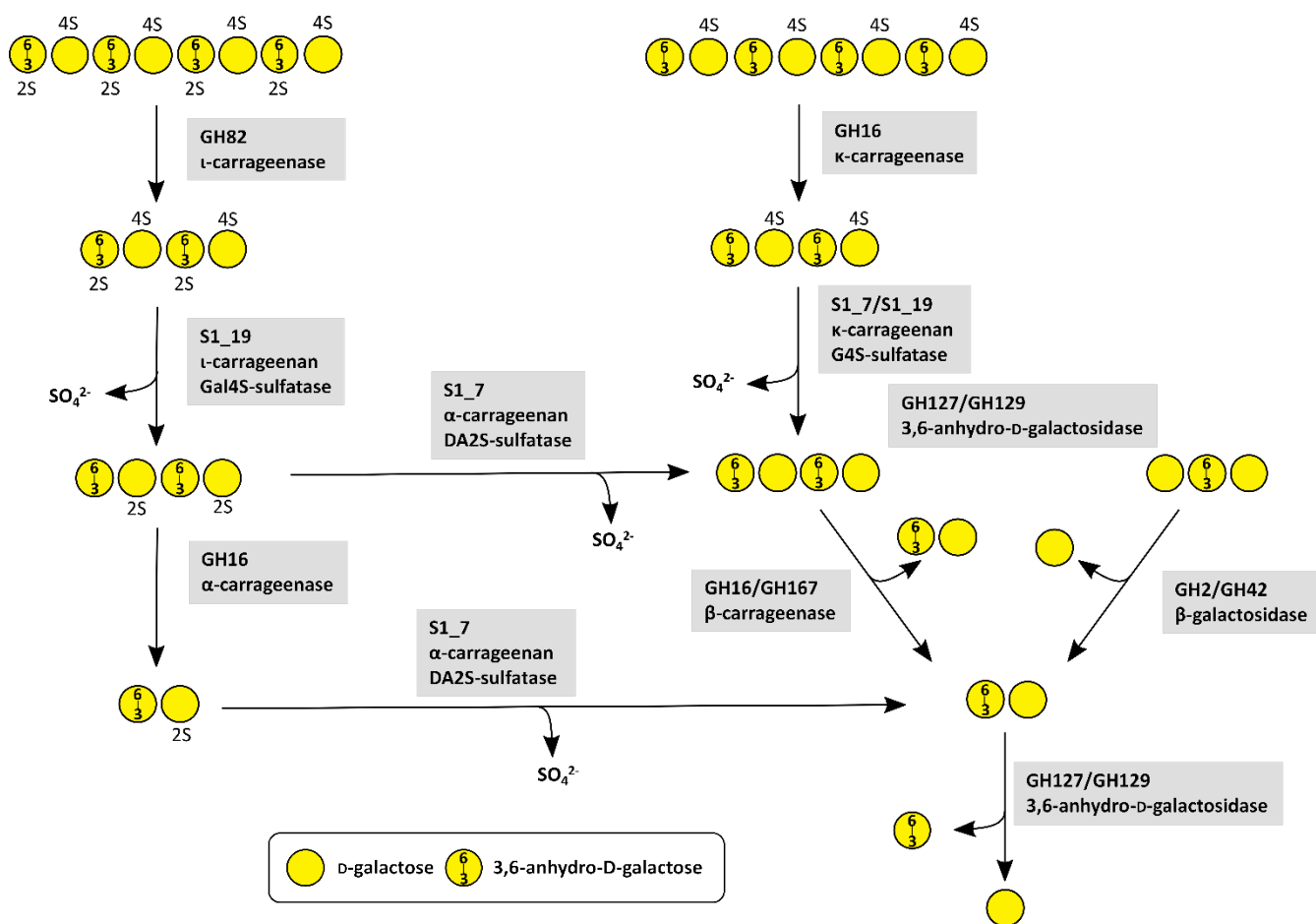


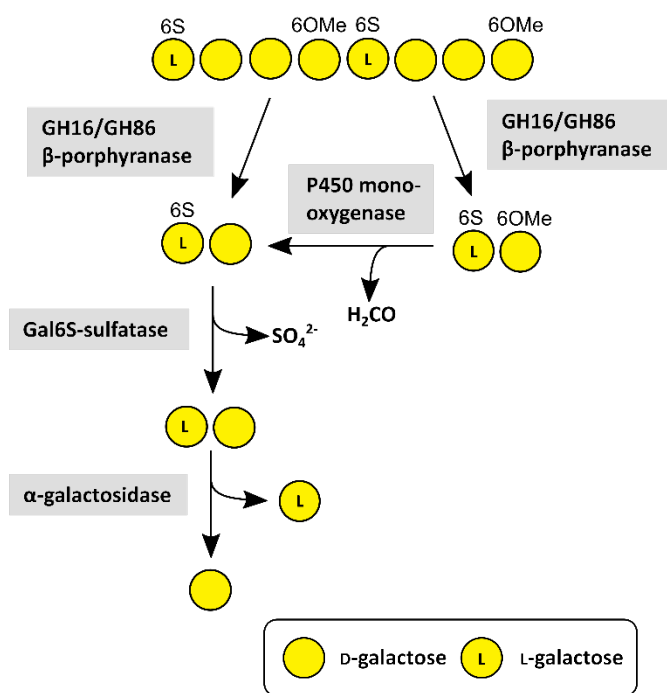
Figure 1. Metabolic carrageenan degradation pathways by CAZymes based on current knowledge.^[44–46,78] The oligosaccharides on the top represent a section of a larger polysaccharide chain. A number in combination with an 'S' attached to a sugar represents the position of sulfate groups.

A prior desulfation of κ -carrageenan or ι -carrageenan by the S1_19 sulfatase allowed depolymerisation by a third GH16 enzyme indicating an α - or β -carrageenases activity. A second S1_19 sulfatase was revealed to be an *exo*-G4S κ -carrageenan sulfatase being inactive on ι -carrageenan.^[82] A β -neocarrabiose releasing *exo*-carrageenase of family GH167 was shown to degrade short κ -carrageenan oligosaccharides after treatment with κ -carrageenan-active sulfatase. This enzyme showed 63% identity^[82] with a formerly described carrageenan-active enzyme from *P. hydrolytica*.^[83] The synergistic degradation of sulfated

carrageenans by GHs and sulfatases was demonstrated for *Z. galactanivorans*.^[78] Here, ι - and κ -carrageenan required a desulfation of the C4 sulfate group of D-galactose by two specialized sulfatases resulting in α - or β -carrageenan. A third sulfatase converts α -carrageenan into desulfated β -carrageenan by removing the C2 sulfate group from anhydro-galactose. Without these desulfations further degradation steps by GHs were blocked.^[78] In detail, the first step of ι -carrageenan degradation is the cleavage of the polysaccharide chain into smaller oligosaccharides by ι -carragenases of family GH82. In κ -

carrageenan this first step is carried out by a κ -carrageenase of family GH16. Oligomeric ι -carrageenan requires a desulfation of the D-galactose residues by an ι -carrageenan G4S-sulfatase from family S1_19 resulting in oligomeric α -carrageenan. The same steps occur in κ -carrageenan. Here, a κ -carrageenan G4S-sulfatase hydrolyses the sulfate ester at D-galactose residues, leading to unsulfated β -carrageenan. To convert α -carrageenan into the unsulfated β -carrageenan a desulfation of the remaining 3,6-anhydro-D-galactose residues by an α -carrageenan DA2S-sulfatase from family S1_17 is required. Finally, the unsulfated β -carrageenan can be successively degraded from the non-reducing end by 3,6-anhydro-D-galactosidases from family GH127 or GH129 proteins and β -galactosidases from family GH2 (Figure 1).^[78] Recently the carrageenan decomposition was investigated in several *Pseudoalteromonas* species.^[82] Two GHs from family GH16 with high identity to a previously described GH16 family κ -carrageenase from *P. carrageenovora* 9T^[44–46] and a β -carrageenan-specific *endo*-hydrolase from *Paraglaciicola hydrolytica* SS66T^[83] were able to degrade κ -carrageenan into even numbered κ -neocarrageenan oligosaccharides. The synergistic activity between previously described S1_19 sulfatase^[84] and these GH16 enzymes were revealed to resemble the previous results for *Z.*

a



b

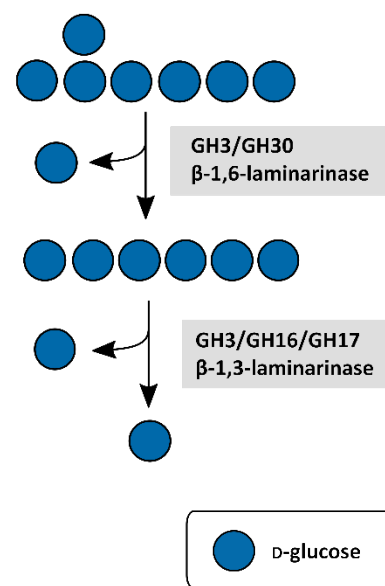


Figure 2. Metabolic porphyran (a) and laminarin (b) degradation pathways by CAZymes based on current knowledge.^[54,55] The oligosaccharides on the top represent a section of a larger polysaccharide chain. A number in combination with an 'S' attached to a sugar represents the position of sulfate groups. A number in combination with an 'OMe' attached to a sugar represents the position of methyl ether groups.

galactanivorans.^[78] A prior desulfation of κ -carrageenan or ι -carrageenan by the S1_19 sulfatase allowed depolymerisation by a third GH16 enzyme indicating an α - or β -carrageenase activity. A second S1_19 sulfatase was revealed to be an *exo*-G4S κ -carrageenan sulfatase being inactive on ι -carrageenan.^[82] A β -neocarrabiase releasing *exo*-carrageenase of family GH167 was shown to degrade short κ -carrageenan oligosaccharides after treatment with κ -carrageenan-active sulfatase. This enzyme showed 63% identity^[82] with a formerly described carrageenan-active enzyme from *P. hydrolytica*.^[83]

3.2 Porphyrin

Agars are galactans from red algae containing α -1,3-linked L-galactose and β -1,4-linked D-galactose. L-galactose is replaced by 3,6-anhydro-L-galactose in agarose and by L-galactose-6-sulfate in porphyran (Scheme 1, Table 1).^[32,85] Porphyran is especially abundant in algae of the genus *Porphyra*. The degradation of agars in general was reviewed before^[86] followed by biochemical characterizations of agarose-degrading pathways.^[87,88]

The marine porphyran degradation was enabled with the investigation of the first marine β -porphyranases from the Bacteroidetes *Z. galactanivorans*.^[59] These enzymes belong to family GH16 and were shown to cleave the β -1,4-linkage between β -D-galactose and α -L-galactose-6-sulfate in purified polymeric porphyran resulting in the disaccharide Gal6S-Gal as the final degradation product.^[59] Later a new β -porphyranase from *Bacteroides plebeius* from family GH86 was identified.^[60] Similar studies indicated the importance of GH16 enzymes in the degradation of porphyran by the investigation of further GH16 porphyranases.^[85,89] Nevertheless, these enzymes are not sufficient on their own for the complete degradation of porphyran (Figure 2). They require a synergistic cleavage of several side group-removing enzymes which deprotect the polysaccharide chain from functional groups and thereby enable further degradation by the porphyranases. As mentioned above, P450 monooxygenases catalyse demethylation of 6-O-methyl-D-galactose – a monosaccharide that replaces D-galactose in porphyran in a random manner.^{[90,91][77]} Hence, these P450s are crucial for the complete decomposition of porphyran. Besides methylated sugars, porphyran is known to contain L-galactose-6-sulfate.^[92] There are reports about putative sulfatase genes in PUL structures presumably targeting porphyran,^[60] but they have not been biochemically characterized and their function is not yet confirmed.

3.3 Laminarin

Laminarin is one of the most abundant marine polysaccharides.^[27] It occurs in brown algae and especially in diatoms (Table 1).^[93] It is a highly water soluble glucan which is composed of linear β -1,3-linked D-glucose with β -1,6-linked D-glucose side chains (Scheme 1).^[93]

Several CAZymes are required for the depolymerisation of laminarin (Figure 2). Laminarinases, the main laminarin-degrading enzymes, are classified into *endo*- β -1,3-glucanases (laminarinases) (EC 3.2.1.6 and EC 3.2.1.39) and *exo*- β -1,3-glucanases (EC 3.2.1.58). *Endo*- β -1,3-glucanases hydrolyse the β -1,3 backbone while *exo*- β -1,3-glucanases cleave off glucose from the nonreducing ends of laminarin oligosaccharides. *Endo*-acting laminarinases are mainly grouped into families GH16, GH17, GH55, GH64, and GH81, while the GH3 family contains *exo*-acting laminarinases.^[54]

While GH16 laminarinases can cleave β -1,3- and β -1,4-linkages, GH17 enzymes are highly specific for undecorated β -1,3 glucans.^[69] Two GH16 and GH17 enzymes from *F. agariphila* KMM 3901^T and a GH30 enzyme from *Formosa* sp. Hel1_33_131 were investigated as well.^[54] The GH16 enzyme had 44% and 43% identity with two *endo*-acting laminarinases from *Z. galactanivorans*. GH17 enzymes are *endo*-type enzymes specific for β -1,3-glucans, while the GH30 family contains enzymes, which are specific for β -1,6-glucans. Thereby, *endo*-acting enzymes from family GH17 are required for the depolymerisation of the laminarin backbone, while *exo*-acting GH30 enzymes hydrolyse the side chains. This combination of GH17 and GH30 enzymes is necessary for an efficient laminarin depolymerization.^[54] Beside this, there has been a report about a promiscuous GH3-like laminarinase from *Vibrio campbellii*, which is able to cleave β -1,3-linkages as well as β -1,4- and β -1,6-linkages (Figure 2).^[55] These laminarin-degrading enzymes are conserved in marine Bacteroidetes. Thus, it was demonstrated that enzymes encoded in both chromosomes of *P. carrageenovora* showed activity on β -1,3-glucans. They contain genes for several GH16 *endo*-1,3- β -glucanases. One of the respective PUL structures is conserved in

47 of 52 analysed *Pseudoalteromonas* genomes.^[94] Genome analyses of 53 marine bacterial isolates revealed 400 PULs from which 46 PULs (ca. 12%) are putatively laminarin-targeting.^[69] Thus, the laminarin decomposition plays an important role in the marine polysaccharide turnover.

3.4 Ulvan

Ulvan is the major cell wall polysaccharide of macroalgae from the genus *Ulva*.^[62] It is a branched, highly sulfated polysaccharide composed of repeating disaccharide units of β -1,4-linked D-glucuronic acid (GlcA) or α -L-iduronic acid (IdoA) to α -1,4-linked L-rhamnose-3-sulfate (Rha3S). The GlcA can be replaced by β -1,4-linked D-xylose (Xyl) or D-xylose-2-sulfate (Xyl2S) (Scheme 1, Table 1). Also, GlcA side chains at position 2 of Rha3s have been reported.^[62]

The first enzymatic decomposition of ulvan by a marine bacterium was reported more than twenty years ago when the first ulvan lyase (EC 4.2.2.-) was discovered.^[95] Several other ulvan lyases from the families PL24, PL25, PL28 and PL40 were described in various Bacteroidetes and Proteobacteria.^[61,62,87,96-105] They catalyse the initial cleavage step for the degradation of ulvan via an elimination mechanism and cleave the α -1,4-linkage between rhamnose-3-sulfate and glucuronic or iduronic acid under the formation of an unsaturated uronic acid residue (Δ) at the non-reducing end. This residue can then be hydrolytically removed by glucuronyl hydrolases of the family GH88 or GH105 (EC 3.2.1.-),^[62,106,106-109] forming 5-dehydro-4-deoxy-D-glucuronate. Enzyme functions for the ulvan degradation system of *F. agariphila* KMM 3901^T were first predicted by similarity with the help of artificial chromogenic substrates^[107] after which the first complete metabolic ulvan degradation pathway was elucidated (Figure 3).^[62]

In *F. agariphila* the initial depolymerization step is catalysed by ulvan lyases of family PL28 and PL40. The PL28 family ulvan lyase exhibits a type IX secretion signal and an additional ulvan binding module, which facilitates the recognition and binding of polymeric ulvan. As these properties are missing in the PL40 family lyase, it is suggested that it is more likely a membrane-associated or periplasmic enzyme. This indicates that its main function is to degrade larger oligosaccharides produced by PL28 family lyase, although it exhibits the same activity against polymeric ulvan.^[62] This strategy most probably avoids smaller substrate molecules diffusing away from the bacterial cell. Instead, they are cleaved immediately before or after their uptake into the periplasm by TonB-dependent transporters (TBDTs). This strategy is also known as the 'selfish' uptake mechanism.^[110] Larger oligosaccharides that are resistant to ulvan lyases, usually contain larger amounts of xylose, as it was shown for xylose-rich ulvan. A novel *endo*-rhamnosidase of family GH39 was demonstrated to degrade them. Up to this point, family GH39 was not described to contain rhamnosidases. A BlastP search revealed that this enzyme shows a rather low identity with all other GH39 enzymes.^[107] Thus, it was described to be a new type of a GH39 enzyme with a novel activity and most presumably different structural motifs due to the confirmed *endo*activity.

Beside these xylose-containing oligosaccharides, uronic acid-containing oligosaccharides were described to be resistant to further ulvan lyase degradation as well.^[62] At higher ulvan concentration the lyases are inhibited by their own products.^[95] Furthermore, the small xylose-containing oligosaccharides, are resistant to further degradation by lyases or glycoside hydrolases. For a complete depolymerization, a removal of any side chains and protective groups from the particular polysaccharide is necessary. The cleavage of sulfate ester bonds requires a set of

specialized sulfatases. On ulvan fragments, sulfatases from the families S1_7, S1_8 and S1_25 showed activity.^[62,111] An endolytic S1_8 family xylose sulfatase was described to desulfate small xylose-containing tri- and tetrasaccharides like Δ -Rha3S-Xyl2S-Rha3S and Rha3S-Xyl-Rha3S. In contrast, exolytic rhamnose- and xylose- sulfatases from family S1_25 and S1_7 are responsible for desulfation of non-reducing end rhamnose or

xylose residues. This enables a further degradation by several other CAZymes. *Exo*-rhamnosidases from family GH78 cleave off the non-reducing end rhamnose residue. In the case of uronic acid-containing fragments, the responsible GH78 rhamnosidase was shown to be a multimodular CAZyme also containing a family S1_25 sulfatase responsible for desulfating the substrate non-reducing end rhamnose residue.^[111]

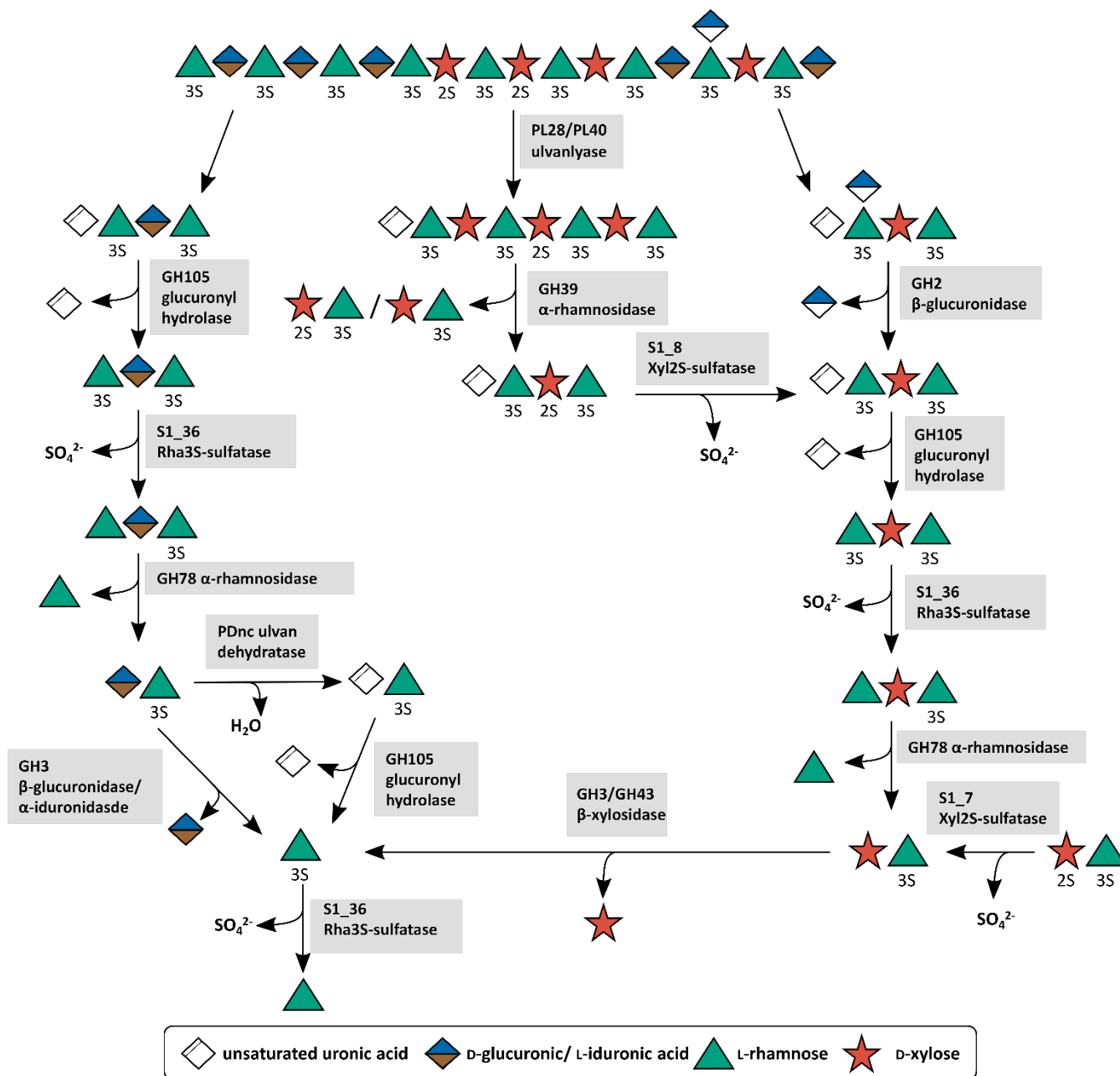


Figure 3. Metabolic ulvan degradation pathway by CAZymes based on current knowledge.^[62,111] The oligosaccharide on the top represents a section of a larger polysaccharide chain. A number in combination with an 'S' attached to a sugar represents the position of sulfate groups. 'Unsaturated uronic acid' represents 4-deoxy- α -L-threo-hex-4-enopyranuronic acid.

This was the first characterized multimodular CAZyme involved in ulvan degradation. Multimodular enzymes were revealed before in other Bacteroidetes even the same combination of GH78 and sulfatase was reported for *N. ulvanivorans*,^[26] which underlines the importance of this combination of enzyme activities for the utilization of ulvan. The last step of the degradation of ulvan to monomeric sugars is cleavage of the disaccharides Xyl-Rha3S and GlcA/IdoA-Rha3S.

β -xylosidases of family GH3 and GH43 were shown to hydrolyse Xyl-Rha3S. Two options for the cleavage of GlcA/IdoA-Rha3S were recently discovered by us.^[111] A β -glucuronidase/ α -iduronase from family GH3 or a novel polysaccharide dehydratase in combination with unsaturated glucuronyl hydrolase from family 105 were able to cleave the uronic acid-containing disaccharides (unpublished).

4. Saccharification processes for marine sugars

The elucidation of these marine polysaccharide utilization systems enables the use of algal biomass for fermentation processes as well as the production of biofuels and high value fine chemicals. Currently, algal biomass is treated as waste and accumulates in very large amounts due to the high growth rate of algae,^[112] making it a cheap and easily accessible source of raw materials. For their use in biotechnological applications, biorefinery concepts were published mostly for the efficient saccharification and fermentation of brown algae carbohydrates. Among these, the metabolic engineering of *Saccharomyces cerevisiae* for the fermentation of mannitol and alginate degradation products to ethanol was reported.^[113] The metabolic engineering of *Escherichia coli* led to the creation of a strain that is able to degrade, take up and metabolize alginate under the production of bioethanol.^[114] The clarification of the metabolic pathway for 3,6-anhydro-L-galactose enabled the use of the red algae polysaccharides agar and carrageenan for the same purpose.^[115] The direct bioconversion of brown algae into ethanol was reported for *Defluviitalea phaphyphila*.^[116] Another milestone in the biofuel production from algae was the engineering of the yeast *S. cerevisiae* for enzymatic hydrolysis of laminarin from brown macroalgae for the production of bioethanol.^[117] Recent reviews summarize biofuel feedstocks including macroalgal^[118] and related genetic engineering.^[119]

The production of meso-2,3-butanediol from glucose was demonstrated by metabolic engineering of *Bacillus licheniformis*.^[120] Glucose can be produced using the widespread glucanases of various organisms to degrade the green algal glucans and the brown algal laminarins. This enables the 2,3-butanediol fermentation using marine polysaccharides as feedstock. Besides ethanol, hydrogen is a promising energy carrier. The fermentative hydrogen evolution was reviewed showing biochemical pathways for the production of hydrogen by various microorganisms.^[121] The reported hydrogen generation systems involving bacteria growing on first or second generation plant sources^[122] can easily be adapted to the use of algal biomass as the investigation of fermentative pathways starts with monosaccharides that can be provided by either land plants or algae. Thus, the hydrogen evolution using the hyperthermophilic bacteria *Thermotoga neapolitana* on biomass of the green alga *Chlamydomonas reinhardtii* was reported.^[123] The production of hydrogen from xylose is also possible using an *in vitro* enzyme cascade,^[124] showing that hydrogen evolution can also work cell-free. To increase the yield of produced biofuels from macroalgae several pretreating methods were described recently.^[125,126]

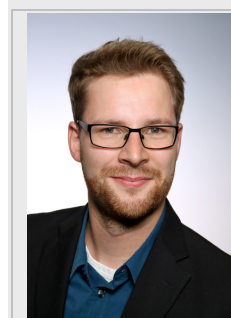
In principle, the application of carbohydrates in biotechnological processes for fermentation requires the possibility to fully degrade the respective carbohydrate to the monomeric level and the ability to metabolize the corresponding monosaccharides released by the polysaccharide degradation. Thus, the more complex the polysaccharide is, the more CAZymes are required for its degradation. If the polymer contains rare sugars, there are often only a few microorganisms which are able to metabolize them. For the production of ethanol, yeasts like *S. cerevisiae* are often used because they exhibit a high ethanol tolerance.^[127] Even if yeasts have been implemented in the production from bioethanol from brown algae^[117] they still often lack genes in the metabolic pathways encoding for proteins for the conversion of pentose sugars like xylose and arabinose.^[128] Enabling a usage of these sugars would require metabolic

engineering,^[129] complicating the use of algal biomass in yeast fermentation processes. However, especially rare sugars can not only serve as carbon source for fermentation, but also be a value product on their own. Thus, 3,6-anhydro-L-galactose can be isolated from red algae and it was reported to exhibit skin whitening and anti-inflammatory properties.^[130]

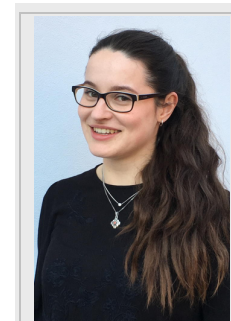
5. Conclusion

In contrast to their terrestrial counterpart, the metabolic degradation of marine polysaccharides is currently still underexplored. Marine microorganisms provide enzymatic toolboxes for the successive degradation of these carbohydrates into monomeric sugars. The acquired insight of the metabolic polysaccharide utilization greatly expands the possibility to use algal waste for recycling in biorefinery processes to high value materials with even beneficial effect for the environment. The research on this topic is still in its infancy regarding the huge diversity of marine polysaccharides and still much scientific work is necessary in this field to overcome the bottlenecks for producing fermentable saccharide fragments from algae for the production of valuable chemicals.

Marcus Bäumen studied biochemistry at Greifswald University where he received his PhD in 2020 in the group of Prof. Bornscheuer. After focusing on research with Baeyer-Villiger monooxygenases, he switched his interest to the marine polysaccharide utilization and functional characterization of marine carbohydrate-active enzymes within the DFG-funded research unit FOR2406 "Proteogenomic of Marine Polysaccharide Utilization" (POMPU).



Theresa Dutschei studied biochemistry at Greifswald University and received her M.Sc degree in the group of Prof. Bornscheuer. In the field of applied marine carbohydrate saccharification and utilization. Since 2019 she also works in the DFG-funded POMPU-project on the functional characterization of marine carbohydrate-active enzymes.



Uwe T. Bornscheuer studied chemistry and received his PhD in 1993 at Hannover University followed by a postdoc at Nagoya University (Japan). In 1998, he completed his Habilitation at Stuttgart University about the use of lipases and esterases in organic synthesis. He has been Professor at the Institute of Biochemistry at Greifswald University since 1999. Beside other awards, he received in 2008 the BioCat2008 Award. He was just recognized as 'Chemistry Europe Fellow'. His current research interest focuses on the discovery and engineering of enzymes from various classes for applications in organic synthesis, lipid modification, degradation of plastics or complex marine polysaccharides.



Acknowledgements

We thank the German Research Foundation (DFG) for funding through the Research Unit FOR2406 "Proteogenomics of Marine Polysaccharide Utilization" (POMPU) (BO 1862/17-1 and BO 1862/17-2) and the BMBF for funding in the frame of the Plant3 project MarZucker (03WIR2205A). We also thank Thomas Schweder, Jan-Hendrik Hehemann and Lukas Reisky for valuable discussions.

Conflict of interest

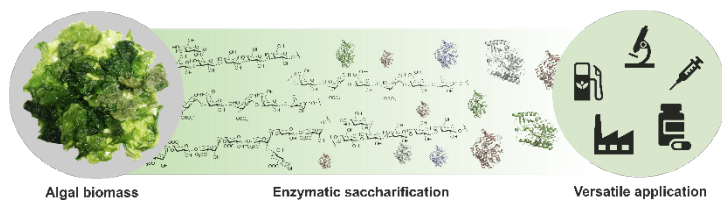
The authors declare no conflict of interest.

Keywords: marine polysaccharides • ulvan • laminarin • porphyrin • carrageenan • CAZymes • enzymatic saccharification

- [1] S. Das, P. S. Lyla, S. A. Khan, *Curr. Sci.* **2006**, *90*, 1325–1335.
- [2] P. J. McCarthy, S. A. Pomponi, **2004**, *22*, 1–2.
- [3] W. M. Post, T.-H. Peng, W. R. Emanuel, A. W. King, V. H. Dale, D. L. DeAngelis, *Am. Sci.* **1990**, *78*, 310–326.
- [4] K. Caldeira, M. E. Wickett, *Nature* **2003**, *425*, 365–365.
- [5] Royal Society (Great Britain), *Ocean Acidification Due to Increasing Atmospheric Carbon Dioxide*, Royal Society, London, **2005**.
- [6] M. Wang, C. Hu, B. B. Barnes, G. Mitchum, B. Lapointe, J. P. Montoya, *Science* **2019**, *365*, 83–87.
- [7] V. Smetacek, A. Zingone, *Nature* **2013**, *504*, 84–88.
- [8] C. Filote, S. C. R. Santos, V. I. Popa, C. M. S. Botelho, I. Volf, *Environ. Chem. Lett.* **2020**, *1*, 1–36.
- [9] M. Fabris, R. M. Abbriano, M. Pernice, D. L. Sutherland, A. S. Commault, C. C. Hall, L. Labeeuw, J. I. McCauley, U. Kuzhiuparambil, P. Ray, T. Kahlke, P. J. Ralph, *Front. Plant Sci.* **2020**, *11*, 279.
- [10] L. M. L. Laurens, M. Lane, R. S. Nelson, *Trends Biotechnol.* **2020**, *11*, 1232–1244.
- [11] A. Trincone, *Molecules* **2018**, *23*, 901.
- [12] H. Ertesvåg, *Front. Microbiol.* **2015**, *6*, 523.
- [13] D. Cheng, C. Jiang, J. Xu, Z. Liu, X. Mao, *Int. J. Biol. Macromol.* **2020**, *164*, 1304–1320.
- [14] B. Piilgaard, M. Vuillemin, J. Holck, C. Wilkens, A. S. Meyer, *J. Fungi* **2021**, *7*, 80.
- [15] B. Subhadra, M. Edwards, *Energy Policy* **2010**, *38*, 4897–4902.
- [16] R. Taylor, R. L. Fletcher, J. A. Raven, *Bot. Mar.* **2001**, *44*, 327–336.
- [17] R. A. Cohen, P. Fong, *Ecol. Appl.* **2006**, *16*, 1405–1420.
- [18] S. Kraan, *Carbohydr. - Compr. Stud. Glycobiol. Glycotechnol.* **2012**, *1*, 1–44.
- [19] M. Murata, J. Nakazoe, *Jpn. Agric. Res. Q. JARQ* **2001**, *35*, 281–290.
- [20] M. Lahaye, J.-L. Gomez-Pinchetti, M. J. del Rio, G. Garcia-Reina, *J. Sci. Food Agric.* **1995**, *68*, 99–104.
- [21] B. Kloareg, R. S. Quatrano, *Oceanogr. Mar. Biol. Annu. Rev.* **1988**, *26*, 259–315.
- [22] K. D. Hoagland, J. R. Rosowski, M. R. Gretz, S. C. Roemer, *J. Phycol.* **1993**, *29*, 537–566.
- [23] G. Jiao, G. Yu, J. Zhang, H. S. Ewart, *Mar. Drugs* **2011**, *9*, 196–223.
- [24] E. Percival, R. H. McDowell, **1967**, *162*, 895–896.
- [25] A. Willis, A. Chiovitti, T. M. Dugdale, R. Wetherbee, *J. Phycol.* **2013**, *49*, 937–949.
- [26] W. Helbert, *Front. Mar. Sci.* **2017**, *4*, 1–10.
- [27] S. Becker, J. Tebben, S. Coffinet, K. Wiltshire, M. H. Iversen, T. Harder, K.-U. Hinrichs, J.-H. Hehemann, *Proc. Natl. Acad. Sci.* **2020**, *117*, 6599–6607.
- [28] R. S. Aquino, A. M. Landeira-Fernandez, A. P. Valente, L. R. Andrade, P. A. S. Mourão, *Glycobiology* **2005**, *15*, 11–20.
- [29] J. L. Olsen, P. Rouzé, B. Verhelst, Y.-C. Lin, T. Bayer, J. Collen, E. Dattolo, E. De Paoli, S. Dittami, F. Maumus, G. Michel, A. Kersting, C. Lauritano, R. Lohaus, M. Töpel, T. Tonon, K. Vanneste, M. Amirebrahimi, J. Brakel, C. Boström, M. Chovatia, J. Grimwood, J. W. Jenkins, A. Jueterbock, A. Mraz, W. T. Stam, H. Tice, E. Bornberg-Bauer, P. J. Green, G. A. Pearson, G. Procaccini, C. M. Duarte, J. Schmutz, T. B. H. Reusch, Y. Van de Peer, *Nature* **2016**, *530*, 331–335.
- [30] J. Wright, A. Colling, *Seawater Its Compos. Prop. Behav.* **1995**.
- [31] M. Bochenek, G. J. Etherington, A. Koprivova, S. T. Mugford, T. G. Bell, G. Malin, S. Kopriva, *New Phytol.* **2013**, *199*, 650–662.
- [32] E. Ficko-Blean, C. Hervé, G. Michel, *Perspect. Phycol.* **2015**, *2*, 51–64.
- [33] L. R. Andrade, R. N. Leal, M. Nosedá, M. E. R. Duarte, M. S. Pereira, P. A. S. Mourão, M. Farina, G. M. Amado Filho, *Mar. Pollut. Bull.* **2010**, *60*, 1482–1488.
- [34] M. Ucko, E. Cohen, H. Gordin, *APPL. ENV. MICROBIOL.* **1989**, *55*, 5.
- [35] M. Jam, D. Flament, J. Allouch, P. Potin, L. Thion, B. Kloareg, M. Czjzek, W. Helbert, G. Michel, T. Barbeyron, *Biochem. J.* **2005**, *385*, 703–713.
- [36] H. T. Kim, S. Lee, K. H. Kim, I.-G. Choi, *Bioresour. Technol.* **2012**, *107*, 301–306.
- [37] W. Y. Shieh, W. D. Jean, *Can. J. Microbiol.* **1998**, *44*, 637–645.
- [38] Y. Chu, Z. Yi, R. Zeng, G. Zhang, *J. Mol. Catal. B Enzym.* **2016**, *129*, 47–53.
- [39] M. Ryu, E. Y. Lee, *J. Ind. Eng. Chem.* **2011**, *17*, 853–858.
- [40] Y. Zhu, L. Wu, Y. Chen, H. Ni, A. Xiao, H. Cai, *Microbiol. Res.* **2016**, *182*, 49–58.
- [41] A. Inoue, R. Nishiyama, T. Ojima, *Algal Res.* **2016**, *19*, 355–362.
- [42] M. Akagawa-Matsushita, M. Matsuo, Y. Koga, K. Yamasato, *Int. J. Syst. Bacteriol.* **1992**, *42*, 621–627.
- [43] E. Ficko-Blean, A. Préchoux, F. Thomas, T. Rochat, R. Larocque, Y. Zhu, M. Stam, S. Génicot, M. Jam, A. Calteau, B. Viart, D. Ropartz, D. Pérez-Pascual, G. Correc, M. Matarad-Mann, K. A. Stubbs, H. Rogniaux, A. Jeudy, T. Barbeyron, C. Médigue, M. Czjzek, D. Vallenet, M. J. McBride, E. Duchaud, G. Michel, *Nat. Commun.* **2017**, *8*, 1685.
- [44] G. Michel, L. Chantalat, E. Duee, T. Barbeyron, B. Henrissat, B. Kloareg, O. Dideberg, *Structure* **2001**, *9*, 513–525.
- [45] T. Barbeyron, B. Henrissat, B. Kloareg, *Gene* **1994**, *139*, 105–109.
- [46] M. Matarad-Mann, T. Bernard, C. Leroux, T. Barbeyron, R. Larocque, A. Préchoux, A. Jeudy, M. Jam, P. Nyvall Collén, G. Michel, M. Czjzek, *J. Biol. Chem.* **2017**, *292*, 19919–19934.
- [47] B. Klippel, A. Lochner, D. C. Bruce, K. W. Davenport, C. Detter, L. A. Goodwin, J. Han, S. Han, M. L. Land, N. Mikhailova, M. Nolan, L. Pennacchio, S. Pitluck, R. Tapia, T. Woyke, S. Wiebusch, A. Basner, F. Abe, K. Horikoshi, M. Keller, G. Antranikian, *J. Bacteriol.* **2011**, *193*, 4547–4548.
- [48] G. Rajagopal, S. Kannan, *Biotechnol. Rep. Amst. Neth.* **2017**, *13*, 30–36.
- [49] B. R. Mohapatra, *Int. J. Biol. Macromol.* **2017**, *98*, 103–110.
- [50] T. Nagao, A. Kumabe, F. Komatsu, H. Yagi, H. Suzuki, T. Ohshiro, *J. Biosci. Bioeng.* **2017**, *124*, 277–282.
- [51] S. Dong, Y. Chang, J. Shen, C. Xue, F. Chen, *Protein Expr. Purif.* **2017**, *129*, 9–17.
- [52] A. Venugopal, C. Sudheer Kumar, N. Siva Kumar, M. J. Swamy, *Int. J. Biol. Macromol.* **2017**, *104*, 432–441.
- [53] H. Hong, D. H. Kim, H. Seo, K. H. Kim, K.-J. Kim, *J. Agric. Food Chem.* **2021**, *69*, 3380–3389.
- [54] S. Becker, A. Scheffel, M. F. Polz, J.-H. Hehemann, *Appl. Environ. Microbiol.* **2017**, *83*, 1–14.
- [55] L. Wang, Y. Chen, H. Huang, Z. Huang, H. Chen, Z. Shao, *Aquac. Res.* **2015**, *46*, 395–404.
- [56] B. R. Mohapatra, *Biotechnol. Biotechnol. Equip.* **2021**, *35*, 150–161.
- [57] J. K. Hobbs, A. G. Hettle, C. Vickers, A. B. Boraston, *Appl. Environ. Microbiol.* **2018**, *85*, 1–16.
- [58] J.-H. Hehemann, L. V. Truong, F. Unfried, N. Welsch, J. Kabisch, S. E. Heiden, S. Junker, D. Becher, A. Thürmer, R. Daniel, R. Amann, T. Schweder, *Environ. Microbiol.* **2017**, *19*, 2320–2333.
- [59] J.-H. Hehemann, G. Michel, T. Barbeyron, M. Czjzek, *Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun.* **2010**, *66*, 413–417.
- [60] J.-H. Hehemann, A. G. Kelly, N. A. Pudlo, E. C. Martens, A. B. Boraston, *Proc. Natl. Acad. Sci.* **2012**, *109*, 19786–19791.
- [61] V. R. Konasani, C. Jin, N. G. Karlsson, E. Albers, *Sci. Rep.* **2018**, *8*, 14713.
- [62] L. Reisky, A. Préchoux, A. K. Zühlke, M. Baumgen, C. S. Robb, N. Gerlach, T. Roret, C. Stanetty, R. Larocque, G. Michel, S. Tao, S. Markert, F. Unfried, M. D. Mihovilovic, A. Trautwein-Schulz, D. Becher, T. Schweder, U. T. Bornscheuer, J.-H. Hehemann, *Nat. Chem. Biol.* **2019**, *16*, 6987–6996.
- [63] B. Guo, P.-Y. Li, Y.-S. Yue, H.-L. Zhao, S. Dong, X.-Y. Song, C.-Y. Sun, W.-X. Zhang, X.-L. Chen, X.-Y. Zhang, B.-C. Zhou, Y.-Z. Zhang, *Mar. Drugs* **2013**, *11*, 1173–1187.
- [64] Y. Umemoto, T. Shibata, T. Araki, *Mar. Biotechnol.* **2012**, *14*, 10–20.
- [65] T. Araki, N. Inoue, T. Morishita, *J. Gen. Appl. Microbiol.* **1998**, *44*, 269–274.
- [66] J. P. Acevedo, M. T. Reetz, J. A. Asenjo, L. P. Parra, *Enzyme Microb. Technol.* **2017**, *100*, 60–70.
- [67] Y. Hsieh, P. J. Harris, *Polymers* **2019**, *11*, 354.
- [68] H. Teeling, B. M. Fuchs, D. Becher, C. Klockow, A. Gardebrecht, C. M. Bennis, M. Kassabgy, S. Huang, A. J. Mann, J. Waldmann, M. Weber, A. Klindworth, A. Otto, J. Lange, J. Bernhardt, C. Reinsch, M. Hecker, J. Peplies, F. D. Bockelmann, U. Callies, G. Gerdt, A. Wichels, K. H. Wiltshire, F. O. Glöckner, T. Schweder, R. Amann, *Science* **2012**, *336*, 608–611.
- [69] L. Kappelmann, K. Krüger, J.-H. Hehemann, J. Harder, S. Markert, F. Unfried, D. Becher, N. Shapiro, T. Schweder, R. I. Amann, H. Teeling, *ISME J.* **2019**, *13*, 76–91.
- [70] V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho, B. Henrissat, *Nucleic Acids Res.* **2014**, *42*, D490–495.

- [71] B. L. Cantarel, P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, *Nucleic Acids Res.* **2009**, *37*, D233–D238.
- [72] G. Davies, B. Henrissat, *Structure* **1995**, *3*, 853–859.
- [73] M.-L. Garron, M. Cygler, *Glycobiology* **2010**, *20*, 1547–1573.
- [74] G. J. Davies, T. M. Gloster, B. Henrissat, *Curr. Opin. Struct. Biol.* **2005**, *15*, 637–645.
- [75] T. Barbeyron, L. Brillet-Guéguen, W. Carré, C. Carrière, C. Caron, M. Czjzek, M. Hoebeke, G. Michel, *PLOS ONE* **2016**, *11*, 1–33.
- [76] A. Levasseur, E. Drula, V. Lombard, P. M. Coutinho, B. Henrissat, *Biotechnol. Biofuels* **2013**, *6*, 41.
- [77] L. Reisky, H. C. Büchenschütz, J. Engel, T. Song, T. Schweder, J.-H. Hehemann, U. T. Bornscheuer, *Nat. Chem. Biol.* **2018**, *14*, 342–344.
- [78] E. Ficko-Blean, A. Préchoux, F. Thomas, T. Rochat, R. Larocque, Y. Zhu, M. Stam, S. Génicot, M. Jam, A. Calteau, B. Viart, D. Ropartz, D. Pérez-Pascual, G. Correc, M. Matard-Mann, K. A. Stubbs, H. Rogniaux, A. Jeudy, T. Barbeyron, C. Médigue, M. Czjzek, D. Vallenet, M. J. McBride, E. Duchaud, G. Michel, *Nat. Commun.* **2017**, *8*, 1685.
- [79] T. Barbeyron, G. Michel, P. Potin, B. Henrissat, B. Kloareg, *J. Biol. Chem.* **2000**, *275*, 35499–35505.
- [80] P. Chauhan, A. Saxena, *3Biotech* **2016**, *6*, 146.
- [81] J. Weigl, W. Yaphe, *Can. J. Microbiol.* **2011**, DOI 10.1139/m66-118.
- [82] A. G. Hettle, J. K. Hobbs, B. Pluvinage, C. Vickers, K. T. Abe, O. Salama-Alber, B. E. McGuire, J.-H. Hehemann, J. P. M. Hui, F. Berrue, A. Banskota, J. Zhang, E. M. Bottos, J. Van Hamme, A. B. Boraston, *Commun. Biol.* **2019**, *2*, 474.
- [83] M. Schultz-Johansen, P. K. Bech, R. C. Hennessy, M. A. Glaring, T. Barbeyron, M. Czjzek, P. Stougaard, *Front. Microbiol.* **2018**, *9*, 839.
- [84] A. G. Hettle, C. Vickers, C. S. Robb, F. Liu, S. G. Withers, J.-H. Hehemann, A. B. Boraston, *Structure* **2018**, *26*, 747–758.
- [85] J.-H. Hehemann, A. B. Boraston, M. Czjzek, *Curr. Opin. Struct. Biol.* **2014**, *28*, 77–86.
- [86] W.-J. Chi, Y.-K. Chang, S.-K. Hong, *Appl. Microbiol. Biotechnol.* **2012**, *94*, 917–930.
- [87] J. Gao, C. Du, Y. Chi, S. Zuo, H. Ye, P. Wang, *Mar. Drugs* **2019**, *17*, 568.
- [88] B. Pluvinage, J. M. Grondin, C. Amundsen, L. Klassen, P. E. Moote, Y. Xiao, D. Thomas, N. A. Pudlo, A. Aniele, E. C. Martens, G. D. Inglis, R. E. R. Uwiera, A. B. Boraston, D. W. Abbott, *Nat. Commun.* **2018**, *9*, 1043.
- [89] Y. Zhang, Y. Chang, J. Shen, C. Xue, *J. Agric. Food Chem.* **2019**, *67*, 9307–9313.
- [90] J. R. Turvey, J. Christison, *Biochem. J.* **1967**, *105*, 317–321.
- [91] M. Duckworth, J. R. Turvey, *Biochem. J.* **1969**, *113*, 687–692.
- [92] S. Peat, D. A. Rees, *Biochem. J.* **1961**, *79*, 7–12.
- [93] S. Becker, A. Scheffel, M. F. Polz, J.-H. Hehemann, *Appl. Environ. Microbiol.* **2017**, *83*, 1–14.
- [94] A. Gobet, T. Barbeyron, M. Matard-Mann, G. Magdelenat, D. Vallenet, E. Duchaud, G. Michel, *Front. Microbiol.* **2018**, *9*, DOI 10.3389/fmicb.2018.02740.
- [95] M. Lahaye, M. Brunel, E. Bonnin, *Carbohydr. Res.* **1997**, *304*, 325–333.
- [96] P. Nyvall Collén, J.-F. Sassi, H. Rogniaux, H. Marfaing, W. Helbert, *J. Biol. Chem.* **2011**, *286*, 42063–42071.
- [97] M. Kopel, W. Helbert, Y. Belnik, V. Buravenkov, A. Herman, E. Banin, *J. Biol. Chem.* **2016**, *291*, 5871–5878.
- [98] E. Foran, V. Buravenkov, M. Kopel, N. Mizrahi, S. Shoshani, W. Helbert, E. Banin, *Algal Res.* **2017**, *25*, 39–46.
- [99] T. Ulaganathan, R. Shi, D. Yao, R.-X. Gu, M.-L. Garron, M. Cherney, D. P. Tieleman, E. Sterner, G. Li, L. Li, R. J. Linhardt, M. Cygler, *Glycobiology* **2017**, *27*, 176–187.
- [100] T. Ulaganathan, W. Helbert, M. Kopel, E. Banin, M. Cygler, *J. Biol. Chem.* **2018**, *293*, 4026–4036.
- [101] T. Ulaganathan, E. Banin, W. Helbert, M. Cygler, *J. Biol. Chem.* **2018**, *293*, 11564–11573.
- [102] H.-M. Qin, P. Xu, Q. Guo, X. Cheng, D. Gao, D. Sun, Z. Zhu, F. Lu, *RSC Adv.* **2018**, *8*, 2610–2615.
- [103] L. Reisky, C. Stanetty, M. D. Mihovilovic, T. Schweder, J.-H. Hehemann, U. T. Bornscheuer, *Appl. Microbiol. Biotechnol.* **2018**, *102*, 6987–6996.
- [104] R. L. J. Melcher, M. Neumann, J. P. Fuenzalida Werner, F. Gröhn, B. M. Moerschbacher, *Sci. Rep.* **2017**, *7*, 44115.
- [105] C. He, H. Muramatsu, S. Kato, K. Ohnishi, *Biosci. Biotechnol. Biochem.* **2017**, *81*, 2145–2151.
- [106] T. Itoh, A. Ochiai, B. Mikami, W. Hashimoto, K. Murata, *J. Mol. Biol.* **2006**, *360*, 573–585.
- [107] A. Salinas, C. E. French, *Algal Res.* **2017**, *27*, 335–344.
- [108] T. Itoh, W. Hashimoto, B. Mikami, K. Murata, *Biochem. Biophys. Res. Commun.* **2006**, *344*, 253–262.
- [109] P. N. Collén, A. Jeudy, J.-F. Sassi, A. Groisillier, M. Czjzek, P. M. Coutinho, W. Helbert, *J. Biol. Chem.* **2014**, *289*, 6199–6211.
- [110] G. Reintjes, C. Arnosti, B. M. Fuchs, R. Amann, *ISME J.* **2017**, DOI 10.1038/ismej.2017.26.
- [111] M. Baumgen, T. Dutschei, L. Reisky, C. Stanetty, D. Bartosik, N. Gerlach, M. D. Mihovilovic, T. Schweder, J.-H. Hehemann, U. Bornscheuer, *unpublished n.d.*
- [112] E.-M. Aro, *Ambio* **2016**, *45*, 24–31.
- [113] M. Enquist-Newman, A. M. E. Faust, D. D. Bravo, C. N. S. Santos, R. M. Raisner, A. Hanel, P. Sarvabhowman, C. Le, D. D. Regitsky, S. R. Cooper, L. Peereboom, A. Clark, Y. Martinez, J. Goldsmith, M. Y. Cho, P. D. Donohoue, L. Luo, B. Lamberson, P. Tamrakar, E. J. Kim, J. L. Villari, A. Gill, S. A. Tripathi, P. Karamchedu, C. J. Paredes, V. Rajgarhia, H. K. Kotlar, R. B. Bailey, D. J. Miller, N. L. Ohler, C. Swimmer, Y. Yoshikuni, *Nature* **2014**, *505*, 239–243.
- [114] A. J. Wargacki, E. Leonard, M. N. Win, D. D. Regitsky, C. N. S. Santos, P. B. Kim, S. R. Cooper, R. M. Raisner, A. Herman, A. B. Sivitz, A. Lakshmanaswamy, Y. Kashiyama, D. Baker, Y. Yoshikuni, *Science* **2012**, *335*, 308–313.
- [115] E. J. Yun, S. Lee, H. T. Kim, J. G. Pelton, S. Kim, H.-J. Ko, I.-G. Choi, K. H. Kim, *Environ. Microbiol.* **2015**, *17*, 1677–1688.
- [116] S.-Q. Ji, B. Wang, M. Lu, F.-L. Li, *Biotechnol. Biofuels* **2016**, *9*, 81.
- [117] D. F. Rocher, R. A. Cripwell, M. Viljoen-Bloom, *Algal Res.* **2021**, *54*, 102233.
- [118] T. V. Ramachandra, D. Hebbale, *Renew. Sustain. Energy Rev.* **2020**, *117*, 109479.
- [119] D. Kaloudas, N. Pavlova, R. Penchovsky, *Environ. Chem. Lett.* **2021**, 1–16.
- [120] Y. Qui, J. Zhang, L. Li, Z. Wen, C. T. Nomura, S. Wu, S. Chen, *Biotechnol. Biofuels* **2016**, *9*, 1–13.
- [121] G. Vardar-Schara, T. Maeda, T. K. Wood, *Microb. Biotechnol.* **2008**, *1*, 107–125.
- [122] T. de Vrije, R. R. Bakker, M. A. Budde, M. H. Lai, A. E. Mars, P. A. Claassen, *Biotechnol. Biofuels* **2009**, *2*, 12.
- [123] T.-A. D. Nguyen, K.-R. Kim, M.-T. Nguyen, M. S. Kim, D. Kim, S. J. Sim, *Int. J. Hydrog. Energy* **2010**, *35*, 13035–13040.
- [124] J. S. Martín del Campo, J. Rollin, S. Myung, Y. Chun, S. Chandrayan, R. Patiño, M. W. Adams, Y.-H. P. Zhang, *Angew. Chem.* **2013**, *125*, 4685–4688.
- [125] Sulfahri, S. Mushlihah, D. R. Husain, A. Langford, A. C. M. A. R. Tassakka, *J. Clean. Prod.* **2020**, *265*, 121763.
- [126] M. Dinesh Kumar, R. Yakesh Kannah, G. Kumar, P. Sivashanmugam, J. Rajesh Banu, *Bioresour. Technol.* **2020**, *301*, 122759.
- [127] M. Ghareib, K. A. Youssef, A. A. Khalil, *Folia Microbiol. (Praha)* **1988**, *33*, 447–452.
- [128] L. Olsson, B. Hahn-Hägerdal, *Enzyme Microb. Technol.* **1996**, *18*, 312–331.
- [129] C. Martín, M. Galbe, C. F. Wahlbom, B. Hahn-Hägerdal, L. J. Jönsson, *Enzyme Microb. Technol.* **2002**, *31*, 274–282.
- [130] E. J. Yun, S. Lee, J. H. Kim, B. B. Kim, H. T. Kim, S. H. Lee, J. G. Pelton, N. J. Kang, I.-G. Choi, K. H. Kim, *Appl. Microbiol. Biotechnol.* **2013**, *97*, 2961–2970.

Entry for the Table of Contents



Algae biomass is a source for many valuable compounds for the pharmaceutical and chemical industry. Especially marine polysaccharides contain rare sugars of interest, but a major bottleneck is access to these sugars. In this mini-review the common knowledge of marine CAZyme-based deconstruction of these polysaccharides and the biotechnological potential of the thus accessible oligo- and monosugars are described.

Institute and/or researcher Twitter usernames: University of Greifswald: @uni_greifswald (https://twitter.com/uni_greifswald)

Article II



A new carbohydrate-active oligosaccharide dehydratase is involved in the degradation of ulvan

Received for publication, August 13, 2021, and in revised form, September 13, 2021. Published, Papers in Press, September 20, 2021, <https://doi.org/10.1016/j.jbc.2021.101210>

Marcus Bäumgen^{1,‡}, Theresa Dutschei^{1,‡}, Daniel Bartosik², Christoph Suster³, Lukas Reisky¹, Nadine Gerlach^{4,5}, Christian Stanetty³, Marko D. Mihovilovic³, Thomas Schweder², Jan-Hendrik Hehemann^{4,5}, and Uwe T. Bornscheuer^{1,*}

From the ¹Department of Biotechnology & Enzyme Catalysis, Institute of Biochemistry, ²Department of Pharmaceutical Biotechnology, Institute of Pharmacy, University Greifswald, Greifswald, Germany; ³Institute of Applied Synthetic Chemistry, TU Wien, Vienna, Austria; ⁴Max Planck-Institute for Marine Microbiology, Bremen, Germany; ⁵Center for Marine Environmental Sciences (MARUM), University of Bremen, Bremen, Germany

Edited by Chris Whitfield

Marine algae catalyze half of all global photosynthetic production of carbohydrates. Owing to their fast growth rates, *Ulva* spp. rapidly produce substantial amounts of carbohydrate-rich biomass and represent an emerging renewable energy and carbon resource. Their major cell wall polysaccharide is the anionic carbohydrate ulvan. Here, we describe a new enzymatic degradation pathway of the marine bacterium *Formosa agariphila* for ulvan oligosaccharides involving unsaturated uronic acid at the nonreducing end linked to rhamnose-3-sulfate and glucuronic or iduronic acid (Δ -Rha3S-GlcA/IdoA-Rha3S). Notably, we discovered a new dehydratase (P29_PDnc) acting on the nonreducing end of ulvan oligosaccharides, *i.e.*, GlcA/IdoA-Rha3S, forming the aforementioned unsaturated uronic acid residue. This residue represents the substrate for GH105 glycoside hydrolases, which complements the enzymatic degradation pathway including one ulvan lyase, one multimodular sulfatase, three glycoside hydrolases, and the dehydratase P29_PDnc, the latter being described for the first time. Our research thus shows that the oligosaccharide dehydratase is involved in the degradation of carboxylated polysaccharides into monosaccharides.

Marine algae catalyze half of the global photosynthetic production of carbohydrates (1). Fast growth makes macroalgae a promising renewable bioresource for the chemical, pharmaceutical, agricultural, and food industry (2–8). Exemplarily, the ubiquitous green seaweed *Ulva* spp. has been recently suggested as a source of bioactive and rare sugars (9, 10). Ulvan, which is branched and highly sulfated, is the major cell wall polysaccharide of the ‘green tide’ causing macroalgae *Ulva* spp. Ulvan can represent up to 30% of the algal dry weight (2). The major disaccharide repeating units are ulvanobiouronic acid A (β -D-glucuronic acid (GlcA)-(1,4)- α -L-rhamnose-3-sulfate (Rha3S)), ulvanobiouronic acid B (α -L-iduronic acid (IdoA)-(1,4)- α -L-rhamnose-3-sulfate), ulvanobiose-3-sulfate (β -D-xylose (Xyl)-

(1,4)- α -L-rhamnose-3-sulfate), and ulvanobiose-2',3-disulfate (β -D-xylose-2-sulfate (Xyl2S)-(1,4)- α -L-rhamnose-3-sulfate). A modification of Rha3S by β -1,2-linked GlcA side chains and the appearance of consecutive GlcA residues have been described as well (2).

In polysaccharide degrading bacteria, the genes encoding for the carbohydrate-active enzymes (CAZymes) often colocalize with transporter genes in gene clusters called ‘polysaccharide utilization loci’ (PULs) (11) as it is the case for the ulvan pathway of the marine *Flavobacterium Formosa agariphila* KMM 3901T (12). Recently, we elucidated the degradation cascade for ulvan consisting of 12 carbohydrate-active enzymes, including two polysaccharide lyases, three sulfatases, and seven glycoside hydrolases (10). This pathway enables the degradation of ulvan into monosaccharides.

Polysaccharide lysases (PL) of the families PL24, PL25, PL28, and PL40 catalyze the initial degradation step of ulvan into oligosaccharides *via* an elimination mechanism. This mechanism forms an unsaturated uronic acid residue at the nonreducing end (10, 13–19), which then can be cleaved by unsaturated glucuronyl hydrolases of the families GH88 or GH105 (2, 10, 20, 21). Importantly, this unsaturated sugar is required for the activity of GH105 and GH88. In the above-mentioned pathway, there are also glycoside hydrolases of family GH78 (P36_GH78) that lead to the formation of a nonreducing end with a saturated uronic acid sugar. This one can only be cleaved by a GH3 family enzyme.

Here, we describe the discovery of a new class of ulvan-active dehydratases. This new enzyme type converts saturated uronic acid sugars such as GlcA/IdoA-Rha3S at the nonreducing end of oligosaccharides into unsaturated sugars enabling cleavage by GH105. This is the first time a dehydratase was reported as an enzyme active on carbohydrates.

Results

The initial degradation step catalyzed by the ulvan lyases (P10_PLnc and P30_PL28) leads to the formation of several oligosaccharides with diverse composition, see also Fig. S1 (10). This includes glucuronic or iduronic acid containing

[‡] These authors contributed equally to this work.

* For correspondence: Uwe T. Bornscheuer, uwe.bornscheuer@uni-greifswald.de.

New carbohydrate dehydratase in ulvan degradation

tetramers, which are the result of an incomplete digestion. Even though these oligosaccharides can be degraded by the ulvan lyases to the dimer Δ -Rha3S (10, 18), the oligosaccharides Δ -Rha3S-GlcA-Rha3S and Δ -Rha3S-IdoA-Rha3S accumulated when ulvan was degraded, because high concentrations of lyase products inhibit the reaction (22). This was our motivation to search for suitable enzyme activities within PUL H of *F. agariphila* that support the conversion of these intermediates. Comparative genomics of 12 bacteroidetal ulvan PULs revealed the presence of a conserved unknown enzyme, P29_PDnc, in *F. agariphila* KMM 3901, which showed an induced expression during cultivation with ulvan as sole carbon source (Fig. 1) (10).

Biochemical characterization led to the discovery of a new enzymatic function in the degradation of uronic polysaccharides such as ulvan. This enabled the complete degradation of Δ -Rha3S-GlcA-Rha3S and Δ -Rha3S-IdoA-Rha3S (Fig. 2, see also Fig. S1). As it was not possible to separate these two oligosaccharides, this mixture was designated as Δ -Rha3S-GlcA/IdoA-Rha3S. At the first step of the degradation cascade, the *exo*-acting unsaturated glucuronyl hydrolase (P33_GH105) cleaves the unsaturated uronyl residue from the nonreducing end (9, 22, 23). The products of this reaction are 5-dehydro-4-deoxy-D-glucuronate and the trisaccharide Rha3S-GlcA/IdoA-

Rha3S. Its chemical structure, which was previously confirmed (10), shows similarity to Rha3S-Xyl-Rha3S and Rha3S-Xyl2S-Rha3S (Table S1 and Figs. S2–S5), which are substrates for the sulfatase domain of P36_S1_25 (10). The P36 is a multi-domain protein, which consists of the sulfatase domain S1_25 and an α -L-rhamnosidase domain GH78 (9). Indeed, the sulfatase converts all three products irrespectively of the sugar species located between the two flanking rhamnose residues. In all three cases, it desulfates the rhamnose residue at the nonreducing end so that this sulfatase is active on ulvan trisaccharides with the general structure Rha3S-XXX-Rha3S (10). The desulfated trisaccharide Rha-GlcA/IdoA-Rha3S was isolated confirming the desulfation at the nonreducing end (Table S2 and Figs. S8–S11). Analogous to the already established pathway, Rha-GlcA/IdoA-Rha3S is now degraded by the α -L-rhamnosidase domain of P36_GH78 leading to the removal of the rhamnose residue at the nonreducing end (Table S3 and Figs. S12–S19). This confirms that both enzyme domains of P36 act in consecutive steps within the ulvan degradation on multiple substrate molecules. This makes P36 the first discovered multimodular enzyme participating in ulvan degradation. The carbohydrate structures along the pathway to the reaction product GlcA/IdoA-Rha3S were elucidated by NMR spectroscopy (Tables S1–S3 and

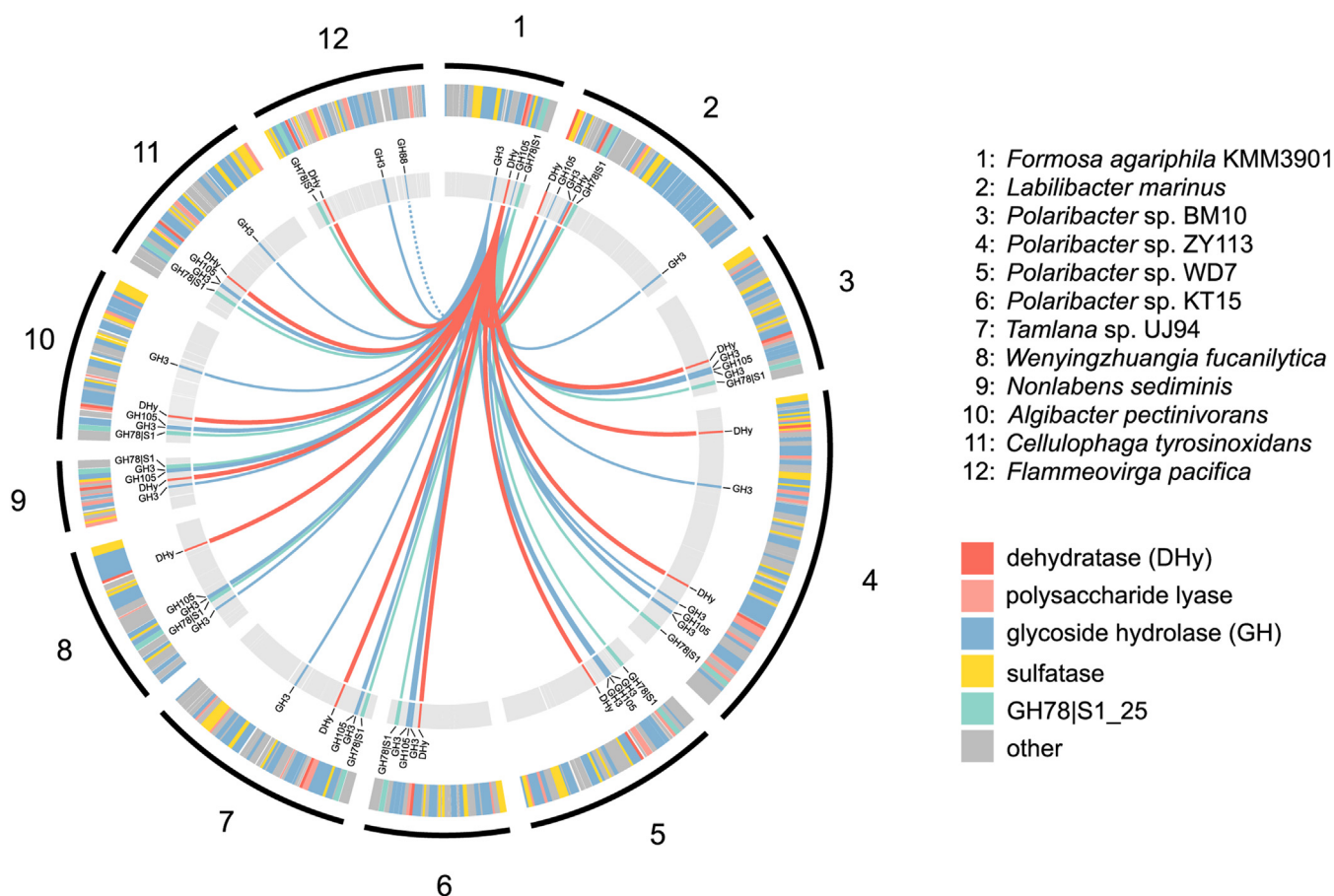


Figure 1. Genomic overview of putative ulvan PULs containing the alternative pathway analogues genes in 12 marine Bacteroidetes. Ulvan PUL annotations are given in the *outer ring*, whereas alternative pathway analogues genes for CAZymes and the GH78|S1_25 hybrid are highlighted within the *inner ring*, linked to the model organism of this study, *F. agariphila* KMM 3901. The *dashed line* indicates the closely related GH88 family of GH105.

New carbohydrate dehydratase in ulvan degradation

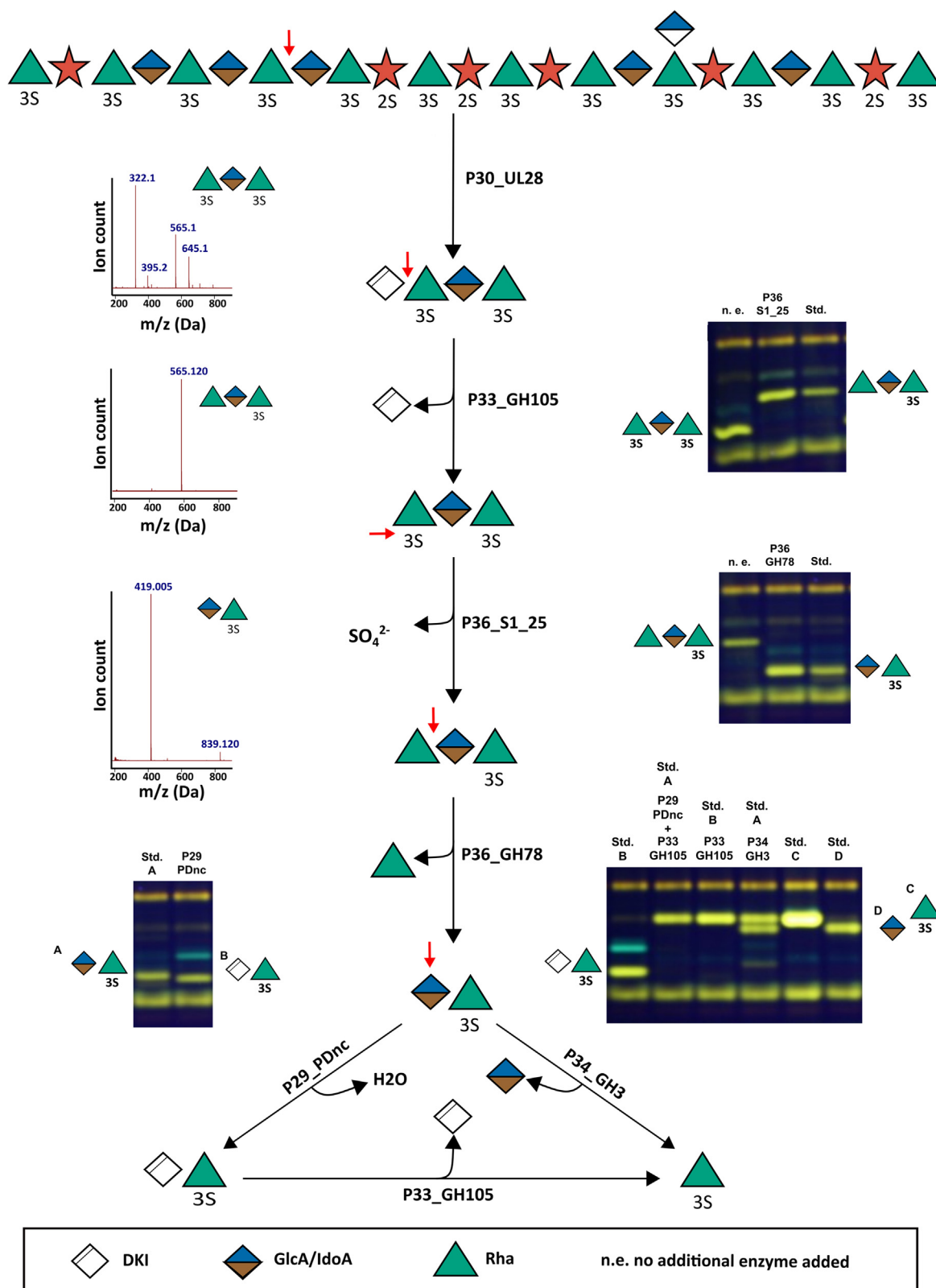


Figure 2. Model of the alternative ulvan degradation pathway of *F. agariphila* based on FACE analyses and MS data. The activity of the P33_GH105 to cleave 4-deoxy- α -L-threo-hex-4-enopyranuronic acid from the tetramer Δ -Rha3S-GlcA/IdoA-Rha3S to the trimers Rha3S-GlcA/IdoA-Rha3S was previously described by Reisky *et al.* (10). In reactions containing P29_PDnc, the other enzymes were heat-inactivated before addition of P29_PDnc, to prevent a degradation of the dehydratase product by P33_GH105. All products and standards (except GlcA) used had been isolated and confirmed by MS and NMR measurement. The MS data were derived from the HPLC-ELS-MS measurement from the purified oligomers from the natural product. The last lane of the *top gel* is the same as the first lane of the *middle gel*, and the last lane of the *middle gel* is the same as the first lane of the *left gel* while the *right lane* of the *left gel* is the first lane of the last *right gel* to ensure continuity in the explanation of the degradation pathway. The full FACE-Gel picture can be found in the [Supporting information](#) (Fig. S20). The standard for GlcA was obtained from the company Carl-Roth. All products represent the mixture of both oligomers containing one of the epimers GlcA or IdoA. The ratio between GlcA- and IdoA-containing oligomers is $\sim 70:30$ (18). The 4-deoxy- α -L-threo-hex-4-enopyranuronic acid is abbreviated with "unsaturated uronic acid."

New carbohydrate dehydratase in ulvan degradation

Figs. S2–S19). Proceeding from this dimeric intermediate, the pathway splits into two subpathways (Fig. 2), which could be identified by screening all produced enzymes encoded by PUL H. The first way to digest the disaccharide GlcA/IdoA-Rha3S is to use P34_GH3 (Fig. 2). This glycoside hydrolase cleaves the glucuronic or iduronic acid residues with release of rhamnose-3-sulfate, making it a promiscuous β -glucuronidase/ α -iduronidase. The second way is to use the conserved hypothetical protein P29_PDnc in the first step (Fig. 2). This enzyme converts the disaccharide GlcA/IdoA-Rha3S into the formerly described disaccharide Δ -Rha3S by elimination of water. The disaccharide Δ -Rha3S can also be produced using the previously described ulvan lyases, see also Fig. S1 (10). P33_GH105 hydrolyzes the formed disaccharide as described before, leading to the formation of Rha3S and 5-dehydro-4-deoxy-D-glucuronate, which confirms the dehydratase activity of P29_PDnc (Fig. 2). Increased abundance of these four CAZymes in ulvan-grown *F. agariphila* KMM 3901 (10), as well as the occurrence in diverse marine Bacteroidetes (Fig. 1), emphasizes the importance of this alternative pathway for the efficient degradation of ulvan.

The two resulting products of both pathways—rhamnose-3-sulfate and glucuronic or iduronic acid—were confirmed by using commercial substrates as standards (Fig. 2). Thereby,

P29_PDnc was identified to be a novel type of ulvan-active dehydratase that participates in the degradation of ulvan. This is the first time a dehydratase was described to be acting in the depolymerization of a carbohydrate. Other described sugar-active dehydratases usually catalyze monosaccharide-related reactions (24–26).

In previous studies, P29_PDnc was reported to be an ulvan lyase with broad substrate spectrum (25). In contrast, in this work no activity of this enzyme against polymeric ulvan from seven different sources could be observed. One difference between the constructs used in our study and the published example is the position of the His-tag at the investigated heterologous produced enzymes. To investigate if the His-tag position influences the enzymatic activity, two different variants, one with N-terminal and the other with C-terminal His-tag, were prepared. A spectrophotometric lyase assay was used to determine the double bond formation, which is characteristic for the lyase activity (18). Carbohydrate polyacrylamide gel electrophoresis (C-PAGE) was used to visualize the breakdown products, and a reducing-end assay was used to estimate the reducing ends resulting from this cleavage process (Fig. 3, A–C and Figs. S21–S27). However, we still could not detect lyase activity in any of the C-terminal or N-terminal His-Tag P29_PDnc variants on ulvans from seven different

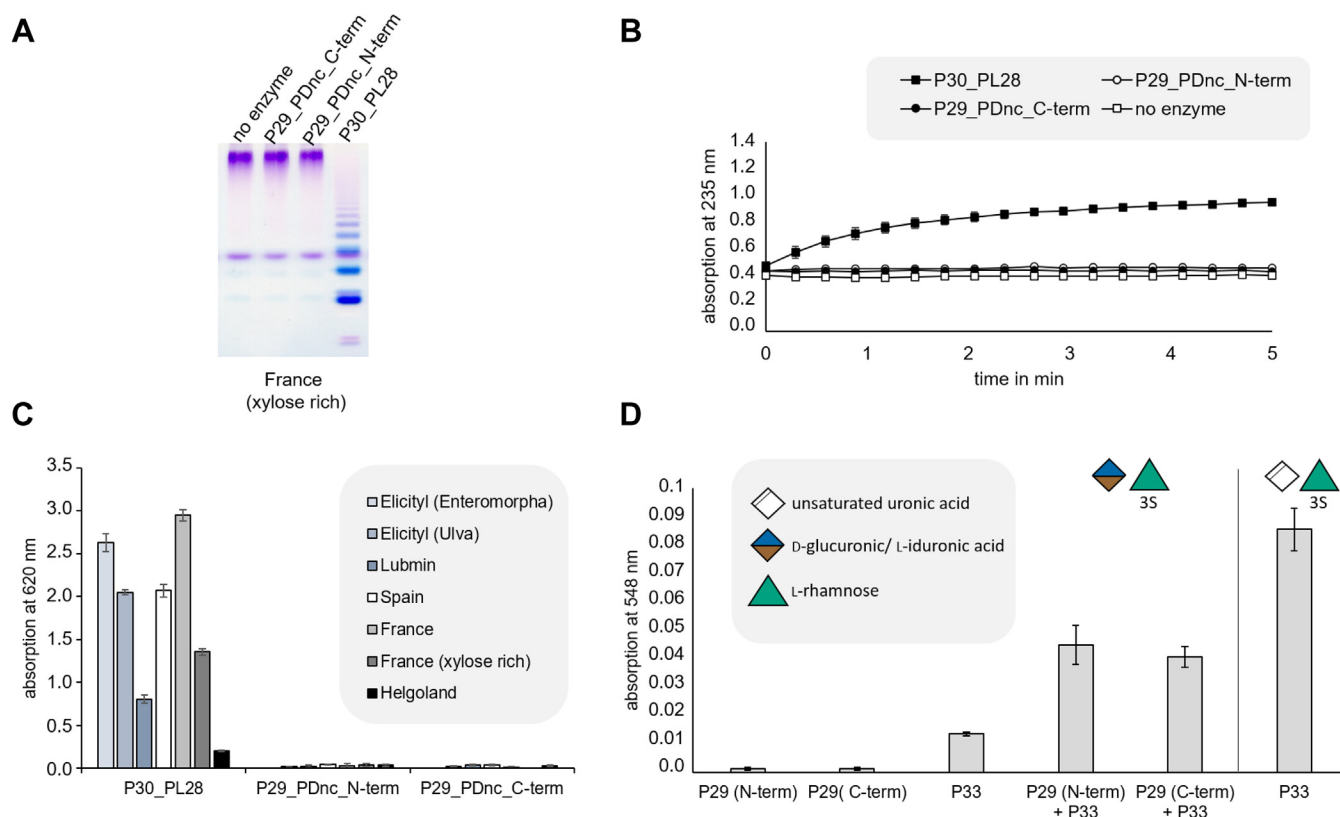


Figure 3. Analysis of lyase and dehydratase activity of P29_PDnc. A, C-PAGE analysis, (B) lyase assay with ulvan from France (xylose rich), (C) reducing-end assay, and (D) thiobarbituric acid assay. Polymeric ulvan from seven different sources has been incubated with both P29_PDnc variants with N- or C-terminal His-tag and P30_PL28 as positive control or without enzymes as negative control. We used in these experiments two commercially available ulvans from Elicityl extracted from *Enteromorpha* sp. or *Ulva* sp., and five self-isolated ulvans from “kulau sea lettuce” containing *Ulva* spp. from Spain, and from self-collected *Ulva* sp. from Helgoland (North Sea), France (Atlantic Ocean) and Lubmin (Baltic Sea) (see Figs. S20–S26). Purified GlcA/IdoA-Rha(3S) was incubated with both P29_PDnc variants with N- or C-terminal His-tag and/or P33_GH105 or purified Δ -Rha(3S) was incubated with P33_GH105 as a positive control. The resulting reaction mixture was investigated using the thiobarbituric acid assay for the determination of 5-dehydro-4-deoxy-D-glucuronate (23).

sources, while the positive control P30_PL28 (18) showed activity (Fig. 2 and Figs. S21–S27). However, the P29_PDnc activity against the disaccharide GlcA/IdoA-Rha3S could be confirmed by a thiobarbituric acid assay (23) (Fig. 3D). Both C-terminal and N-terminal His-Tag P29_PDnc variants and the supporting enzyme P33_GH105, which releases the 5-dehydro-4-deoxy-D-glucuronate after the dehydratase reaction, were incubated with the target disaccharide. As a positive control P33_GH105 was also used in biocatalysis reactions on the lyase-produced disaccharide Δ -Rha3S, which is the reaction product of the dehydratase reaction. The batches containing only a C-terminal or N-terminal His-Tag variant of P29_PDnc showed no absorption as the P33_GH105, which is supposed to release the formed 5-dehydro-4-deoxy-D-glucuronate, is missing. P33_GH105 on its own induced a very small absorption, but only a combination of P29_PDnc and P33_GH105 led to a significant signal. However, this is far from the absorption values observed for the positive control using P33_GH105 on the disaccharide Δ -Rha3S (Fig. 3D). The thiobarbituric acid assay thereby confirms the results obtained by FACE analysis that P29_PDnc converts GlcA/IdoA-Rha3S to Δ -Rha3S, which can be targeted by P33_GH105. As the substrate GlcA/IdoA-Rha3S is converted by both, P29_PDnc and P34_GH3, it was possible to reverse the dehydratase reaction by shifting the equilibrium to the educt side (Fig. S28). When the reaction was carried out in deuterium

oxide (D_2O), the deuterium was inserted at the double bond, which could be identified and confirmed by mass spectrometry (Fig. S29).

However, the true mechanism was elusive, as an oligosaccharide dehydratase of this type had never been reported before. When comparing the two most closely related mechanisms—that of PLs and of monosaccharide dehydratases—there are less similarities of P29_PDnc with the sugar dehydratases, as they all dehydrate side chains of monosugars, leading to an elimination of water from a side chain hydroxyl group, forming a desoxy sugar. The monosugar dehydratases convert their substrate *via* an oxidation mechanism (26, 27). Furthermore, P29_PDnc seems to be cofactor-independent as the reaction worked without addition of any supplements. The activity of P29_PDnc leads to a dehydration of a ring hydroxyl group resulting in the formation of an unsaturated hexenuronic acid residue. It seems to require a C5 carboxyl group like common PLs. Thus, the mechanism follows most probably the general PL mechanism, but eliminating water instead of a sugar residue. To reveal residues that are presumably involved in this catalysis, alignment studies were performed. Variants of P29_PDnc were produced with mutation of single functional amino acid residues (Fig. 4) to residues with similar structure, but different chemical properties, to ensure that these mutations lead to a loss of activity, because the residue is important for the catalysis and that not a structural change of the active

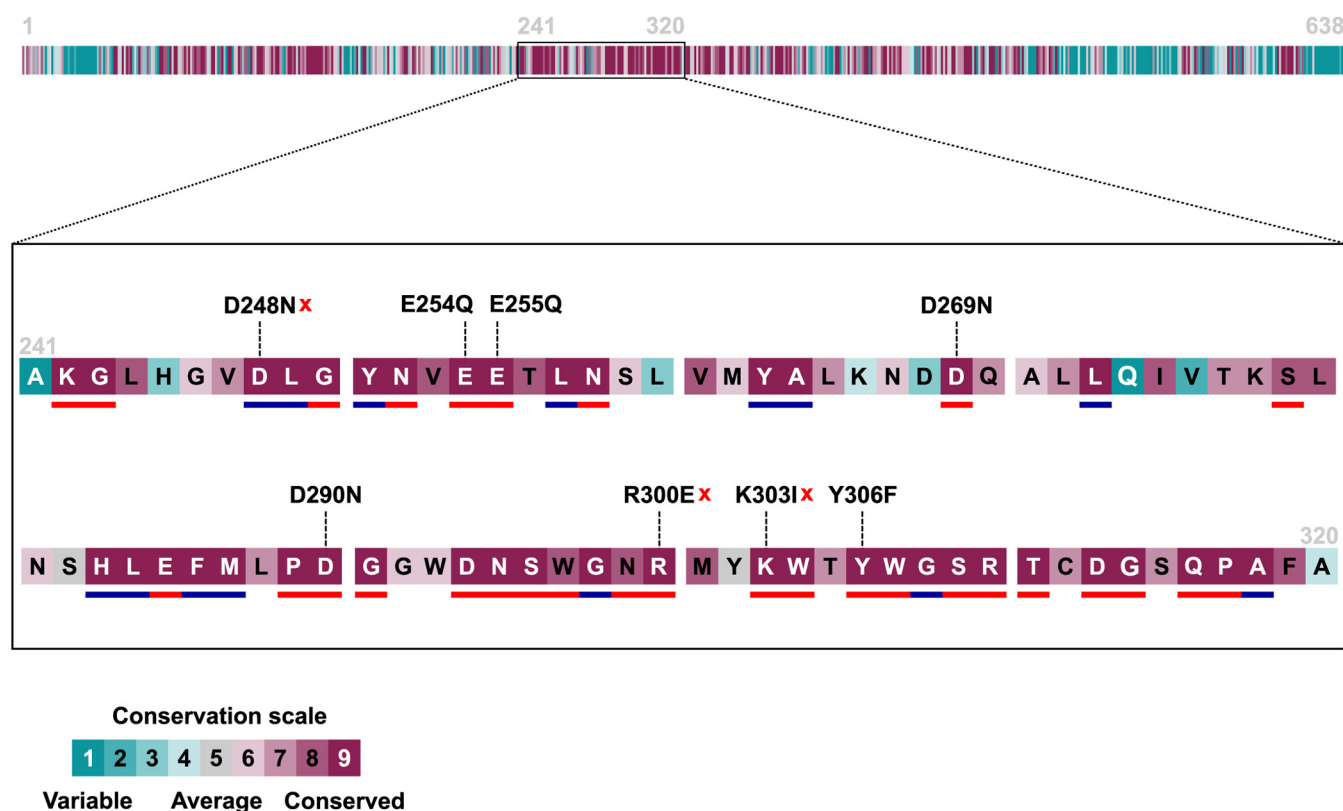


Figure 4. Conserved residues of P29_PDnc sequence WP_038530528.1. Conservation was determined using the ConSurf Server with the multiple sequence alignment of 25 phmmer hits *versus* UniProt reference proteomes given in Figure 6. This graphic was adapted from the graphical output of the ConSurf Server. The amino acids with a high conservation score (purple) and functional prediction (red lines under the main sequence) were chosen for the mutations to investigate for functionality of the P29_PDnc. Mutations, which lead to inactivation of the P29_PDnc, are marked with a red cross. Blue underlined residues are predicted as structural residue.

New carbohydrate dehydratase in ulvan degradation

site leads to the loss of activity. The P29_PDnc variants D248N, E254Q, E255Q, D269N, D290N, R300E, K303I, and Y306F were generated (Fig. 4). Interestingly, the residue Y306F showed still activity, after mutating it to phenylalanine as there was full conversion as shown in the FACE analysis (Fig. 5). Tyrosine is the catalytic residue in ulvan lyases as well as in many dehydratases (15, 26, 27), which led to the assumption that this is also the case for the novel ulvan PUL encoded dehydratase described here. As the mutation did not inactivate this variant, a catalytic participation of Y306 is rather unlikely. Three mutations led to a loss of activity indicated by the lack of conversion observed by the FACE analysis: D248N, R230E and K303I (Fig. 5). The arginine (R230) is presumably involved in the stabilization of the C5 carboxyl group of the substrate as it was reported to be the case in some PLs (13). The aspartate (D248N) most probably serves as the catalytic base abstracting the ring proton, which

in turn initializes the formation of the enolate intermediate like it is the case in the lyase mechanism (13). The lysine (K303I) presumably provides a proton for the leaving group water and takes the role of the catalytic acid (13). These results led to the proposed preliminary mechanism shown in Figure 5.

Discussion

In this study, we were able to complement the complex ulvan degradation pathway previously described by Reisky *et al.* (10) by elucidating an alternative enzyme cascade, which is able to fully degrade uronic-acid-containing oligosaccharides resulting from an incomplete degradation by ulvan lyases. Biochemical characterization of each step of the cascade with purified enzymes and structural determination of the produced intermediates enabled us to discover a new branch of

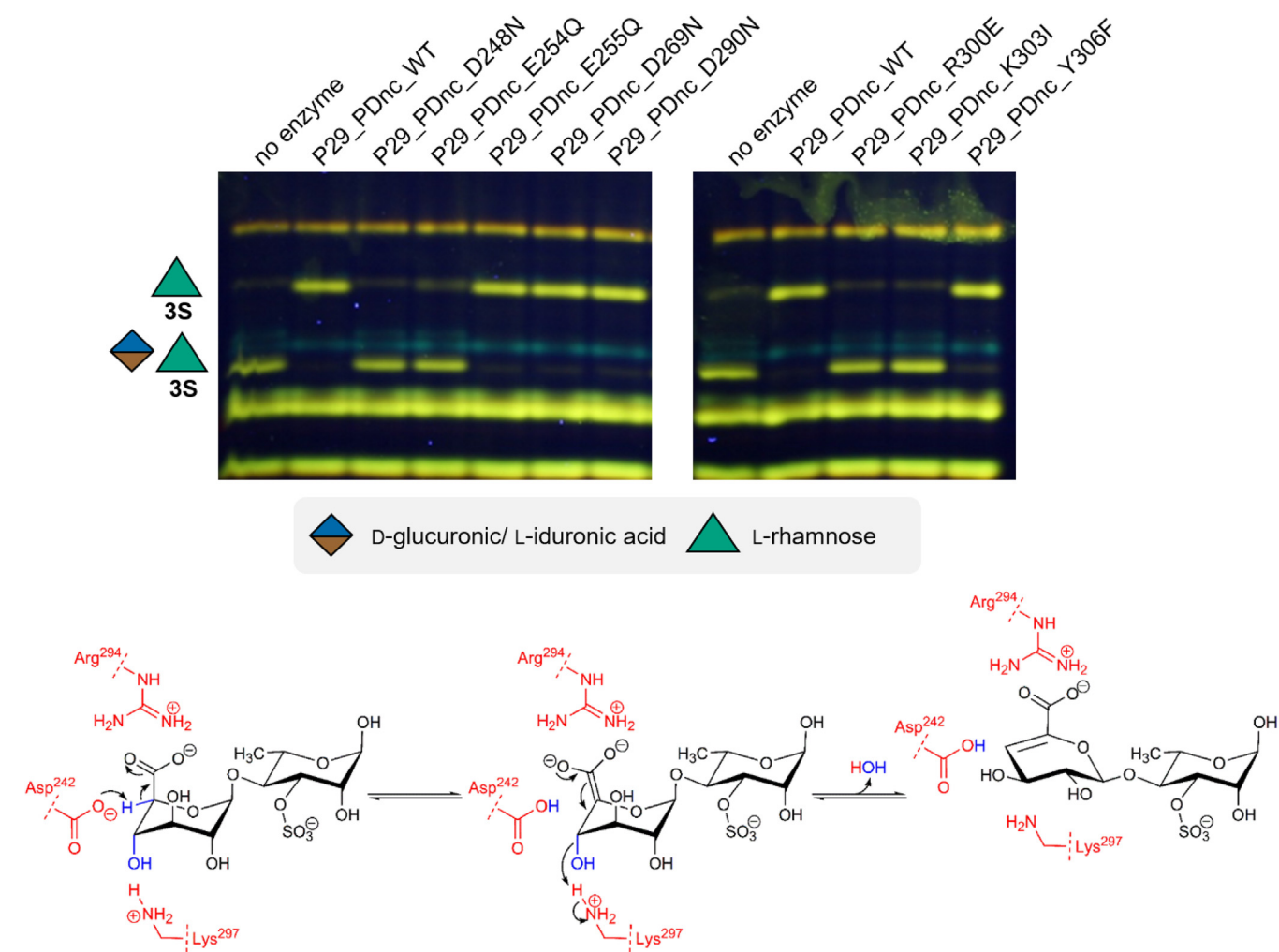


Figure 5. Proposed preliminary reaction mechanism for the polysaccharide dehydratase P29_PDnc. The variants of P29_PDnc were incubated with the disaccharide GlcA/IdoA-Rha3S produced from Δ -Rha3S-GlcA/IdoA-Rha3S with enzymes from the alternative pathway leading to the formation of the free α -keto acid and Rha3S. For the negative control, no P29_PDnc-single point mutation variant was added to the enzyme reaction with the enzymes from the alternative pathway. The P33_GH105 was not heat-inactivated as the produced rhamnose-3-sulfate can be distinguished much easier from GlcA/IdoA-Rha3S. A number in combination with an 'S' attached to a sugar represents the position of sulfate groups. "Unsaturated uronic acid" represents 4-deoxy- α -L-threo-hex-4-enopyranuronic acid. The mechanism shows the proposed reaction for the disaccharide β -D-glucuronic acid 1,4-linked to α -L-rhamnose-3-sulfate. For simplification, only the preliminary functional amino acid side chains are shown (red). The C5 proton (blue) is abstracted by Asp242, while Arg294 stabilizes the oxanion intermediate. With support of Lys297, which protonates the glycosidic oxygen atom, water is eliminated under formation of the characteristic 5-dehydro-4-deoxy-D-glucuronate.

the complex ulvan degradation pathway in *F. agariphila* (Fig. 2, see Fig. S1 for a comparison). In addition to the previously described ulvan-degrading enzymes (P10_PL40, P17_GH2, P18_S1_7, P20_GH78, P24_GH3, P27_GH43, P30_PL28, P31_GH39, P32_S1_8, P33_GH105, P36_S1_25), we were able to elucidate the function of three further enzymes in the ulvan utilization comprising two glucoside hydrolases (P34_GH3, P36_GH78) and an oligosaccharide dehydratase (P29_PDnc). P36_GH78 could also be confirmed to be the first multimodular enzyme participating in ulvan degradation as enzyme P36 combines a sulfatase domain with a GH78 domain. These two domains were analyzed separately. Furthermore, a new substrate specificity in this pathway could be observed for the previously described sulfatase of family S1_25 (P36_S1_25).

We found that the here described new alternative ulvan specific degradation pathway is conserved in other marine Bacteroidetes (Fig. 1), whereas sediment-associated Planctomycetes encode a more distant dehydratase homolog without a GH105, GH88, and GH3 context (Fig. 6).

This elucidation of an alternative degradation pathway illustrates the complexity of the biological systems for marine ulvan degradation. It indicates the necessity of backup mechanisms for metabolic processes in order to get access and compete for the diversity of complex marine carbon sources in nature. Several small degradation cascades complement each other to break substrate compounds down to

the monomeric level for the use of the structurally diverse polysaccharide ulvan. The here described ulvan specific subpathway in general and the newly described dehydratase activity in particular enable a more efficient ulvan utilization. The higher the ulvan concentration is, the more the ulvan lyases are inhibited by their own products. This is prevented by the proposed alternative pathway, which is able to take over the function of the inhibited lyase cascade to ensure a more efficient utilization of ulvan sugars as energy and carbon sources.

The findings of this study expand our insights into the metabolic processes of the degradation of a complex marine polysaccharide and thus help to elucidate specific molecular mechanisms of the ocean's carbon cycle. The characterizations of ulvan-active enzymes and the clarification of their substrate scopes allow using these enzymes for the production of ulvan-derived chemical products from currently rarely used green algal biomass.

Experimental procedures

Comparative genomics

Bacterial genomes were chosen according to (10) and downloaded from NCBI-GenBank. Ulvan PUL encoded carbohydrate-active enzymes were identified using the dbCAN meta server (<http://bcb.unl.edu/dbCAN2>) integrated tools HMMER, DIAMOND, and Hotpep (tool versions and databases as of 09/28/19) and assigned to CAZyme families

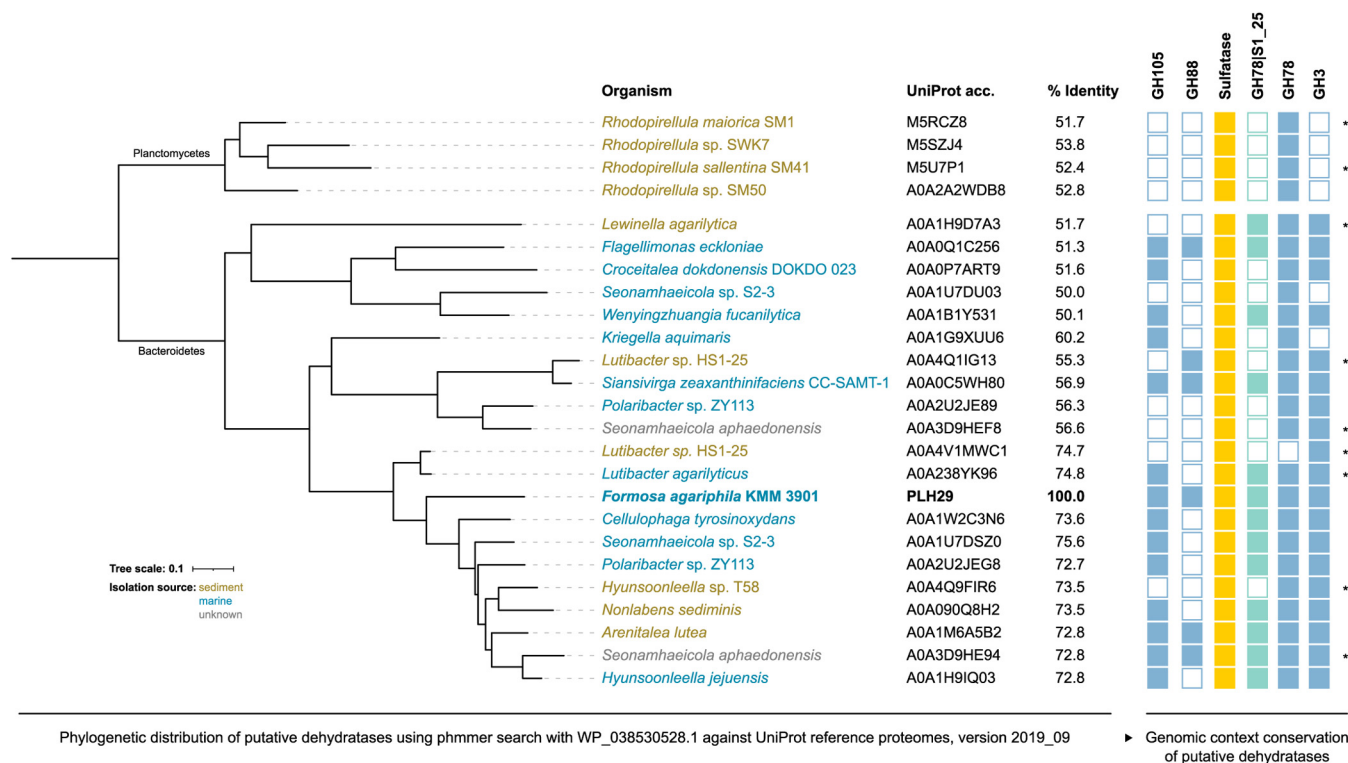


Figure 6. Phylogenetic distribution of putative dehydratases in different organisms. Phylogenetic analysis was done using the phmmer web server search with WP_038530528.1 against UniProt reference proteomes, version 2019_09. Only sequences with a percentage identity of at least 50% are shown. The right panel shows the presence or absence of alternative pathway encoded CAZymes and sulfatases of the respective dehydratases. Gene cluster highlighted with an asterisk (*) might be incomplete due to limited contig length.

New carbohydrate dehydratase in ulvan degradation

if the carbohydrate-active enzyme was found by at least two tools (28). Putative dehydratases were predicted using the PL37 family. Sulfatases were annotated using HMMER v3.2.1 (29) against the Pfam profile PF00884. Pfam hits were further filtered with the dbCAN hmmer-scan-parser script and assigned to a sulfatase family using Protein-Protein BLAST v2.11.0+ (30) against the Sulfatlas database v1.1 (31) with an expect value threshold of 1E-5. Ulvan PULs were predicted by searching annotations of ulvan lyases and the surrounding genes using an up- and downstream distance to any CAZyme (excluding glycosyl transferases) or sulfatase of up to seven genes. Circos was used to visualize the results (32).

In order to determine a broader phylogenetic distribution and therefore substrate specificity, the P29_PDnc sequence WP_038530528.1 was searched against the uniprotrefprot database v2019_09 using phmmer v3.3.2 with the following settings “-E 1-domE 1-incE 0.01-incdomE 0.03-mx BLOSUM62-pextend 0.4-popen 0.02”, including all taxa (33). Hits with a percentage identity of at least 50% were aligned using CLUSTALW (34) with default settings. Maximum-likelihood phylogeny was estimated using PhyML webserver (default settings) (35, 36). The resulting tree was visualized with iTOL (37). Genomes of putative dehydratase encoding organisms were downloaded and analyzed as described above. Functionally and structurally important residues of WP_038530528.1 were analyzed and visualized using ConSurf web server (38, 39).

Gene cloning

Expression constructs were prepared using the FastCloning strategy (40) with genomic DNA from *F. agariphila* KMM 3901T (collection number DSM15362 at DSMZ) as template for the amplification of the inserts. Generally, the pET28 constructs were prepared as described previously (10) with the gene primers shown in Table S4. To clone the gene for the formylglycine-generating enzyme (FGE) from *F. agariphila*, the vector backbone was amplified with the primers 5'-AATA GCGC CGTC GACC ATCA TCAT CATC ATCAT-3' and 5'-CATG GTTA ATTC CTCC TGTT AGCC CAAA AA-3' from pBAD/myc-his A (Addgene). The gene of the sulfatase P36_S1_25 was ordered codon-optimized for *Escherichia coli* from Genscript and subcloned into pET28 with *NheI* and *XhoI*. The optimized nucleotide sequence is provided in the Supporting information.

The gene of the glycoside hydrolase P36_GH78 was a kind gift from Gurvan Michel (Station Biologique de Roscoff, Roscoff, France).

Enzyme production

E. coli BL21(DE3) was transformed with pET28-based plasmids harboring the required genes. For the overexpression, 50 ml LB or TB medium with 100 µg ml⁻¹ kanamycin was inoculated from an overnight culture in LB containing 50 µg ml⁻¹ kanamycin. The culture was grown at 37 °C and 180 rpm until the OD600 nm reached 0.8. The expression was

then induced by adding 0.5 or 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and the culture was cooled to 20 °C (or to 16 °C in case of the P29-variants), for 24 h. For the expression of sulfatase genes, the FGE from *F. agariphila* was coexpressed. LB medium with 100 µg ml⁻¹ ampicillin and 50 µg ml⁻¹ kanamycin was inoculated from an overnight culture in the same medium and incubated at 37 °C and 180 rpm until the OD600 nm reached 0.3 to 0.5. After the addition of 1.5 mM L-arabinose and incubation for 90 min at 37 °C, the culture was cooled to 18 °C for 2 h before 0.5 mM IPTG was added, and the culture was incubated overnight at 18 °C.

SDS-PAGE

Samples from the cultivations equivalent to a volume of 7/ OD600 nm in ml were taken before harvest and the cells were collected by centrifugation (13,000g, 4 °C, 2 min). Pellets were resuspended in 500 µl 50 mM HEPES with 100 mM NaCl (pH 7.4). After lysis with FastPrep cell disruptor (MP Biomedicals), whole cell protein (W) samples were obtained prior to removal of the cell debris by centrifugation (13,000g, 4 °C, 10 min). Samples of the soluble protein fraction (S) were taken from the respective supernatant. For the SDS-PAGE, 12.5% acrylamide gels were used containing 1% (v/v) 2,2,2-trichloroethanol for the visualization of proteins under UV light (41). Electrophoresis was carried out at 200 V and gels were placed on a UV transilluminator for 2 min to develop the fluorescence after which pictures were taken. Alternatively, proteins were stained with Coomassie Blue (PhastGel Blue R, Sigma Aldrich).

Enzyme purification

The cell pellets of a 50 ml culture were thawed on ice and resuspended in 10 ml of ice-cold Tris-HCl buffer (50 mM, pH 7.4 + 300 mM NaCl +10 mM imidazole) (wash buffer). The cells were lysed by ultrasonication on ice (2 × 3 min, 50% power, 50% cycle time), and the cell debris was removed by centrifugation (15 min at 10,000g). Rotigaroze-His/Ni beads (Carl Roth) incubated with the clarified lysate were used in gravity flow columns. After washing, the protein was eluted with Tris-HCl-buffer (50 mM, pH 7.4 + 100 mM NaCl +300 mM imidazole). Fractions containing the protein of interest were pooled and desalted using PD-10 columns (GE Healthcare) equilibrated with 50 mM Tris-HCl (pH 7.4 + 10 mM NaCl). Alternatively, the same buffers with pH 8.0 instead of 7.4 could be used without any effect on the protocol. The desalted enzymes were aliquoted in tubes flash frozen in liquid nitrogen and stored at -20 °C. The protein concentration was determined with the Roti-Nanoquant kit with an albumin standard (0–100 µg/ml).

Purification of ulvan

Ulva sp. was collected near Roscoff (France), Lubmin (Germany), or Helgoland (Germany) and dried. Alternatively, dried *Ulva* biomass from the Atlantic coast in Spain was purchased as organic sea lettuce (Kulau). Ulvan was extracted

according to the literature (42). The dialysis step was exchanged by precipitation with acetone (80% (v/v) final concentration). After washing, acetone-precipitated ulvan was dissolved in deionized water and freeze-dried.

Fluorophore-assisted carbohydrate electrophoresis

Fluorophore-assisted carbohydrate electrophoresis (FACE) was performed with 2-aminoacridone (AMAC) as fluorophore as shown previously (43).

Carbohydrate polyacrylamide gel electrophoresis

For carbohydrate polyacrylamide gel electrophoresis (C-PAGE), samples were mixed with an equal volume of FACE loading buffer (42). Gels and running conditions were identical to FACE. Carbohydrates were visualized by staining with Stains-All solution (0.25 g l⁻¹ in 1.7 mM Tris-HCl pH 7.5 + 25% (v/v) isopropanol). The gels were destained with 35% (v/v) isopropanol in deionized water.

Purification of oligomers and structure determination

Ulvan (300 mg) was digested with purified enzymes (100 µg/ml) in 35 mM Tris-HCl buffer (pH 8.0 + 50 mM NaCl) at room temperature overnight. Oligomers were separated on two XK 26/100 (GE Healthcare) in row filled with Bio-Gel P-2 (Rio-Rad) using 50 mM (NH₄)₂CO₃ as mobile phase at a flow rate of 1 ml min⁻¹. After lyophilization of the fractions containing the products, oligomers were dissolved in D₂O and lyophilized two times before NMR spectra were recorded on a Bruker Avance III HD 600 (600 MHz) spectrometer (Bruker) in D₂O solutions. The oligomers—with the confirmed NMR structures—served as standards for activity screening of PUL H enzymes and fractioning after size-exclusion chromatography. The degradation was performed stepwise. The structures were independently elucidated based on 1D and 2D (COSY, HSQC, HMBC, TOCSY) methods, and the assigned ¹H and ¹³C-NMR signals were then compared with literature data, showing excellent consistency (18, 22, 44). For samples containing uronic acid structures, it was required to neutralize the otherwise acidic NMR samples with Na₂HPO₄ to pH 7 to 8 (pH-electrode calibrated to H⁺) in order to achieve fully resolved signals for the carboxylic acid and neighboring positions (¹³C). HPLC-ELS-MS analysis was performed by injection of ~0.1% solutions (1–5 µl) on a Nexera UHPLC system from Shimadzu (equipped with two binary LC-30AD pumps plus degassers, a CTO-20 column oven) and an LCMS-2200 EV MS-detector and an additional ELS-detector (JASCO ELS-2041). Analysis was performed with mobile phase A = H₂O (0.1% HCOOH) and mobile phase B = CH₃CN on a C₁₈ column (XSelect CSH XP C18 2.5 µm 3 × 50 mm) at 40 °C. Flow rate was 1.3 ml min⁻¹ (0–3 min) with 5% B from 0 to 0.15 min, 5% to 98% B from 0.15 to 2.2 min and 98% to 5% B from 2.2 to 2.5 min.

Enzyme assays

Generally, reactions were performed in 35 mM Tris-buffer (pH 8.0 + 50 mM NaCl) ensuring addition of sufficient amounts of the respective enzyme. The ulvan degradation products and the conversion of purified oligomers were analyzed by FACE. Untreated ulvan was generally at a concentration of 1 g l⁻¹ while purified oligomers were used at 0.25 mg ml⁻¹. Incubation was performed overnight at room temperature.

For ulvan lyase activity detection, the respective purified enzymes (15 µg/ml, see also Table S5) were added to an ulvan solution of 1 g l⁻¹ in Tris-buffer (35 mM, pH 8.0, 50 mM NaCl), and the increase of absorbance at 235 nm was measured over time. A sample of the breakdown products was analyzed with the MBTH-assay (45) adapted for reduced volumina (200 µl) and C-PAGE. For the detection of 5-dehydro-4-deoxy-D-glucuronate after reaction with P29_PDnc variants with N-terminal or C-terminal His-tag and/or P33_GH105, the thiobarbituric acid assay (23) adapted for reduced volumina has been used. In total, 37.5 µl of the reaction mixture was mixed with an equal volume of 2% (w/v) sodium acetate in 0.5 N HCl, followed by the addition of 150 µl 0.3% (w/v) thiobarbituric acid in distilled H₂O.

Data availability

All data of this study are contained within the article.

Supporting information—This article contains [supporting information](#) (10, 18).

Acknowledgments—We thank Gurvan Michel for the provision of the plasmid for P36_GH78 and Jens Harder for providing *Ulva* sp. biomass from Helgoland. We thank Daniel Schultz for helping with the MS measurement.

Author contributions—T. S., J.-H. H., and U. T. B. conceptualization; D. B., Christoph Suster, L. R., Christian Stanetty, and U. T. B. formal analysis; T. S., J.-H. H., and U. T. B. funding acquisition; M. B., T. D., Christoph Suster, L. R., N. G., Christian Stanetty, J.-H. H., and U. T. B. investigation; M. B., T. D., Christoph Suster, N. G., Christian Stanetty, and J.-H. H. methodology; D. B. software; M. D. M., J.-H. H., and U. T. B. supervision; J.-H. H. validation; D. B. visualization; M. B. and T. D. writing—original draft; M. B., T. D., D. B., Christoph Suster, L. R., N. G., Christian Stanetty, M. D. M., T. S., J.-H. H., and U. T. B. writing—review and editing.

Funding and additional information—This study was supported by the German Research Foundation (DFG) for the Research Unit FOR2406 “Proteogenomics of Marine Polysaccharide Utilization” (POMPU) by grants of U. T. B. (BO 1862/17-2), J.-H. H. (HE 7217/2-2), and T. S. (SCHW 595/10-2). J.-H. H. acknowledges funding by the Emmy-Noether-Program of the DFG, grant number HE 7217/1-1.

Conflict of interest—The authors declare no conflict of interest.

Abbreviations—The abbreviations used are: C-PAGE, carbohydrate electrophoresis; CAZyme, carbohydrate active enzyme; DHy,

New carbohydrate dehydratase in ulvan degradation

Dehydratase; FACE, fluorophore-assisted carbohydrate electrophoresis; GH, glycoside hydrolase; PL, polysaccharide lyase; PUL, polysaccharide utilization loci.

References

- Field, C. B. (1998) Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* **281**, 237–240
- Lahaye, M., and Robic, A. (2007) Structure and functional properties of ulvan, a polysaccharide from green seaweeds. *Biomacromolecules* **8**, 1765–1774
- Wargacki, A. J., Leonard, E., Win, M. N., Regitsky, D. D., Santos, C. N. S., Kim, P. B., Cooper, S. R., Raisner, R. M., Herman, A., Sivitz, A. B., Lakshmanaswamy, A., Kashiwayama, Y., Baker, D., and Yoshikuni, Y. (2012) An engineered microbial platform for direct biofuel production from brown macroalgae. *Science* **335**, 308–313
- Leiro, J. M., Castro, R., Arranz, J. A., and Lamas, J. (2007) Immunomodulating activities of acidic sulphated polysaccharides obtained from the seaweed *Ulva rigida* C. Agardh. *Int. Immunopharmacol.* **7**, 879–888
- Kim, H. T., Lee, S., Kim, K. H., and Choi, I.-G. (2012) The complete enzymatic saccharification of agarose and its application to simultaneous saccharification and fermentation of agarose for ethanol production. *Biores. Technol.* **107**, 301–306
- Hehemann, J.-H., Boraston, A. B., and Czjzek, M. (2014) A sweet new wave: Structures and mechanisms of enzymes that digest polysaccharides from marine algae. *Curr. Opin. Struct. Biol.* **28**, 77–86
- Song, M., Duc Pham, H., Seon, J., and Chul Woo, H. (2015) Marine brown algae: A conundrum answer for sustainable biofuels production. *Renew. Sustain. Energy Rev.* **50**, 782–792
- Cesário, M. T., da Fonseca, M. M. R., Marques, M. M., and de Almeida, M. C. M. D. (2018) Marine algal carbohydrates as carbon sources for the production of biochemicals and biomaterials. *Biotechnol. Adv.* **36**, 798–817
- Kidgell, J. T., Magnusson, M., de Nys, R., and Glasson, C. R. K. (2019) Ulvan: A systematic review of extraction, composition and function. *Algal Res.* **39**, 101422
- Reisky, L., Préchoux, A., Zühlke, M.-K., Bäumgen, M., Robb, C. S., Gerlach, N., Roret, T., Stanetty, C., Larocque, R., Michel, G., Song, T., Markert, S., Unfried, F., Mihovilovic, M. D., Trautwein-Schult, A., et al. (2019) A marine bacterial enzymatic cascade degrades the algal polysaccharide ulvan. *Nat. Chem. Biol.* **15**, 803–812
- Terrapon, N., Lombard, V., Gilbert, H. J., and Henrissat, B. (2015) Automatic prediction of polysaccharide utilization loci in Bacteroidetes species. *Bioinformatics* **31**, 647–655
- Mann, A. J., Hahnke, R. L., Huang, S., Werner, J., Xing, P., Barbeyron, T., Huettel, B., Stüber, K., Reinhardt, R., Harder, J., Glöckner, F. O., Amann, R. L., and Teeling, H. (2013) The genome of the algae-associated marine flavobacterium *Formosa agariphila* KMM 3901T reveals a broad potential for the degradation of algal polysaccharides. *Appl. Environ. Microbiol.* **79**, 6813–6822
- Ulaganathan, T., Shi, R., Yao, D., Gu, R.-X., Garron, M.-L., Cherney, M., Tieleman, D. P., Sterner, E., Li, G., Li, L., Linhardt, R. J., and Cygler, M. (2017) Conformational flexibility of PL12 family heparinases: Structure and substrate specificity of heparinase III from *Bacteroides thetaiotaomicron* (BT4657). *Glycobiology* **27**, 176–187
- Ulaganathan, T., Helbert, W., Kopel, M., Banin, E., and Cygler, M. (2018) Structure–function analyses of a PL24 family ulvan lyase reveal key features and suggest its catalytic mechanism. *J. Biol. Chem.* **293**, 4026–4036
- Ulaganathan, T., Banin, E., Helbert, W., and Cygler, M. (2018) Structural and functional characterization of PL28 family ulvan lyase NLR48 from *Nonlabens ulvanivorans*. *J. Biol. Chem.* **293**, 11564–11573
- Melcher, R. L. J., Neumann, M., Fuenzalida Werner, J. P., Gröhn, F., and Moerschbacher, B. M. (2017) Revised domain structure of ulvan lyase and characterization of the first ulvan binding domain. *Sci. Rep.* **7**, 44115
- Qin, H.-M., Xu, P., Guo, Q., Cheng, X., Gao, D., Sun, D., Zhu, Z., and Lu, F. (2018) Biochemical characterization of a novel ulvan lyase from *Pseudoalteromonas* sp. strain PLSV. *RSC Adv.* **8**, 2610–2615
- Reisky, L., Stanetty, C., Mihovilovic, M. D., Schweder, T., Hehemann, J.-H., and Bornscheuer, U. T. (2018) Biochemical characterization of an ulvan lyase from the marine flavobacterium *Formosa agariphila*. *Appl. Microbiol. Biotechnol.* **102**, 6987–6996
- Foran, E., Buravenkov, V., Kopel, M., Mizrahi, N., Shoshani, S., Helbert, W., and Banin, E. (2017) Functional characterization of a novel “ulvan utilization loci” found in *Alteromonas* sp. LOR genome. *Algal Res.* **25**, 39–46
- Collén, P. N., Jeudy, A., Sassi, J.-F., Groisillier, A., Czjzek, M., Coutinho, P. M., and Helbert, W. (2014) A novel unsaturated β -glucuronyl hydrolase involved in ulvan degradation unveils the versatility of stereochemistry requirements in family GH105. *J. Biol. Chem.* **289**, 6199–6211
- Salinas, A., and French, C. E. (2017) The enzymatic ulvan depolymerisation system from the alga-associated marine flavobacterium *Formosa agariphila*. *Algal Res.* **27**, 335–344
- Lahaye, M., Brunel, M., and Bonnin, E. (1997) Fine chemical structure analysis of oligosaccharides produced by an ulvan-lyase degradation of the water-soluble cell-wall polysaccharides from *Ulva* sp. (Ulvales, Chlorophyta). *Carbohydr. Res.* **304**, 325–333
- Itoh, T., Ochiai, A., Mikami, B., Hashimoto, W., and Murata, K. (2006) A novel glycoside hydrolase family 105: The structure of family 105 unsaturated rhamnogalacturonyl hydrolase complexed with a disaccharide in comparison with family 88 enzyme complexed with the disaccharide. *J. Mol. Biol.* **360**, 573–585
- Jongkees, S. A. K., and Withers, S. G. (2011) Glycoside cleavage by a new mechanism in unsaturated glucuronyl hydrolases. *J. Am. Chem. Soc.* **133**, 19334–19337
- Konasani, V. R., Jin, C., Karlsson, N. G., and Albers, E. (2018) A novel ulvan lyase family with broad-spectrum activity from the ulvan utilisation loci of *Formosa agariphila* KMM 3901. *Sci. Rep.* **8**, 14713
- Somoza, J. R., Menon, S., Schmidt, H., Joseph-McCarthy, D., Dessen, A., Stahl, M. L., Somers, W. S., and Sullivan, F. X. (2000) Structural and kinetic analysis of *Escherichia coli* GDP-mannose 4,6 dehydratase provides insights into the enzyme’s catalytic mechanism and regulation by GDP-fucose. *Structure* **8**, 123–135
- Allard, S. T. M., Beis, K., Giraud, M.-F., Hegeman, A. D., Gross, J. W., Wilmoth, R. C., Whitfield, C., Graninger, M., Messner, P., Allen, A. G., Maskell, D. J., and Naismith, J. H. (2002) Toward a structural understanding of the dehydratase mechanism. *Structure* **10**, 81–92
- Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z., Busk, P. K., Xu, Y., and Yin, Y. (2018) dbCAN2: A meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **46**, W95–W101
- Eddy, S. R. (2011) Accelerated profile HMM Searches. *PLoS Comput. Biol.* **7**, e1002195
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402
- Barbeyron, T., Brillet-Guéguen, L., Carré, W., Carrière, C., Caron, C., Czjzek, M., Hoebeke, M., and Michel, G. (2016) Matching the diversity of sulfated biomolecules: Creation of a classification database for sulfatases reflecting their substrate specificity. *PLoS One* **11**, e0164846
- Krzywinski, M., Schein, J., Biro, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S. J., and Marra, M. A. (2009) Circos: An information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645
- Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., and Finn, R. D. (2018) HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200–W204
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680

35. Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321
36. Lefort, V., Longueville, J.-E., and Gascuel, O. (2017) SMS: Smart model selection in PhyML. *Mol. Biol. Evol.* **34**, 2422–2424
37. Letunic, I., and Bork, P. (2007) Interactive tree of life (iTOL): An online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127–128
38. Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016) ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350
39. Berezin, C., Glaser, F., Rosenberg, J., Paz, I., Pupko, T., Fariselli, P., Casadio, R., and Ben-Tal, N. (2004) ConSeq: The identification of functionally and structurally important residues in protein sequences. *Bioinformatics* **20**, 1322–1324
40. Li, C., Wen, A., Shen, B., Lu, J., Huang, Y., and Chang, Y. (2011) FastCloning: A highly simplified, purification-free, sequence- and ligation-independent PCR cloning method. *BMC Biotechnol.* **11**, 92
41. Ladner, C. L., Yang, J., Turner, R. J., and Edwards, R. A. (2004) Visible fluorescent detection of proteins in polyacrylamide gels without staining. *Anal. Biochem.* **326**, 13–20
42. Robic, A., Gaillard, C., Sassi, J.-F., Lerat, Y., and Lahaye, M. (2009) Ultrastructure of ulvan: A polysaccharide from green seaweeds. *Biopolymers* **91**, 652–664
43. Hehemann, J.-H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., and Michel, G. (2010) Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* **464**, 908–912
44. Lahaye, M. (1998) NMR spectroscopic characterisation of oligosaccharides from two *Ulva rigida* ulvan samples (Ulvales, Chlorophyta) degraded by a lyase. *Carbohydr. Res.* **314**, 1–12
45. Horn, S., and Eijssink, V. G. H. (2004) A reliable reducing end assay for chito-oligosaccharides. *Carbohydr. Polym.* **56**, 35–39

A new carbohydrate-active oligosaccharide dehydratase is involved in the degradation of ulvan

Marcus Bäumgen^{1,‡}, Theresa Dutschei^{1,‡}, Daniel Bartosik², Christoph Suster³, Lukas Reisky¹, Nadine Gerlach^{4,5}, Christian Stanetty³, Marko D. Mihovilovic³, Thomas Schweder², Jan-Hendrik Hehemann^{4,5}, and Uwe T. Bornscheuer^{1,*}

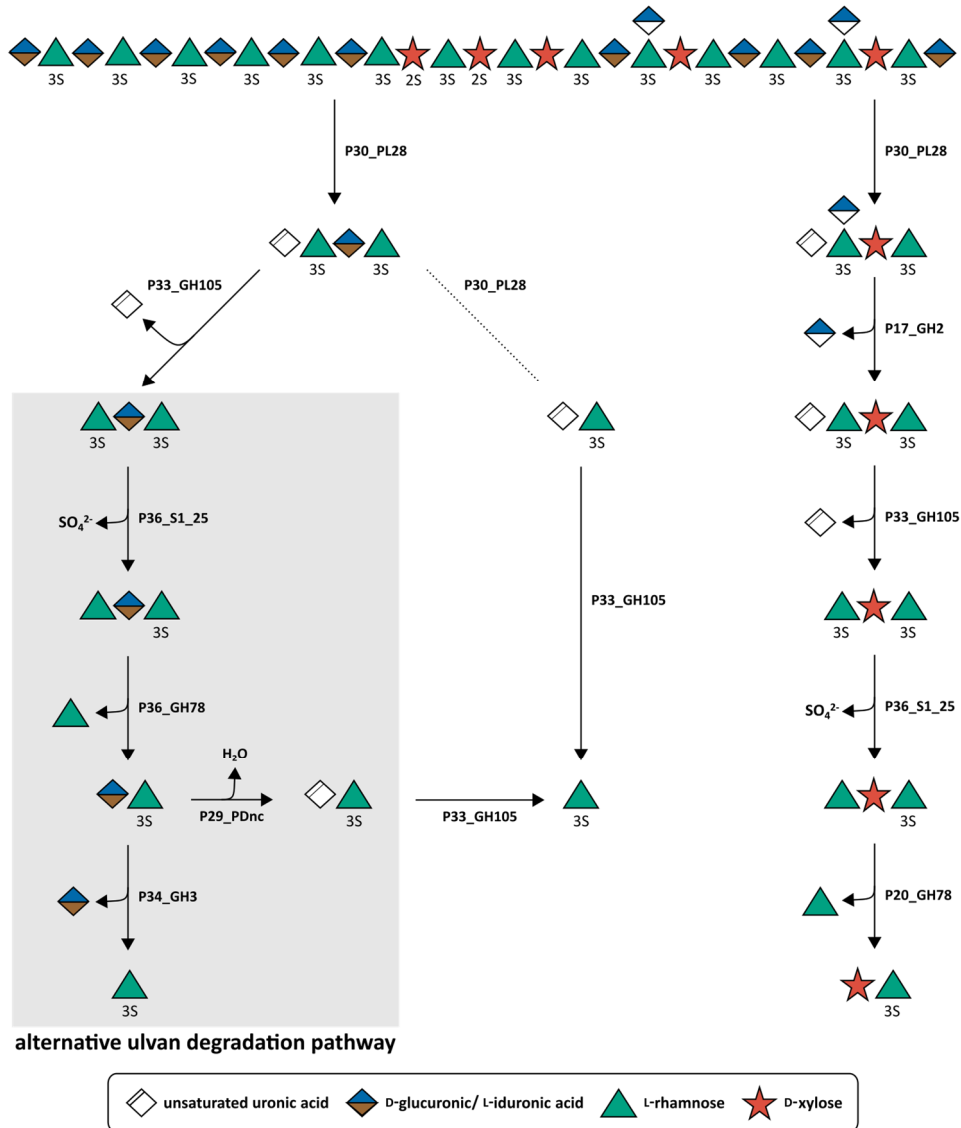
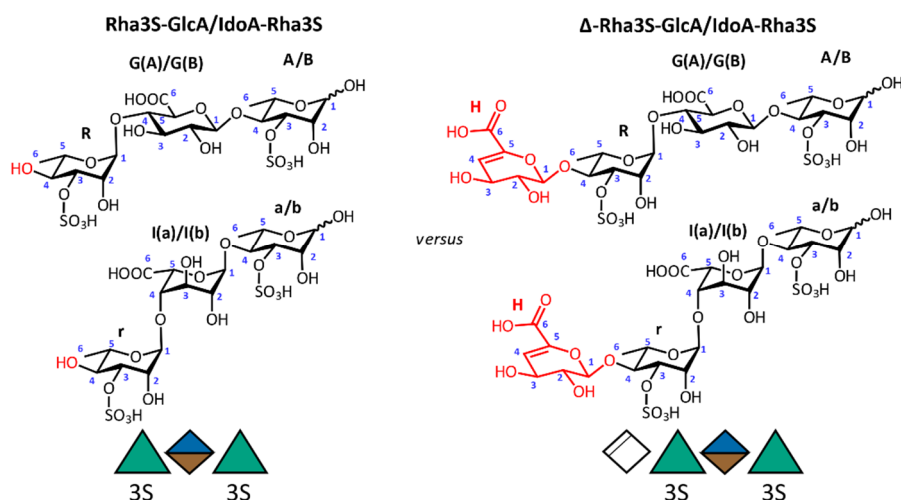


Figure S1. Overview of the degradation of uronic acid containing ulvan oligosaccharides from *Formosa agariphila* KM3901T by PUL H enzymes. The left side summarizes the novel found activities which were presented in this study, which complements the pathway described by Reisky *et al.* 2019 (10). This alternative pathway fills the gap in the degradation of the ulvan oligosaccharides as the previous pathway lacked the degradation of small uronic acid containing oligosaccharide as the ulvan lyase P30_PL28 cannot be active due to product inhibition (dotted line) (18).

Table S1a. NMR shifts of GlcA-trimer contained in a mixture of Rha3S-GlcA-Rha3S and Rha3S-IdoA-Rha3S (ratio ~7:3) originating from the digestion of starting tetramer (Δ -Rha3S-GlcA-Rha3S) with Enzyme P33_GH105. The chemical shifts are compared to the starting material (10,18). Indicative shifts of chemical shifts that support the claimed structure are marked in red. NMR spectra are shown in Figs. S2 to S4. ESI-MS measurements support the structure (Fig. S6 and Fig. S7). The shifts of the IdoA-component are shown in Table S1b.



	¹ H-Shifts (ppm)			¹³ C Shifts (ppm)			
	Rha3S-GlcA Rha3S	Ref Δ -Rha3S- GlcA-Rha3S (10)	Delta (A1=A1)	Rha3S-GlcA Rha3S	Ref Δ -Rha3S-GlcA- Rha3S (10)	Delta (A1=A1)	
A α -Rhamnose	1	5.08	5.08	0.00	96.1	94.62	0.00
	2	4.23	4.21	0.02	71.98	70.50	0.00
	3	4.61	4.61	0.00	80.73	79.25	0.00
	4	3.76	3.77	-0.01	81.02	79.48	0.06
	5	3.98	3.98	0.00	70.3	68.82	0.00
	6	1.31	1.30	0.01	19.89	18.41	0.00
B β -Rhamnose	1	4.89	4.89	0.00	95.67	94.19	0.00
	2	4.23	4.23	0.00	72.41	70.93	0.00
	3	4.43	4.42	0.01	82.68	81.20	0.00
	4	3.69	3.68	0.01	80.58	79.03	0.07
	5	3.53	3.52	0.01	73.54	72.05	0.01
	6	1.32	1.31	0.01	19.89	18.41	0.00
G(A) Glucuronic acid (of α -anomer)	1	4.65	4.64	0.01	105.85	104.35	0.02
	2	3.33	3.31	0.02	76.44	74.84	0.12
	3	3.65	3.63	0.02	76.44	74.96	0.00
	4	3.58	3.56	0.02	81.51	80.11	-0.08
	5	3.85	3.84	0.01	77.8	76.32	0.00
	6				177.13	175.57	0.08
G(B) Glucuronic acid (of β -anomer)	1	4.65	4.64	0.01	105.87	104.32	0.07
	2	3.32	3.30	0.02	76.44	74.88	0.08
	3	3.65	3.63	0.02	76.44	74.96	0.00
	4	3.58	3.55	0.03	81.51	80.11	-0.08
	5	3.85	3.84	0.01	77.8	76.32	0.00
	6				177.13	175.57	0.08
R Rhamnoside	1	4.73	4.72	0.01	102.97	101.34	0.15
	2	4.23	4.19	0.04	71.42	70.1	-0.16
	3	4.42	4.58	-0.16	81.17	79.93	-0.24
	4	3.53	3.77	-0.24	72.52	77.67	-6.63
	5	4.11	4.12	-0.01	71.63	68.6	1.55
	6	1.25	1.11	0.14	19.23	17.63	0.12

Table S1b. NMR shifts of IdoA-trimer contained in a mixture of Rha3S-GlcA-Rha3S and Rha3S-IdoA-Rha3S (ratio ~7:3) originating from the digestion of starting tetramer (Δ -Rha3S-IdoA-Rha3S) with Enzyme P33_GH105. The chemical shifts are compared to the starting material². Indicative shifts of chemical shifts that support the claimed structure are marked in red. NMR spectra are shown in Figs. S2 to S5. ESI-MS measurements support the structure shown in Fig. S6 and S7). The shifts of the GlcA-component are shown in Table S1a. Due to the low concentration of the IdoA component, some signals were not assignable with certainty, and are therefore omitted in the Table.

	¹ H-Shifts (ppm)			¹³ C Shifts (ppm)			
	Rha3S-IdoA Rha3S	Ref Δ -Rha3S-IdoA- Rha3S (10)	Delta (A1=A1)	Rha3S-IdoA Rha3S	Ref Δ -Rha3S-IdoA- Rha3S (10)	Delta (A1=A1)	
a α -Rhamnose	1	5.08	5.08	0.00	96.34	94.86	0.00
	2	4.23	4.22	0.01	72.09	70.61	0.00
	3	4.61	4.60	0.01	81.36	79.26	0.62
	4	3.76	3.76	0.00	79.07	77.51	0.08
	5	3.98	3.96	0.02	69.83	68.34	0.01
	6	1.2	1.18	0.02	19.71	18.23	0.00
b β -Rhamnose	1	4.89	4.88	0.01	95.82	94.34	0.00
	2	4.23	4.22	0.01	72.46	70.97	0.01
	3	4.43	4.41	0.02	83.22	81.74	0.00
	4	3.69	3.68	0.01	78.75	77.18	0.09
	5	3.53	3.51	0.02	73.16	71.68	0.00
	6	1.22	1.20	0.02	19.71	18.23	0.00
I(A) Iduronic acid (of α -anomer)	1	5.05	5.01	0.04	105.74	104.19	0.07
	2	3.63	3.59	0.04	73.29	71.79	0.02
	3	3.81	3.76	0.05	74.16	72.58	0.10
	4	3.98	3.96	0.02	82.12	80.61	0.03
	5	*	4.46		73.22	*	
	6				176.06	174.57	0.01
I(B) Iduronic acid (of β -anomer)	1	5.05	5.01	0.04	105.71	104.16	0.07
	2	3.63	3.59	0.04	73.29	71.72	0.09
	3	3.81	3.76	0.05	74.08	72.50	0.10
	4	3.98	3.96	0.02	82.06	80.54	0.04
	5	*	4.46		73.22	*	
	6				176.06	174.57	0.01
r Rhamnoside	1	4.86	4.82	0.04	104.03	102.39	0.16
	2	4.18	4.15	0.03	71.33	70.07	-0.22
	3	4.4	4.56	-0.16	81.17	79.88	-0.19
	4	3.53	3.76	-0.23	72.63	77.62	-6.47
	5	3.98	3.96	0.02	71.77	68.78	1.51
	6	1.27	1.12	0.15	19.29	17.71	0.1
* no unambiguous assignment possible	<i>Referenced to HDO: 4.79 ppm</i>						

¹H NMR (D₂O, 600 MHz)

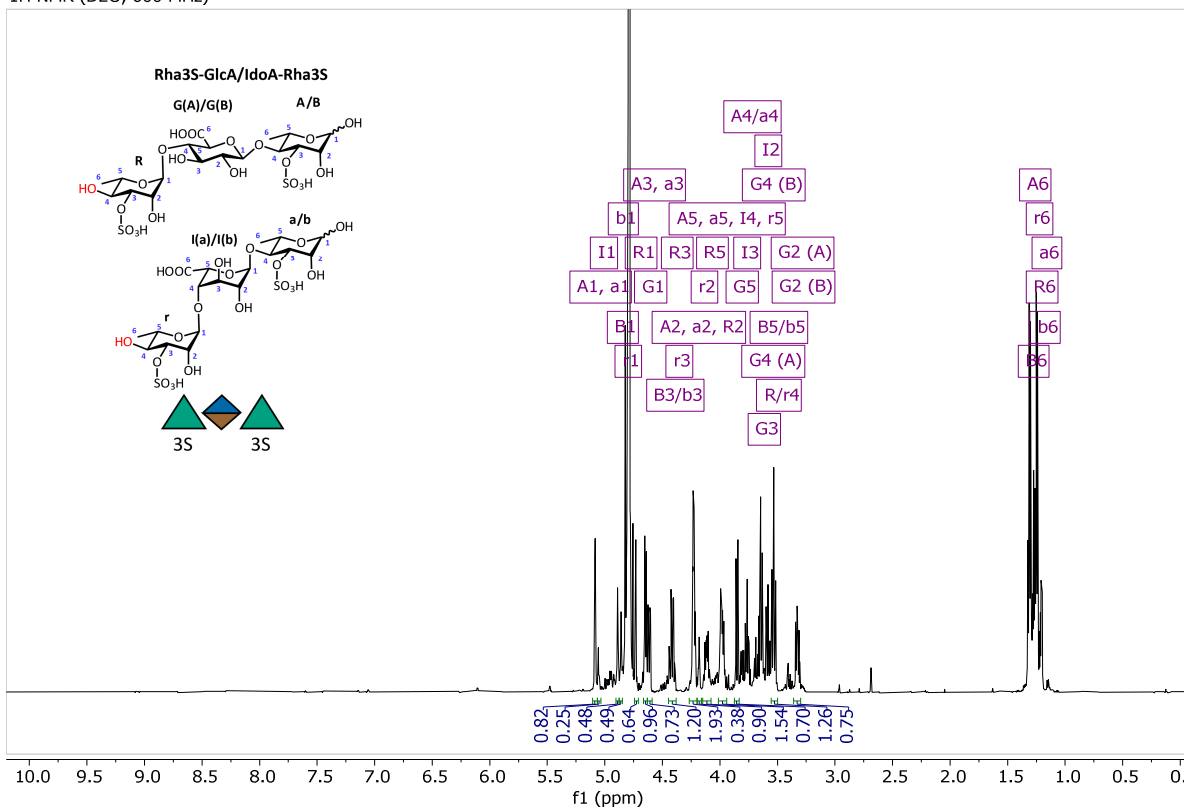


Figure S2. ¹H-NMR of the purified trimer mixture containing Rha3S-GlcA-Rha3S and Rha3S-IdoA-Rha3S (ratio ~7:3) – Full View

¹H NMR (D₂O, 600 MHz)

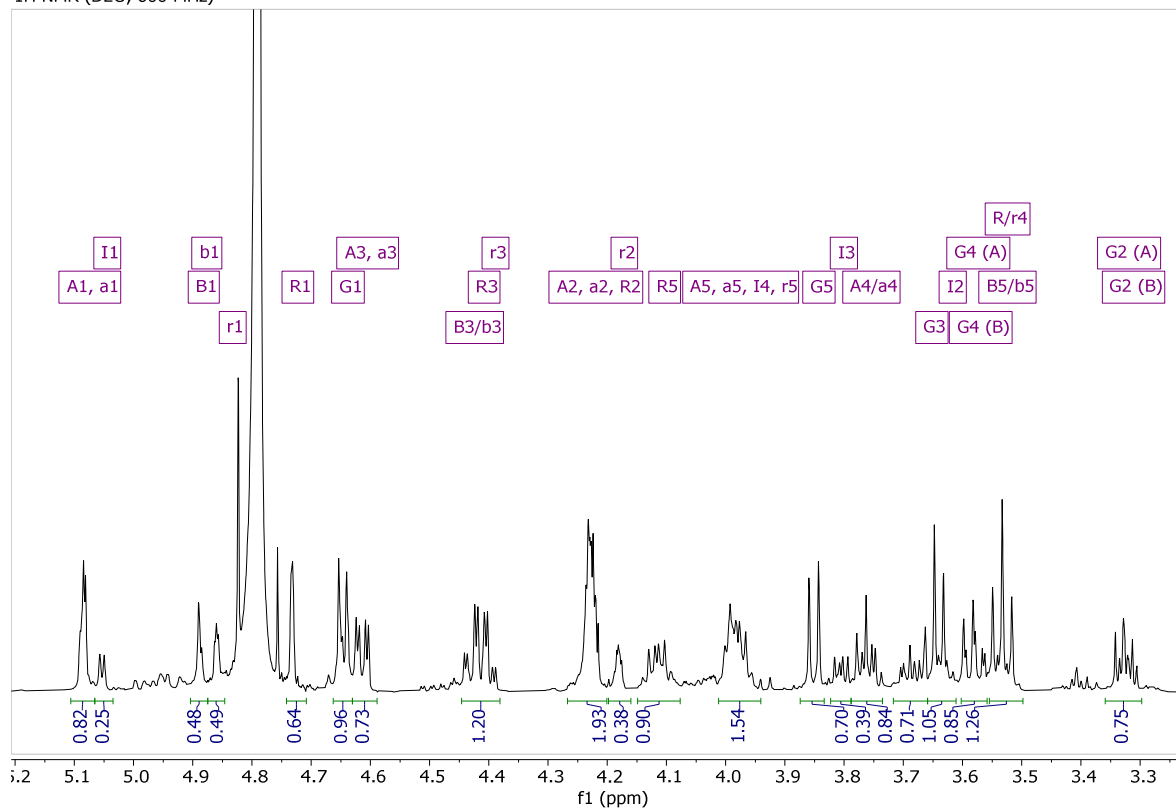


Figure S3. ¹H-NMR of the purified trimer mixture containing Rha3S-GlcA-Rha3S and Rha3S-IdoA-Rha3S (ratio ~7:3) – Zoom into the carbohydrate region

¹³C NMR (D₂O, 150.92 MHz)

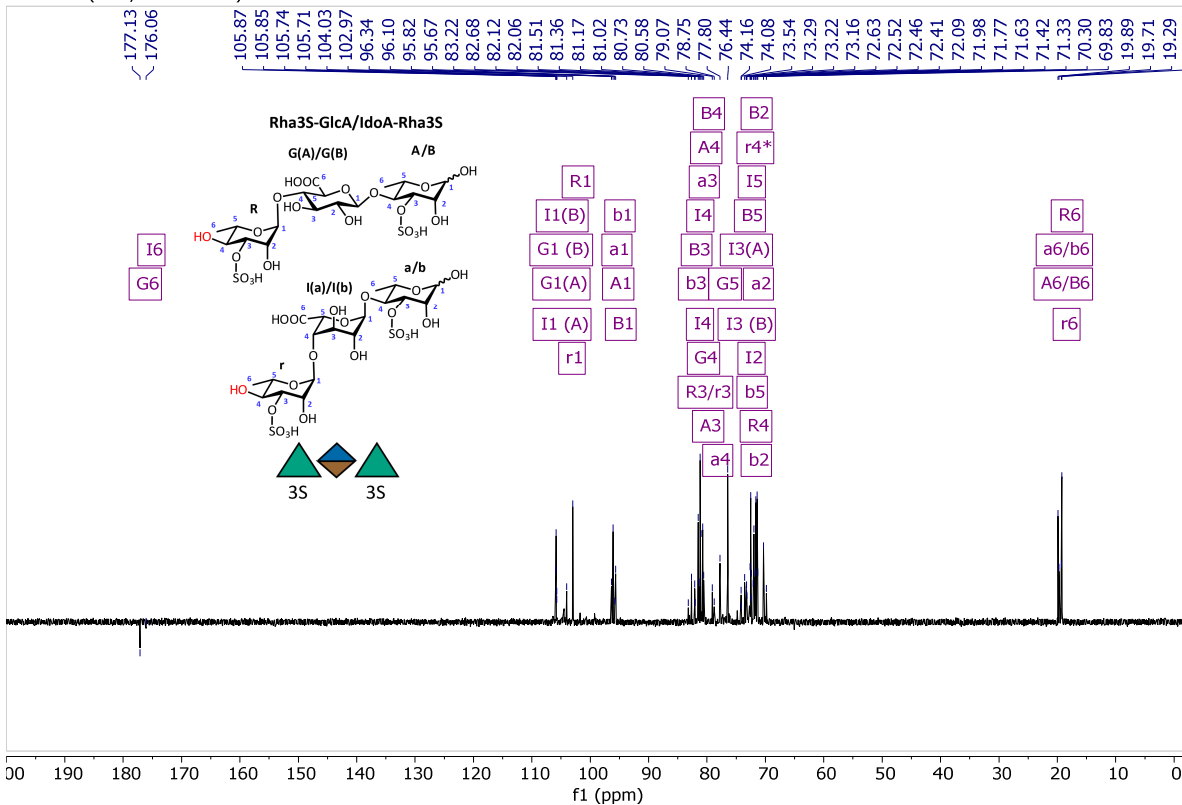


Figure S4. ¹³C-NMR of the purified trimer mixture containing Rha3S-GlcA-Rha3S and Rha3S-IdoA-Rha3S (ratio ~7:3) – FullView

¹³C NMR (D₂O, 150.92 MHz)

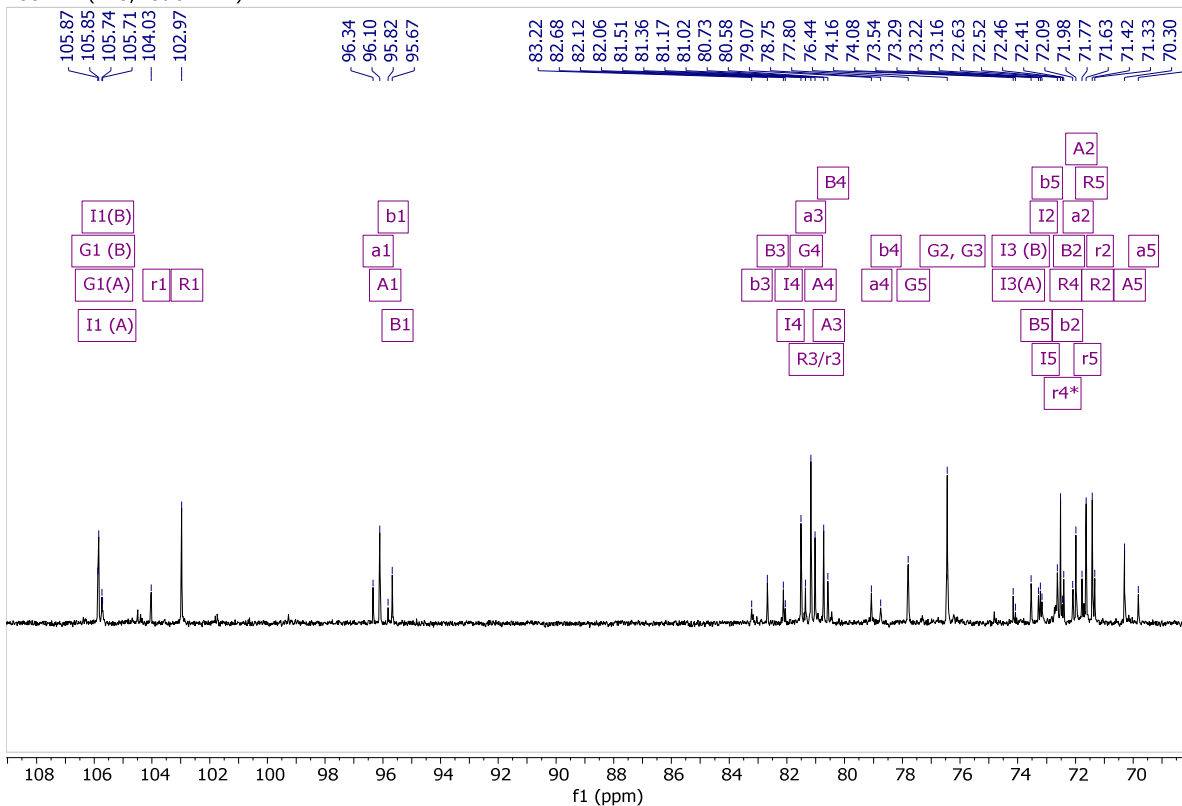


Figure S5. ¹³C-NMR of the purified trimer mixture containing Rha3S-GlcA-Rha3S and Rha3S-IdoA-Rha3S (ratio ~7:3) – Zoom into the carbohydrate region

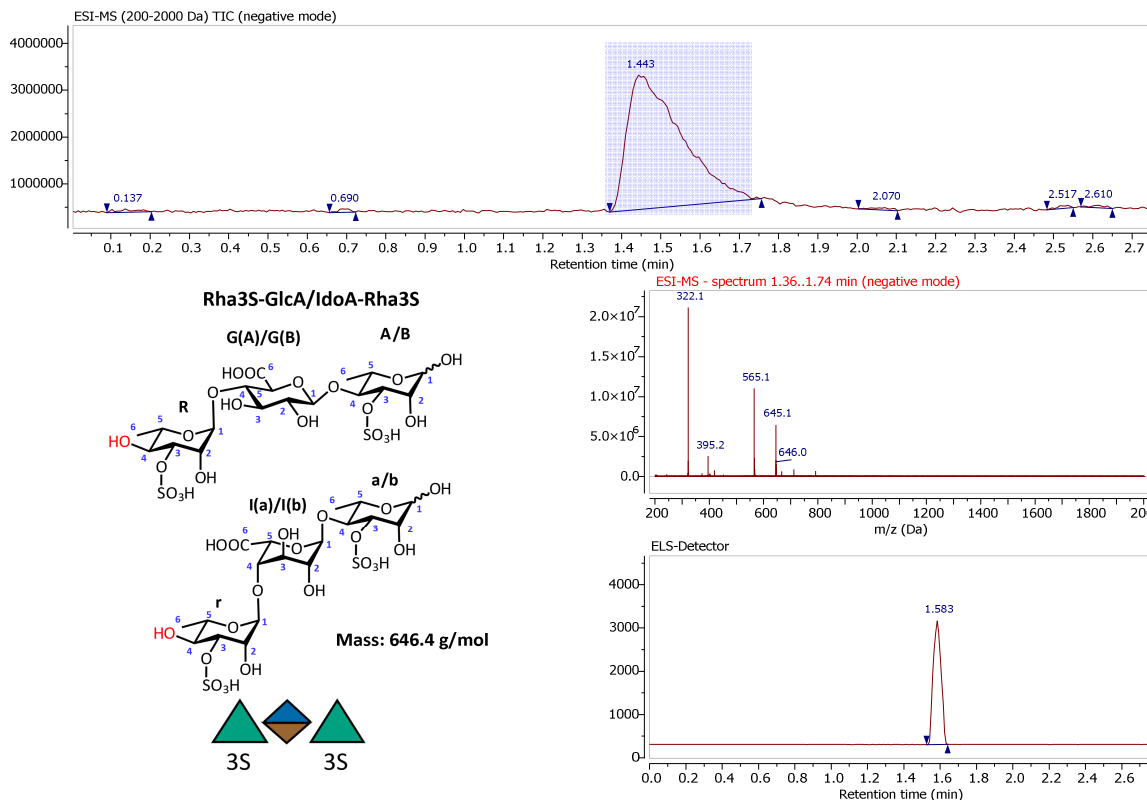


Figure S6. HPLC-MS (ESI-) measurement of the trimer mixture containing Rha3S-GlcA-Rha3S and Rha3S-I doA-Rha3S (ratio ~7:3). Showing the [M-1] molecule peak of the compounds and also the [M-80] signal, which was shown to be only present in oligosaccharide structures containing SO₃ on a non-reducing sugar.

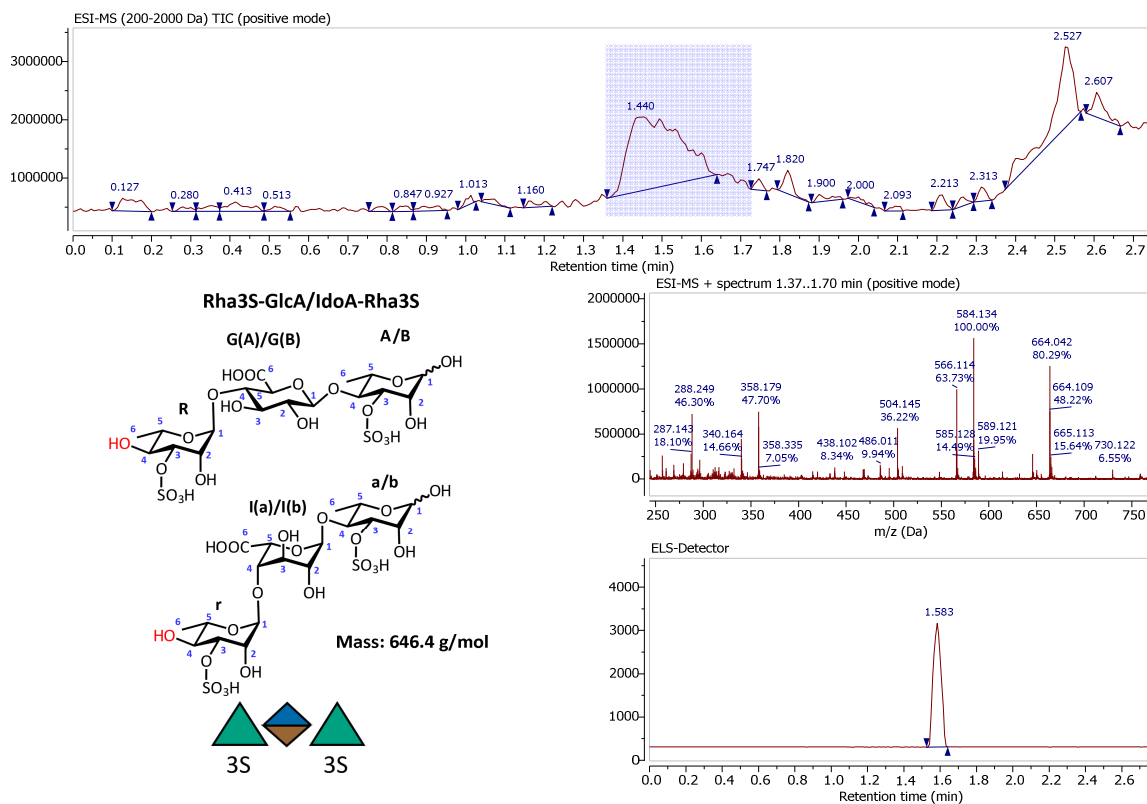
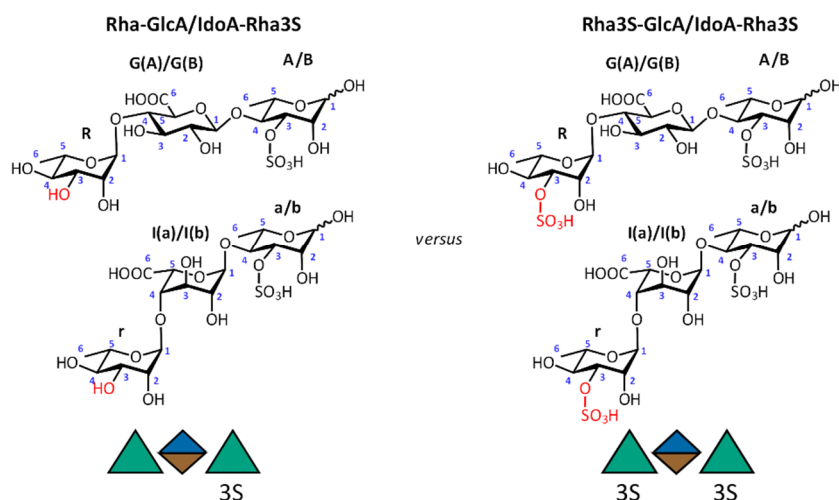


Figure S7. HPLC-MS (ESI+) measurement of the trimer mixture containing Rha3S-GlcA-Rha3S and Rha3S-I doA-Rha3S (ratio ~7:3). Showing the [M+18] molecule peak of the compounds and also the [M-80+18] signal, which was shown to be only present in oligosaccharide structures containing SO₃ on a non-reducing sugar.

Table S2a. NMR shifts of GlcA-trimer contained in a mixture of Rha-GlcA-Rha3S and Rha-IdoA-Rha3S (ratio ~7:3) originating from the digestion of Rha3S-GlcA-Rha3S with Enzyme P36_S1_25. The chemical shifts are compared to the starting material. Indicative shifts of chemical shifts that support the claimed structure are marked in red. NMR spectra are shown in Figs. S8 to S11. ESI-MS measurements support the structure shown in (Figs. S12 and S13). The shifts of the IdoA-component are shown in Table S2b.



	¹ H-Shifts (ppm)				¹³ C Shifts (ppm)		
	Rha-GlcA Rha3S	Ref Rha3S-GlcA- Rha3S	Delta (A1=A1)	Rha-GlcA Rha3S	Ref Rha3S-GlcA- Rha3S	Delta (A1=A1)	
A α-Rhamnose	1	5.08	5.08	0.00	96.12	96.1	0.00
	2	4.22	4.23	-0.01	71.98	71.98	-0.02
	3	4.61	4.61	0.00	80.76	80.73	0.01
	4	3.75	3.76	-0.01	80.88	81.02	-0.16
	5	3.99	3.98	0.01	70.33	70.3	0.01
	6	1.31	1.31	0.00	19.91	19.89	0.00
B β-Rhamnose	1	4.89	4.89	0.00	95.67	95.67	-0.02
	2	4.23	4.23	0.00	72.42	72.41	-0.01
	3	4.43	4.43	0.00	82.7	82.68	0.00
	4	3.68	3.69	-0.01	80.46	80.58	-0.14
	5	3.52	3.53	-0.01	73.57	73.54	0.01
	6	1.32	1.32	0.00	19.91	19.89	0.00
G(A) Glucuronic acid (of α-anomer)	1	4.61	4.65	-0.04	105.83	105.85	-0.04
	2	3.33	3.33	0.00	76.48	76.44	0.02
	3	3.58	3.65	-0.07	76.7	76.44	0.24
	4	3.55	3.58	-0.03	81.86	81.51	0.33
	5	3.7	3.85	-0.15	78.85	77.8	1.03
	6				178.15	177.13	1.00
G(B) Glucuronic acid (of β-anomer)	1	4.61	4.65	-0.04	105.83	105.87	-0.06
	2	3.32	3.32	0.00	76.51	76.44	0.05
	3	3.58	3.65	-0.07	76.7	76.44	0.24
	4	3.53	3.58	-0.05	81.86	81.51	0.33
	5	3.7	3.85	-0.15	78.85	77.8	1.03
	6				178.15	177.13	1.00
R Rhamnoside	1	4.69	4.73	-0.04	103.36	102.97	0.37
	2	3.89	4.23	-0.34	72.97	71.42	1.53
	3	3.74	4.42	-0.68	72.65	81.17	-8.54
	4	3.38	3.53	-0.15	74.6	72.52	2.06
	5	3.99	4.11	-0.12	71.56	71.63	-0.09
	6	1.21	1.25	-0.04	19.1	19.23	-0.15
Referenced to HDO: 4.79 ppm							

Table S2b. NMR shifts of IdoA-trimer contained in a mixture of Rha-GlcA-Rha3S and Rha-IdoA-Rha3S (ratio ~7:3) originating from the digestion of Rha3S-IdoA-Rha3S with enzyme P36_S1_25. The chemical shifts are compared to the starting material. Indicative shifts of chemical shifts that support the claimed structure are marked in red. NMR spectra are shown in Fig. S8 to S11. ESI-MS measurements support the structure shown in (Figs. S12 and S13). The shifts of the IdoA-component are shown in Table S2a. Due to the low concentration of the IdoA component, some signals were not assignable with certainty, and are therefore omitted in the Table.

	¹ H-Shifts (ppm)			¹³ C Shifts (ppm)			
	Rha-IdoA Rha3S	Ref Rha3S-IdoA- Rha3S	Delta (A1=A1)	Rha-IdoA Rha3S	Ref Rha3S-IdoA- Rha3S	Delta (A1=A1)	
a α-Rhamnose	1	5.08	5.08	0.00	96.38	96.34	0.00
	2	4.22	4.23	-0.01	72.1	72.09	-0.03
	3	4.61	4.61	0.00	81.47	81.36	0.07
	4	3.75	3.76	-0.01	*	79.07	
	5	3.99	3.98	0.01	*	69.83	
	6	*	1.2		19.7	19.71	
b β-Rhamnose	1	4.89	4.89	0.00	95.83	95.82	-0.03
	2	4.22	4.23	-0.01	72.6	72.46	0.10
	3	4.43	4.43	0.00	*	83.22	
	4	3.68	3.69	-0.01	78.62	78.75	-0.17
	5	3.52	3.53	-0.01	*	73.16	
	6	*	1.22		19.7	19.71	-0.05
I(A) Iduronic acid (of α-anomer)	1	5.02	5.05	-0.03	105.68	105.74	-0.10
	2	3.64	3.63	0.01	73.19	73.29	-0.14
	3	3.77	3.81	-0.04	*	74.16	
	4	3.99	3.98	0.01	*	82.12	
	5	*			*	73.22	
	6				177.25	176.06	1.15
I(B) Iduronic acid (of β-anomer)	1	5.02	5.05	-0.03	105.68	105.71	-0.07
	2	3.62	3.63	-0.01	73.19	73.29	-0.14
	3	3.77	3.81	-0.04	*	74.08	
	4	3.99	3.98	0.01	*	82.06	
	5	*	*		*	73.22	
	6				177.25	176.06	1.15
r Rhamnoside	1	*	4.86		104.47	104.03	0.40
	2	3.86	4.18	-0.32	72.94	71.33	1.57
	3	3.74	4.4	-0.66	72.62	81.17	-8.59
	4	3.38	3.53	-0.15	74.67	72.63	2.00
	5	3.89	3.98	-0.09	71.67	71.77	-0.14
	6	1.24	1.27	-0.03	19.16/19.33*	19.29	
*... assignment not possible	<i>Referenced to HDO: 4.79 ppm</i>						

¹H NMR (D₂O, 600 MHz)

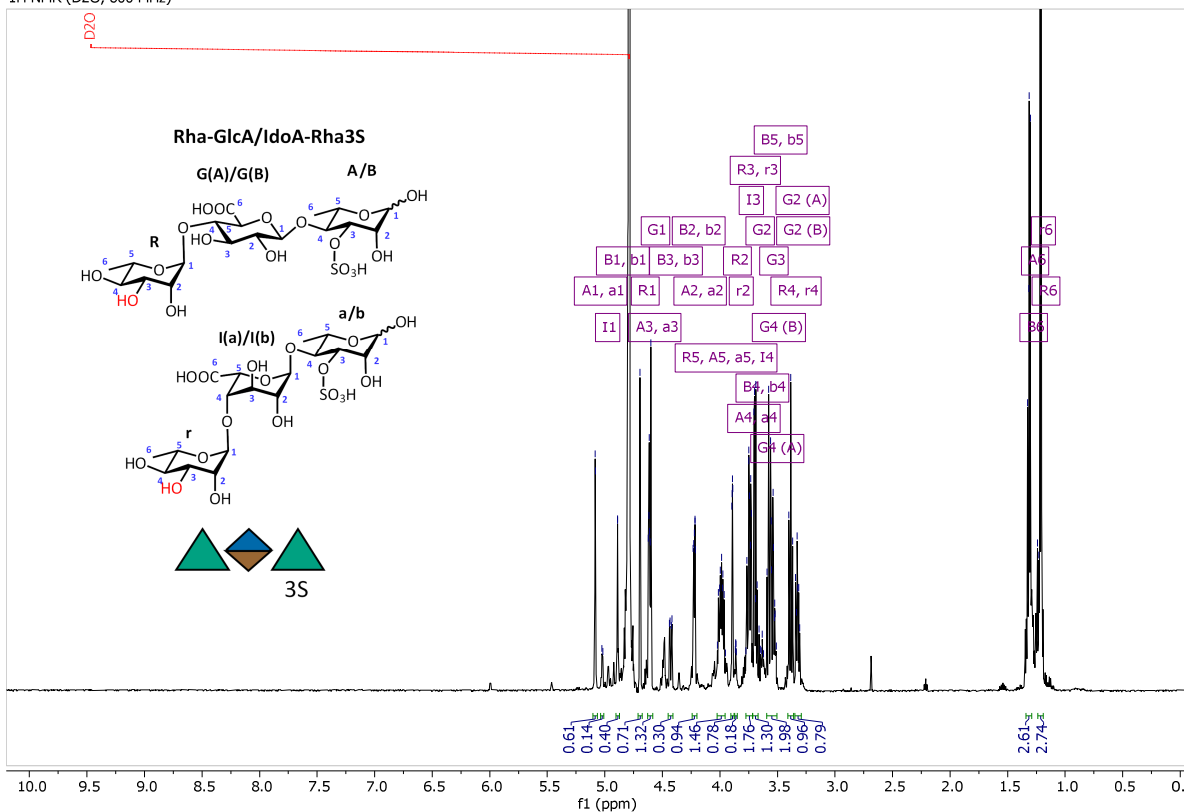


Figure S8. ¹H-NMR of the purified trimer mixture containing Rha-GlcA-Rha3S and Rha-IdoA-Rha3S (ratio ~7:3) – Full View

¹H NMR (D₂O, 600 MHz)

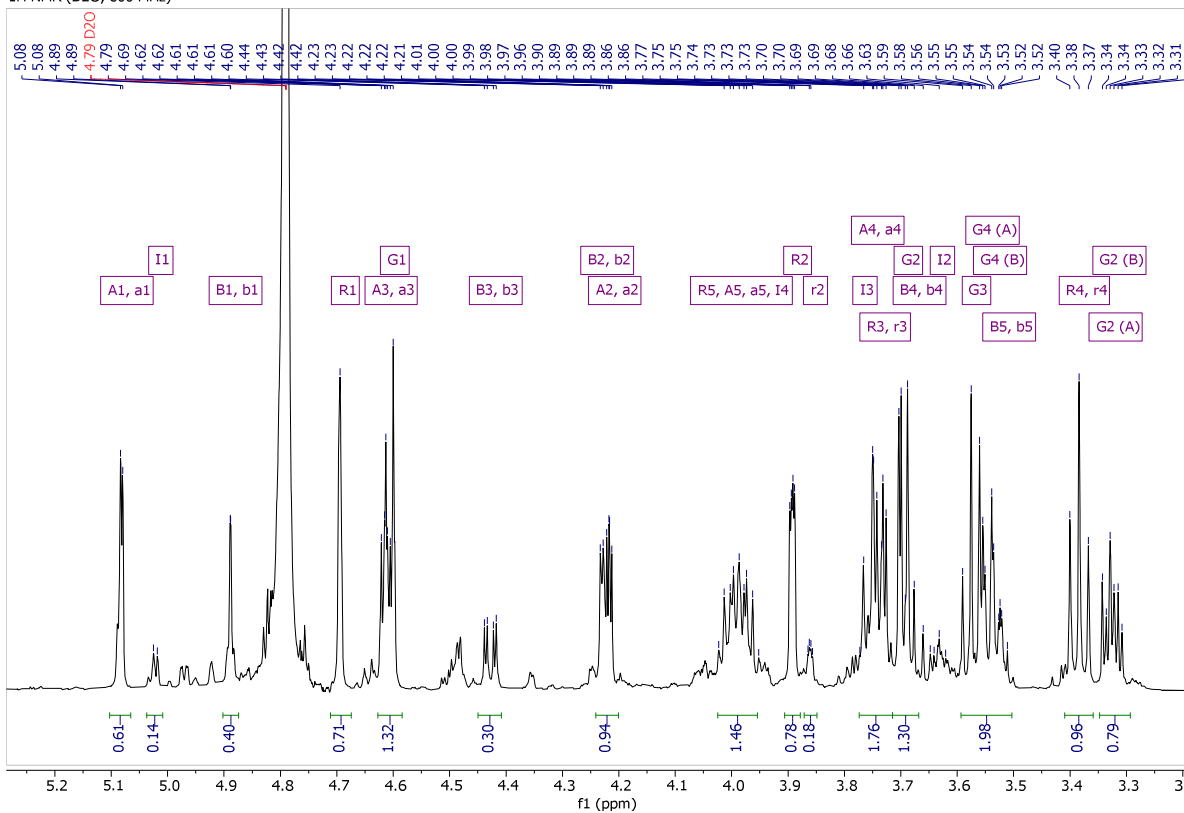


Figure S9. ¹H-NMR of the purified trimer mixture containing Rha-GlcA-Rha3S and Rha-IdoA-Rha3S (ratio ~7:3) – Zoom into the carbohydrate region

¹³C NMR (D₂O, 151 MHz)

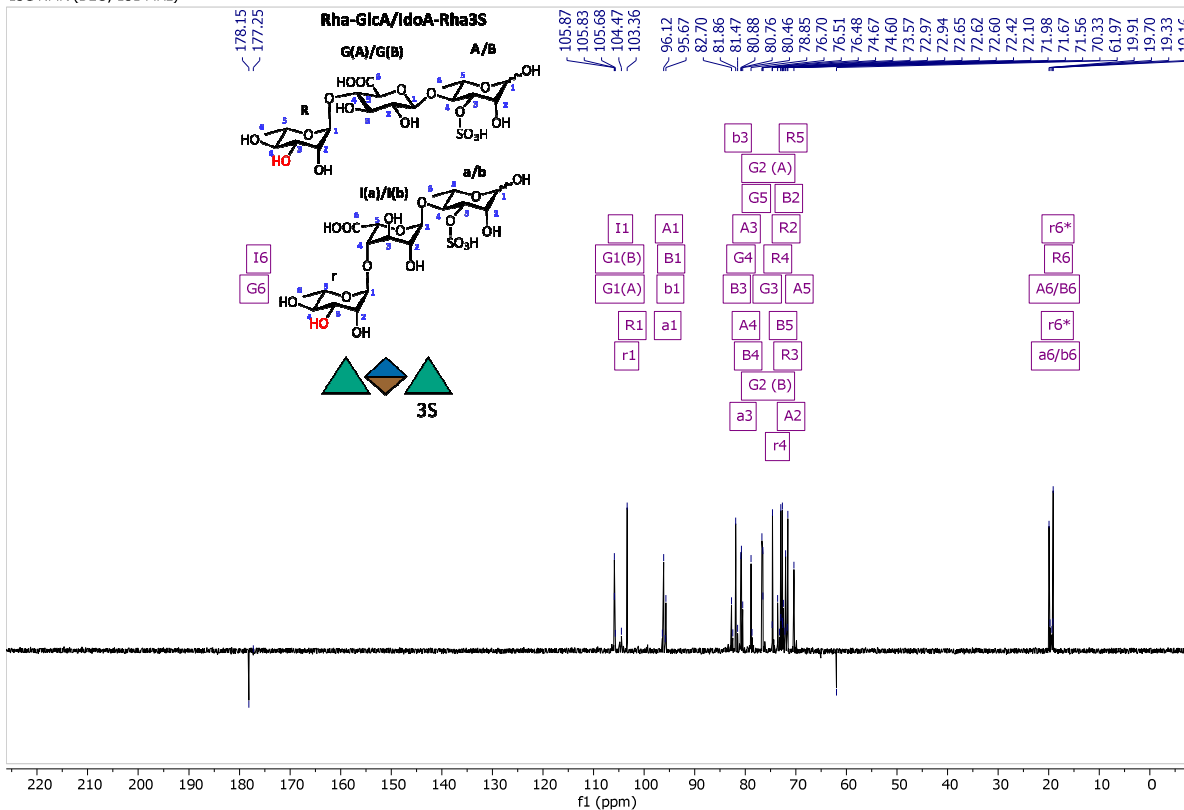


Figure S10. ¹³C-NMR of the purified trimer mixture containing Rha3S-GlcA-Rha3S and Rha3S-IdoA-Rha3S (ratio ~7:3) – FullView

¹³C NMR (D₂O, 151 MHz)

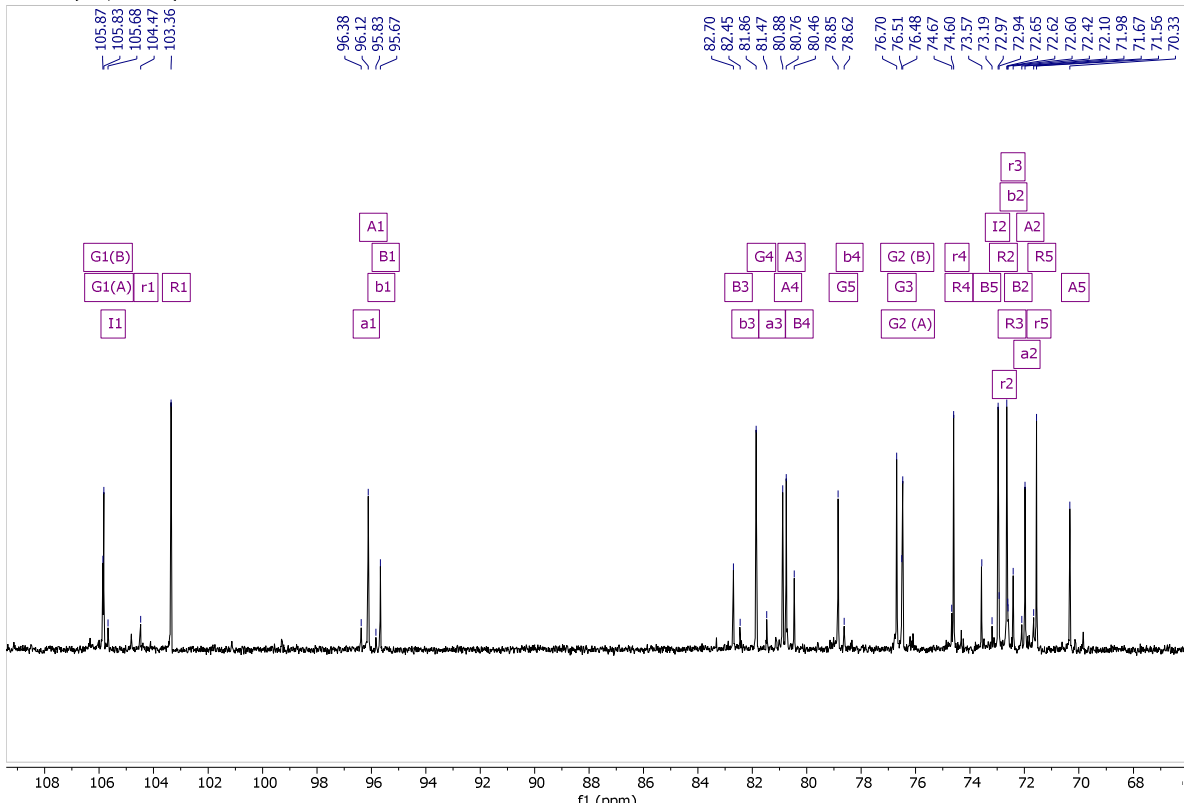
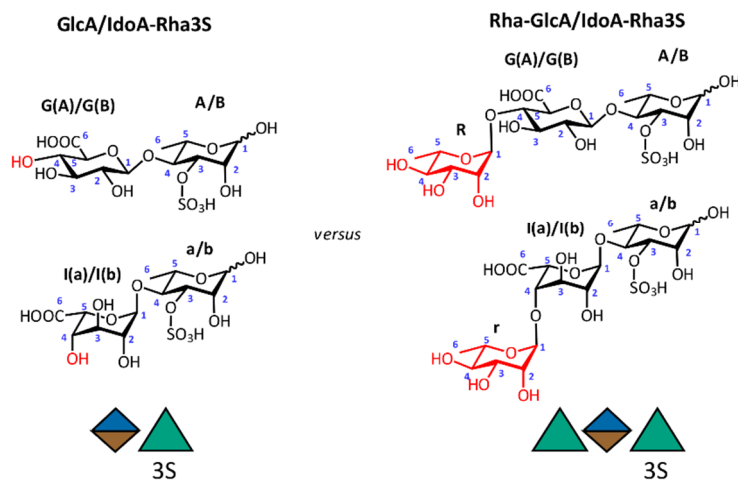


Figure S11. ¹³C-NMR of the purified trimer mixture containing Rha3S-GlcA-Rha3S and Rha3S-IdoA-Rha3S (ratio ~7:3) – Zoom into the carbohydrate region

Table S3a. NMR shifts of GlcA-dimer contained in a mixture of GlcA-Rha3S and IdoA-Rha3S (ratio ~7:3) originating from the digestion of Rha-GlcA-Rha3S with Enzyme P36_GH78. The chemical shifts are compared to the starting material. Indicative shifts of chemical shifts that support the claimed structure are marked in red. NMR spectra are shown in Figs. S14 to S17. ESI-MS measurements support the structure shown in Figs. S18 and S19). The shifts of the IdoA-component are shown in Table S3b.



		¹ H-Shifts (ppm)			¹³ C Shifts (ppm)		
		GlcA Rha3S	Ref Rha-GlcA-Rha3S	Delta (A1=A1)	GlcA Rha3S	Ref Rha-GlcA-Rha3S	Delta (A1=A1)
A α-Rhamnose	1	5.13	5.08	0.00	93.46	96.12	0
	2	4.27	4.22	0.00	69.36	71.98	0.04
	3	4.66	4.61	0.00	78.18	80.76	0.08
	4	3.82	3.75	0.02	78.27	80.88	0.05
	5	4.02	3.99	-0.02	67.76	70.33	0.09
	6	1.37	1.31	0.01	17.28	19.91	0.03
B β-Rhamnose	1	4.93	4.89	-0.01	93.04	95.67	0.03
	2	4.27	4.23	-0.01	69.8	72.42	0.04
	3	4.45	4.43	-0.03	80.14	82.7	0.1
	4	3.75	3.68	0.02	77.85	80.46	0.05
	5	3.55	3.52	-0.02	70.97	73.57	0.06
	6	1.37	1.32	0.00	17.28	19.91	0.03
G(A) Glucuronic acid (of α-anomer)	1	4.66	4.61	0.00	103.25	105.83	0.08
	2	3.34	3.33	-0.04	73.52	76.48	-0.3
	3	3.55	3.58	-0.08	75.44	76.7	1.4
	4	3.55	3.55	-0.05	71.89	81.86	-7.31
	5	3.71	3.7	-0.04	76.63	78.85	0.44
	6			-0.05	175.87	178.15	0.38
G(B) Glucuronic acid (of β-anomer)	1	4.66	4.61	0.00	103.25	105.83	0.08
	2	3.34	3.32	-0.03	73.52	76.51	-0.33
	3	3.55	3.58	-0.08	75.44	76.7	1.4
	4	3.55	3.53	-0.03	71.89	81.86	-7.31
	5	3.71	3.7	-0.04	76.63	78.85	0.44
	6			-0.05	175.87	178.15	0.38
		Referenced to HDO: 4.79 ppm					

Table S3b. NMR shifts of GlcA-dimer contained in a mixture of GlcA-Rha3S and IdoA-Rha3S (ratio ~7:3) originating from the digestion of Rha-GlcA-Rha3S with Enzyme P36_GH78. The chemical shifts are compared to the starting material NMR spectra are shown in Figs. S14 to S17. ESI-MS measurements support the structure shown in Figs. S18 and S19). The shifts of the GlcA-component are shown in Table S3a. Due to the very low concentration of the sample, complete assignment of the minor IdoA component was not possible for a significant amount of atoms.

		¹ H-Shifts (ppm)			¹³ C Shifts (ppm)		
		Rha-IdoA Rha3S	Ref Rha3S-IdoA- Rha3S	Delta (A1=A1)	Rha-IdoA Rha3S	Ref Rha3S-IdoA- Rha3S	Delta (A1=A1)
a α-Rhamnose	1	93.68	96.38	0	93.68	96.38	0
	2	69.36	72.1	-0.04	69.36	72.1	-0.04
	3	78.67	81.47	-0.1	78.67	81.47	-0.1
	4	78.83			78.83		
	5	*			*		
	6	17.05	19.7	0.05	17.05	19.7	0.05
b β-Rhamnose	1	93.15	95.83	0.02	93.15	95.83	0.02
	2	*	72.6		*	72.6	
	3	80.69			80.69		
	4	*	78.62		*	78.62	
	5	*			*		
	6	17.05	19.7	0.05	17.05	19.7	0.05
I(A) Iduronic acid (of α-anomer)	1	102.81	105.68	-0.17	102.81	105.68	-0.17
	2	70.14	73.19	-0.35	70.14	73.19	-0.35
	3	76.73			76.73		
	4	71.51		74.21	71.51		
	5	*			*		
	6	*	177.25		*	177.25	
I(B) Iduronic acid (of β-anomer)	1	102.81	105.68	-0.17	102.81	105.68	-0.17
	2	70.14	73.19	-0.35	70.14	73.19	-0.35
	3	76.73			76.73		
	4	71.51		74.21	71.51		
	5	*			*		
	6	*	177.25		*	177.25	
*... unambiguous assignment not possible		<i>Referenced to HDO: 4.79 ppm</i>					

¹H NMR (D₂O, 600.15 MHz)

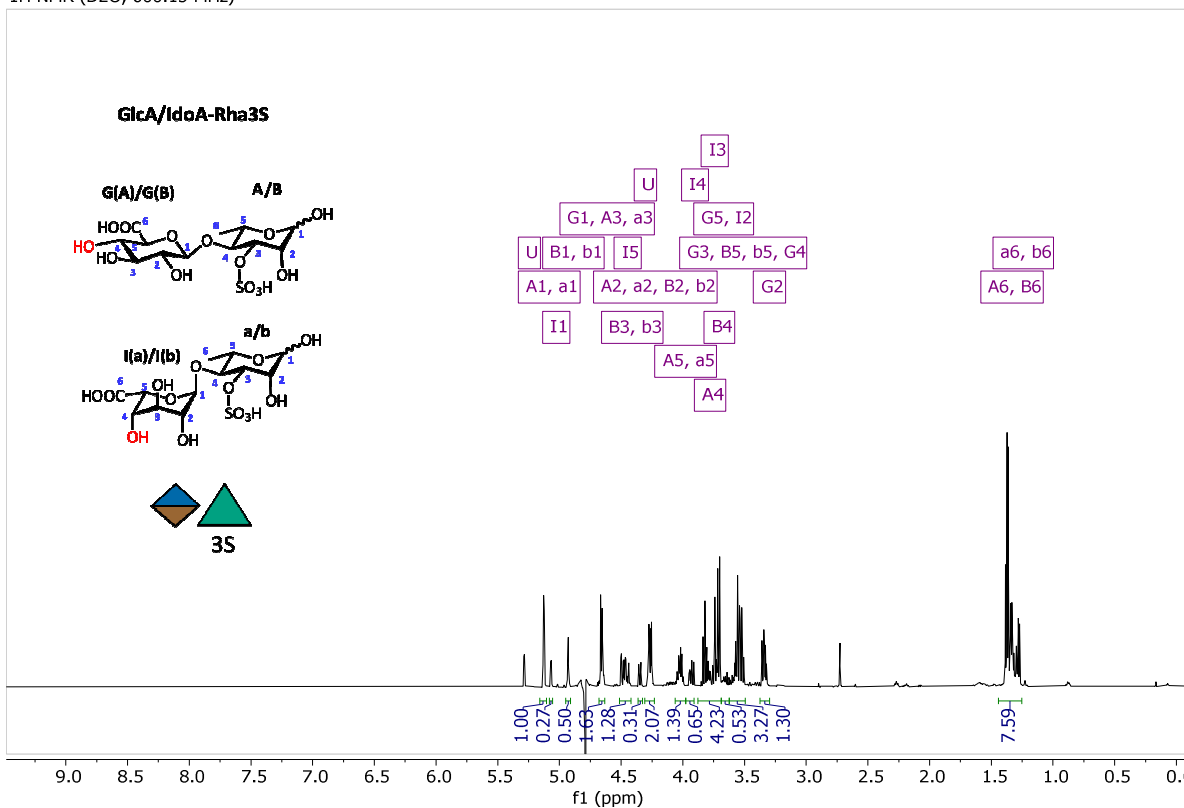


Figure S14. ¹H-NMR (with water suppression) of the dimer mixture containing GlcA-Rha3S and IdoA-Rha3S (ratio ~7:3) together with an unidentified substance (U) – Full View

¹H NMR (D₂O, 600.15 MHz)

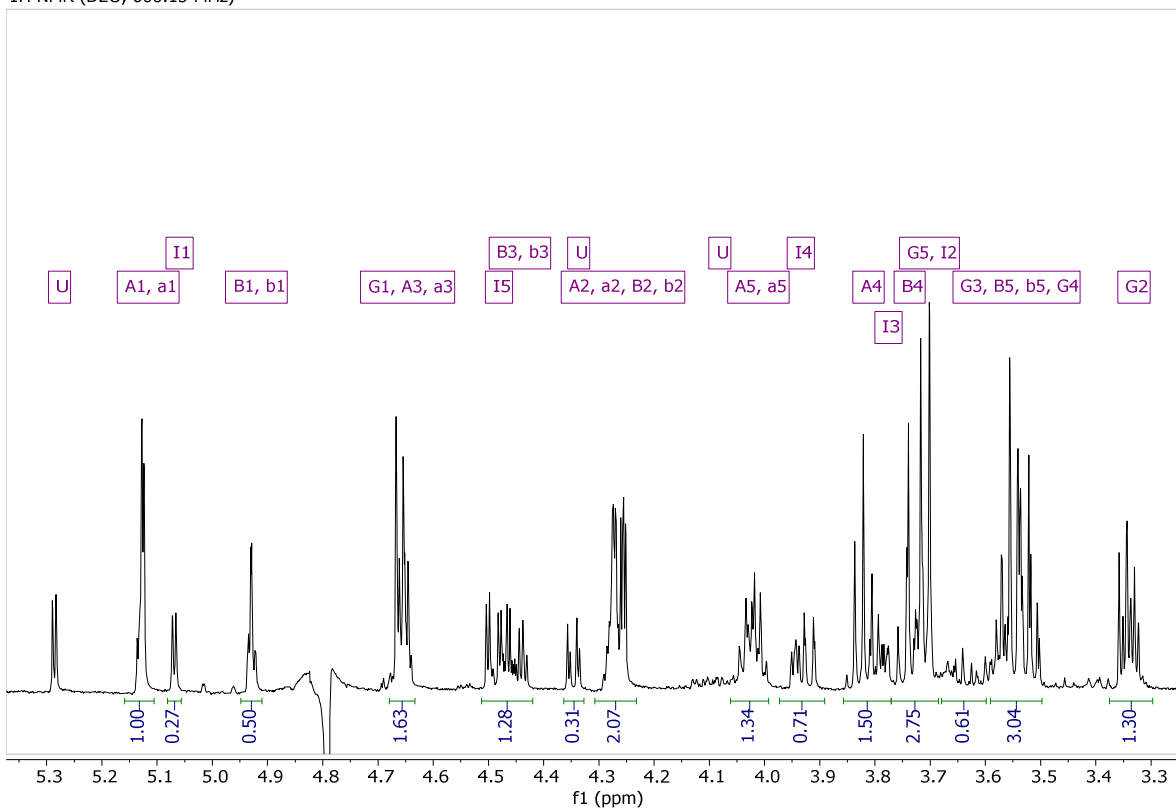


Figure S15. ¹H-NMR (with water suppression) of the dimer mixture containing GlcA-Rha3S and IdoA-Rha3S (ratio ~7:3) together with an unidentified substance (U) – Zoom into the carbohydrate region

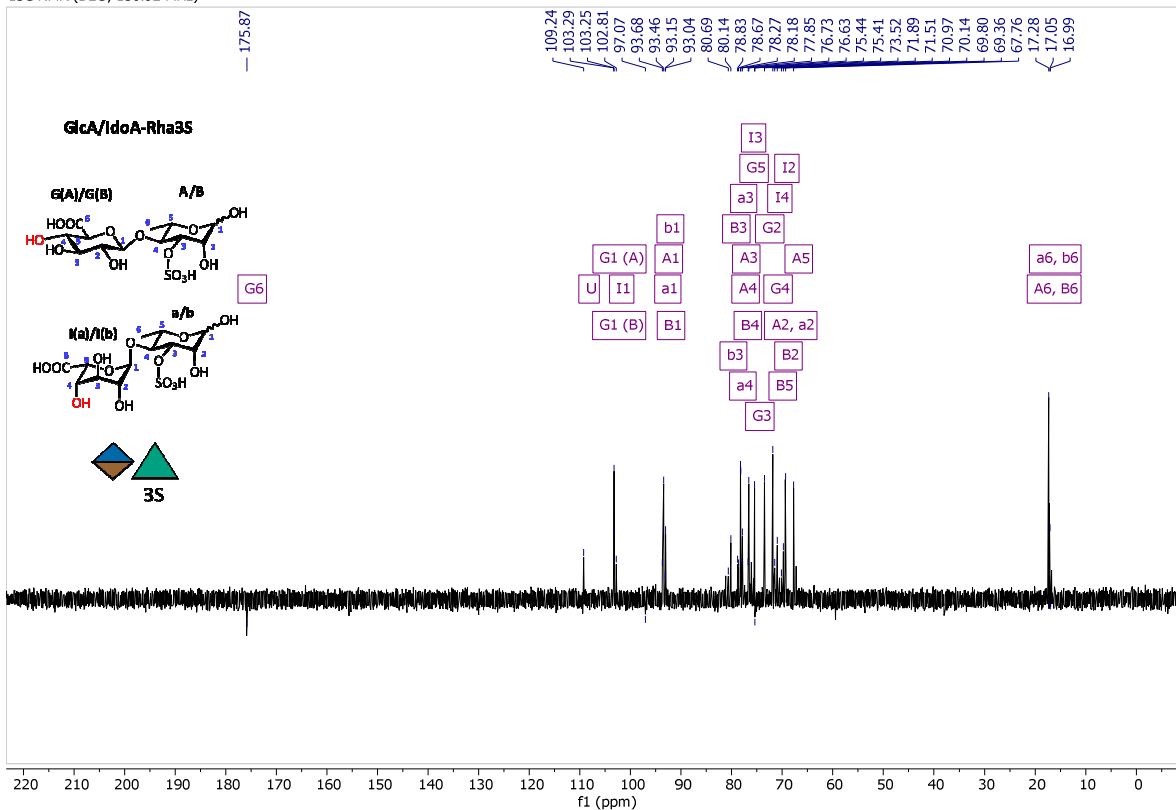


Figure S16. ¹³C-NMR of the dimer mixture containing GlcA-Rha3S and IdoA-Rha3S (ratio ~7:3) together with an unidentified substance (U) – FullView

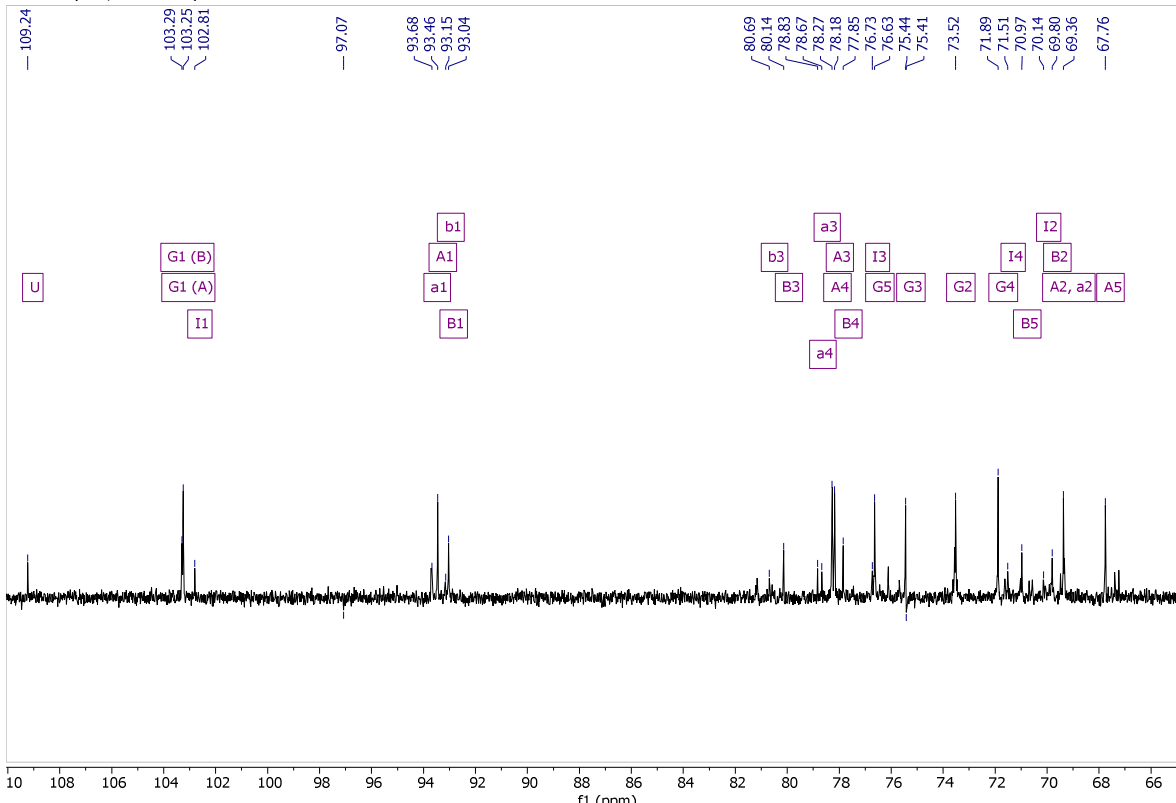


Figure S17. ¹³C-NMR of the dimer mixture containing GlcA-Rha3S and IdoA-Rha3S (ratio ~7:3) together with an unidentified substance (U) – Zoom into the carbohydrate region

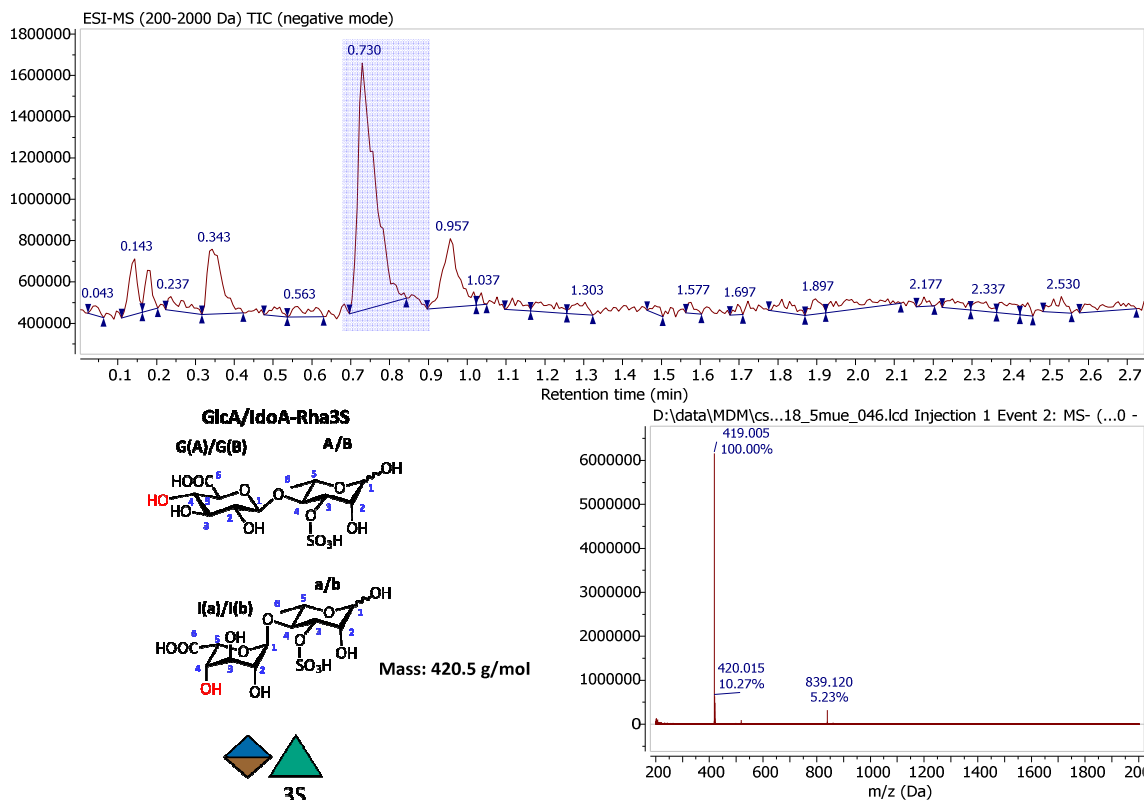


Figure S18. HPLC-MS (ESI-) measurement of the trimer mixture containing GlcA-Rha3S and IdoA-Rha3S (ratio ~7:3). Showing the [M-1] molecule peak of the compounds

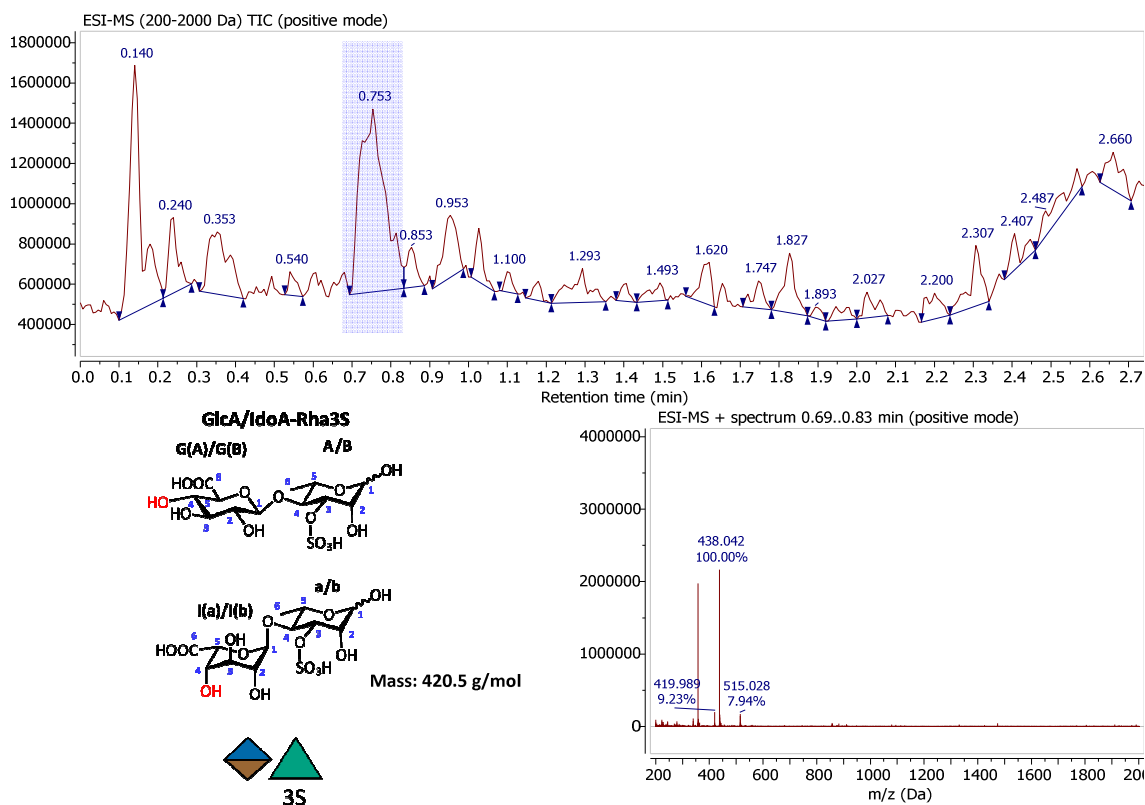


Figure S19. HPLC-MS (ESI+) measurement of the trimer mixture containing GlcA-Rha3S and IdoA-Rha3S (ratio ~7:3). Showing the [M+18] molecule peak of the compounds.

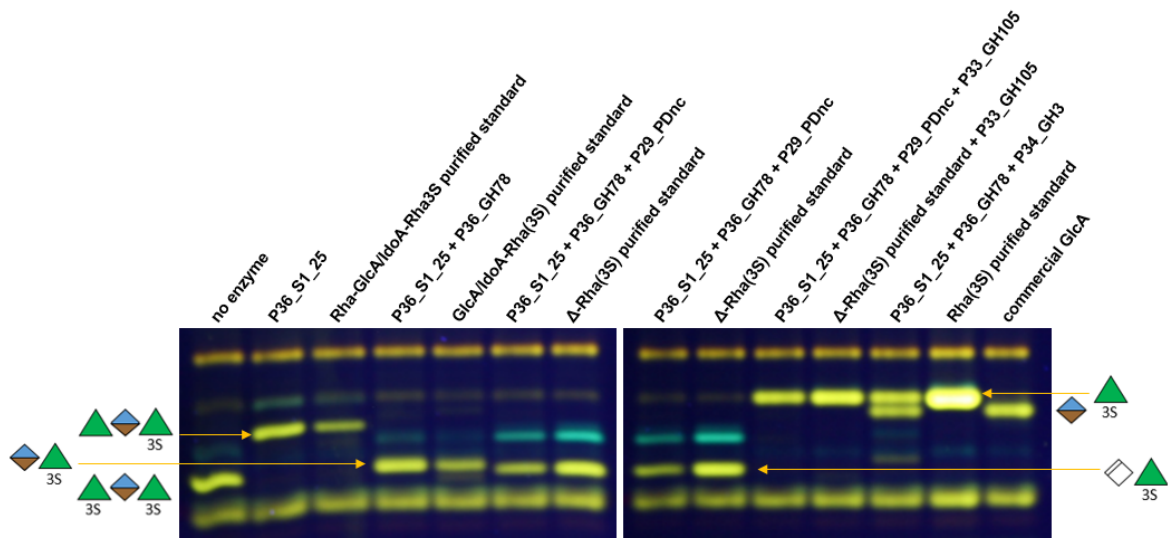


Figure S20. Complete-Gel FACE-analysis of the alternative ulvan degradation pathway of *Formosa agariphila*. In reactions containing P29_PDnc the other enzymes were heat-inactivated before addition of P29_PDnc to prevent a degradation of the dehydratase product by P33_GH105. All used products and standards (except GlcA) have been isolated and confirmed by MS and NMR measurement. The standard for GlcA was obtained from Roth. All products represent the mixture of both oligomers containing one of the epimers GlcA or IdoA. The ratio between GlcA- and IdoA-containing oligomers is ~70:30 (10, 18).

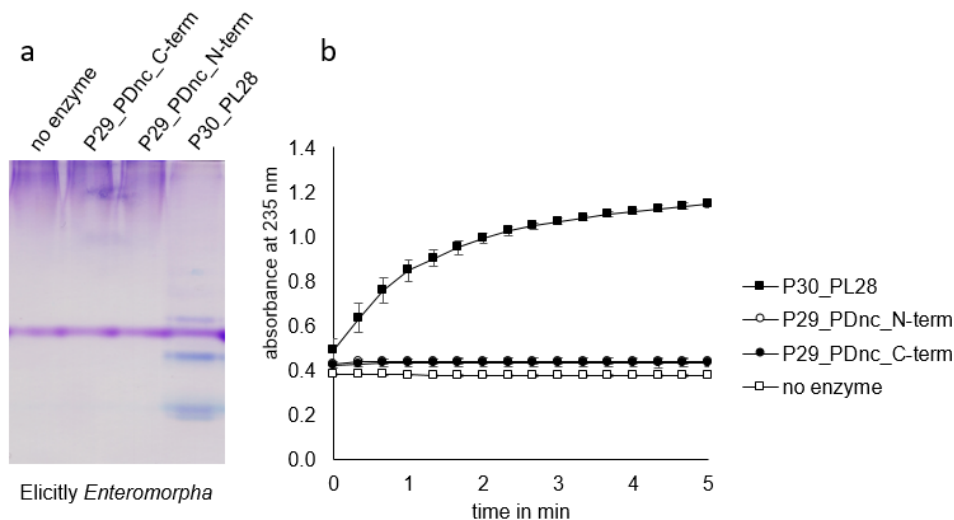


Figure S21. Analysis of lyase activity of P29_PDnc on commercial ulvan (Elicityl, France) from *Enteromorpha* sp. (left) C-PAGE analysis and (right) lyase assay. Polymeric ulvan from seven different sources has been incubated with both P29_PDnc variants with N-terminal or C-terminal His-tag and P30_PL28 as positive control or without enzymes as negative control. We compared two commercially available ulvans from Elicityl extracted from *Enteromorpha* sp. (S20) or *Ulva* sp. (S21), and five self-isolated ulvans from “kulau sea lettuce” containing *Ulva* spp. from Spain (S23), and from self-collected *Ulva* sp. from Lubmin (Baltic Sea) (S22), France (Atlantic Ocean) (S24, S25) and Helgoland (North Sea) (S26) (see the following Fig. S21-S26).

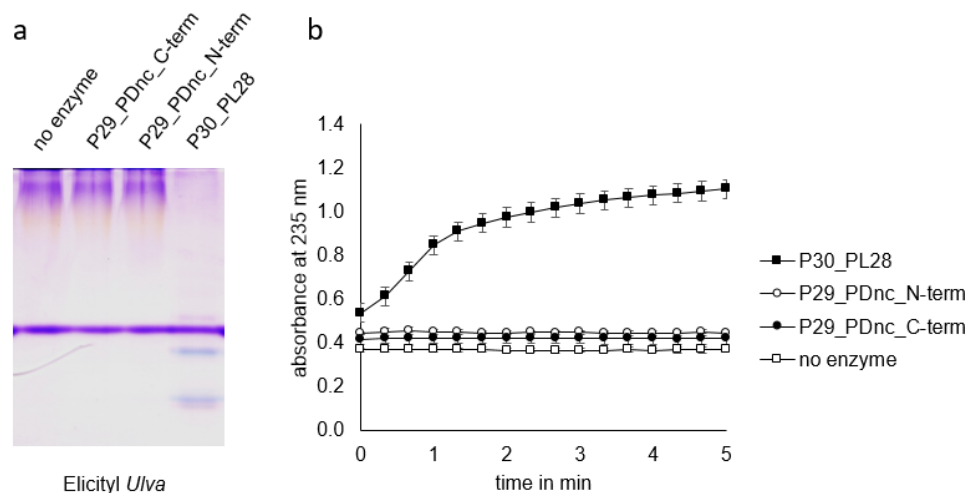


Figure S22. Analysis of lyase activity of P29_PDnc on commercial ulvan (Elicityl, France) from *Ulva* sp. (a) C-PAGE analysis and (b) lyase assay. Polymeric ulvan has been incubated with both P29_PDnc variants with *N*-terminal or *C*-terminal His-tag and P30_PL28 as positive control or without enzymes as negative control.

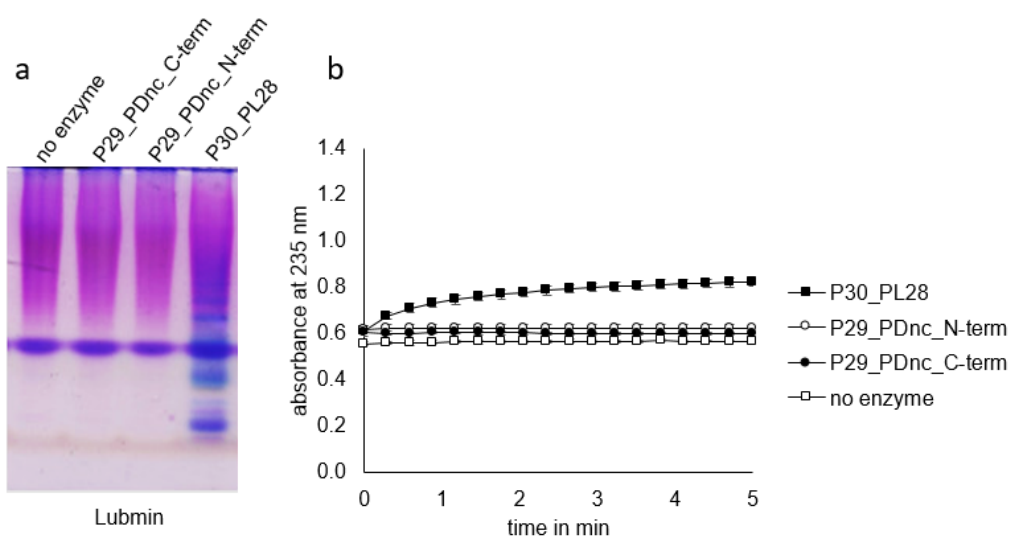


Figure S23. Analysis of lyase activity of P29_PDnc on self-isolated ulvan from *Ulva* sp. from Lubmin (Baltic Sea). (a) C-PAGE analysis and (b) lyase assay. Polymeric ulvan has been incubated with both P29_PDnc variants with *N*-terminal or *C*-terminal His-tag and P30_PL28 as positive control or without enzymes as negative control.

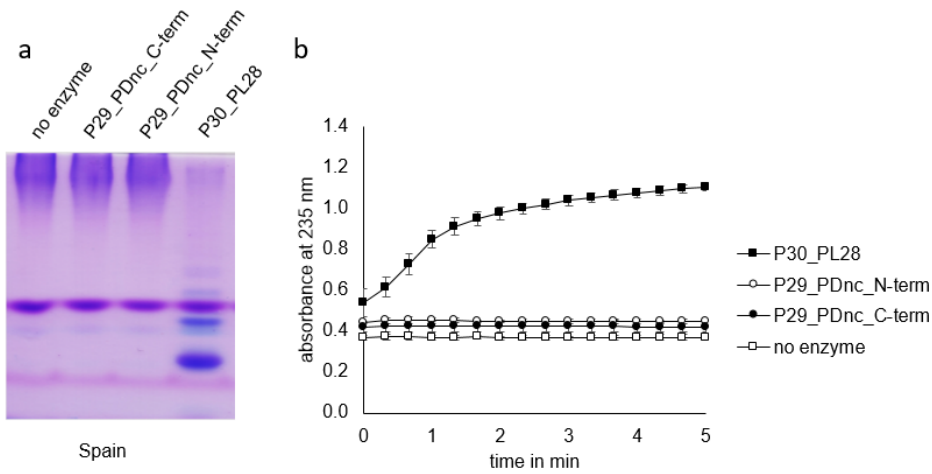


Figure S24. Analysis of lyase activity of P29_PDnc on self-isolated ulvan from *Ulva* sp. from Spain. (a) C-PAGE analysis and (b) lyase assay. Polymeric ulvan has been incubated with both P29_PDnc variants with *N*-terminal or *C*-terminal His-tag and P30_PL28 as positive control or without enzymes as negative control.

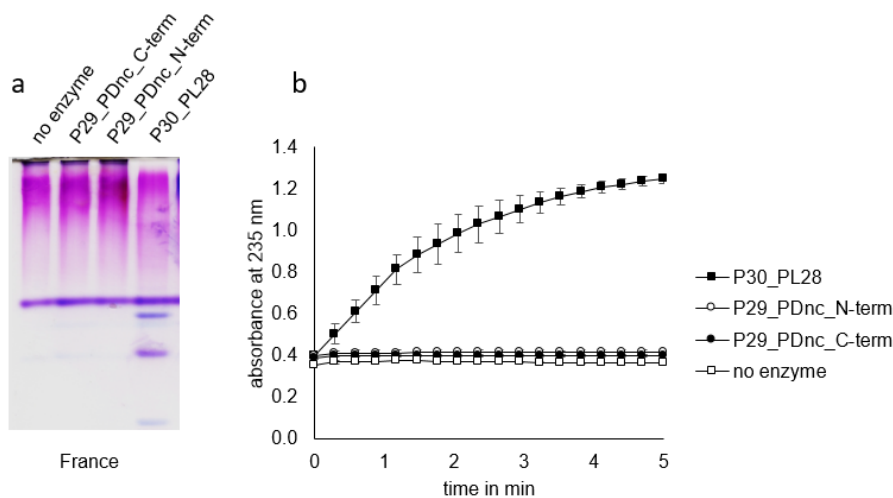


Figure S25. Analysis of lyase activity of P29_PDnc on self-isolated ulvan from self-collected *Ulva* sp. from France (Atlantic Ocean). (a) C-PAGE analysis and (b) lyase assay. Polymeric ulvan has been incubated with both P29_PDnc variants with *N*-terminal or *C*-terminal His-tag and P30_PL28 as positive control or without enzymes as negative control.

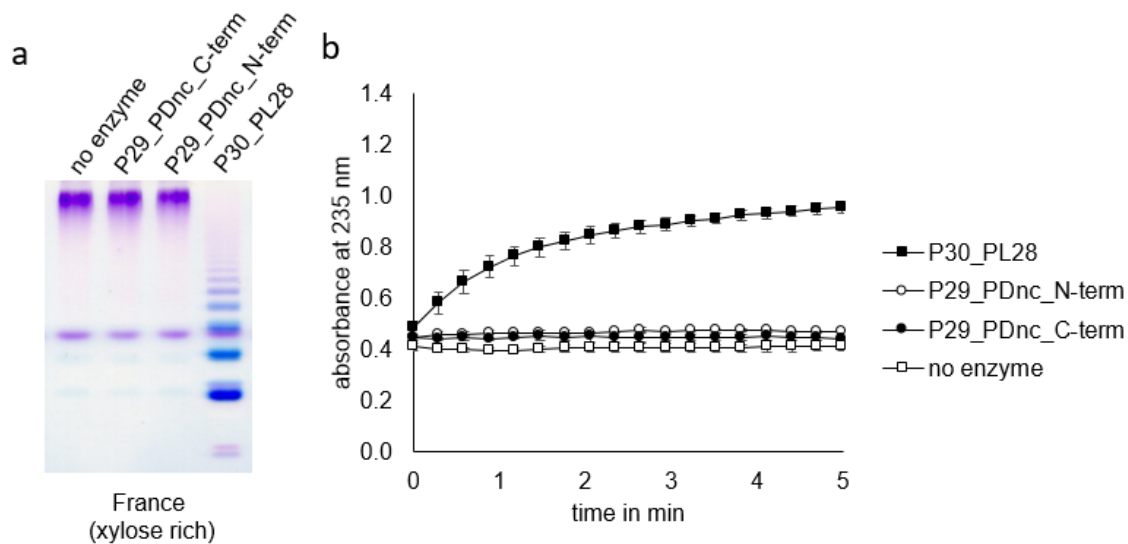


Figure S26. Analysis of lyase activity of P29_PDnc on self-isolated xylose-rich ulvan from self-collected *Ulva* sp. from France (Atlantic Ocean). (a) C-PAGE analysis and (b) lyase assay. Polymeric ulvan has been incubated with both P29_PDnc variants with *N*-terminal or *C*-terminal His-tag and P30_PL28 as positive control or without enzymes as negative control. Figure S26 shows the same data as in Figures 3A and 3B to allow a direct comparison with Figures 3C and 3D in the main publication as well as with Figure S27

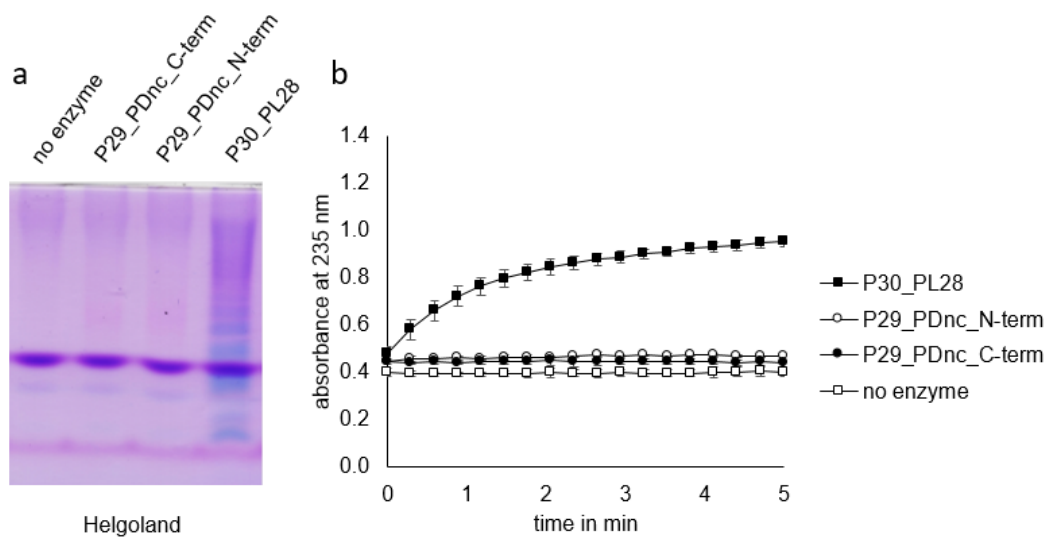


Figure S27. Analysis of lyase activity of P29_PDnc on self-isolated ulvan from self-collected *Ulva* sp. from Helgoland. (a) C-PAGE analysis and (b) lyase assay. Polymeric ulvan has been incubated with both P29_PDnc variants with *N*-terminal or *C*-terminal His-tag and P30_PL28 as positive control or without enzymes as negative control.

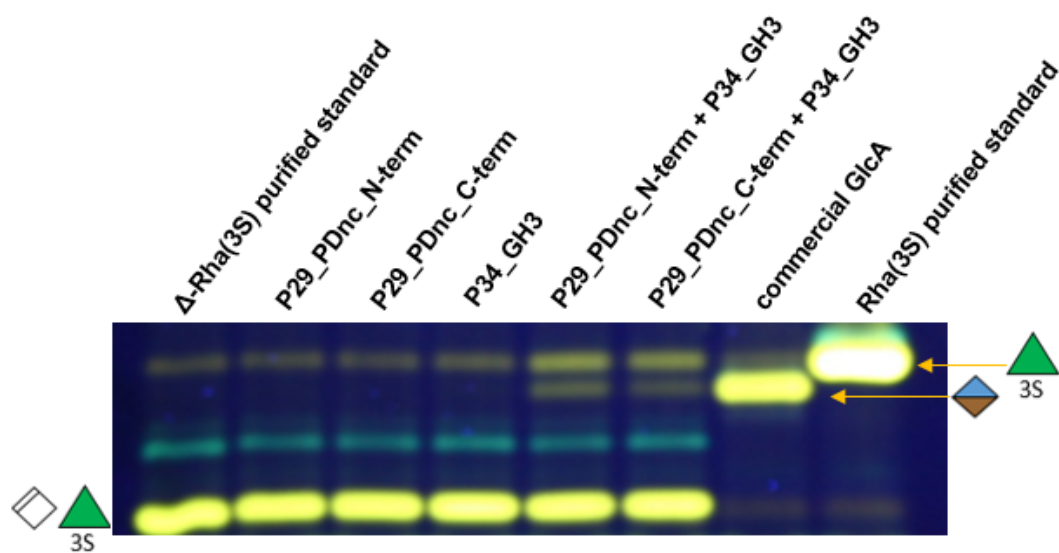


Figure S28. Hydration of Δ -Rha3S using P29_PDnc + P34_GH3 with FACE-analysis. P29_PDnc and P34_GH3 have been incubated on the disaccharide Δ -Rha3S leading to the formation of GlcA or IdoA and Rha3S in monomeric form by shifting the chemical equilibrium of the dehydration step by P29_PDnc to the educt by degrading the produced disaccharide GlcA/IdoA with P34_GH3.

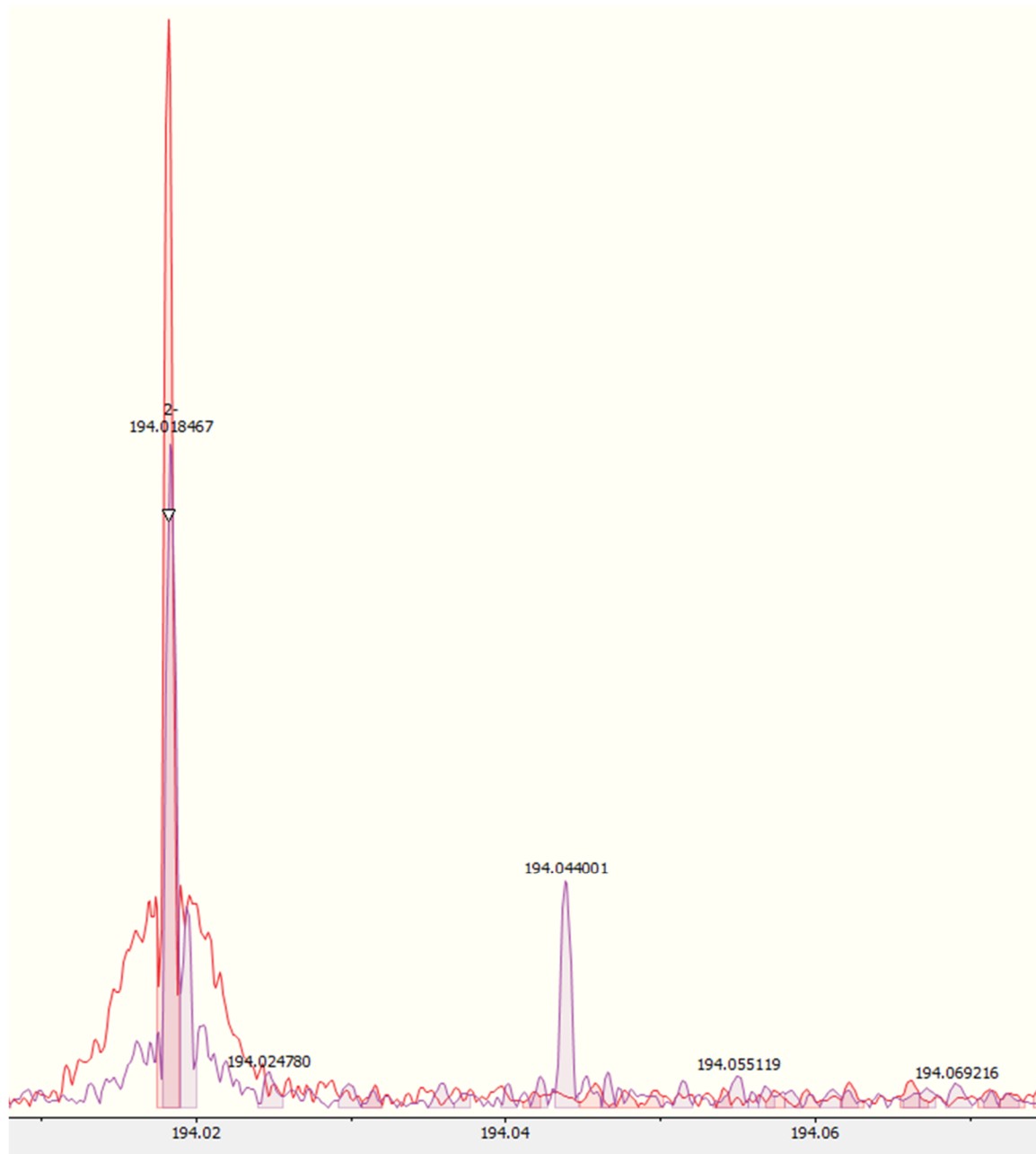


Figure S29. Mass spectrum for the hydration of Δ -Rha3S using P29_PDnc + P34_GH3. P29_PDnc and P34_GH3 were incubated on the disaccharide Δ -Rha3S leading to the formation of GlcA or IdoA and Rha3S in monomeric form by shifting the chemical equilibrium of the dehydration step by P29_PDnc to the educt by degrading the produced disaccharide GlcA/IdoA with P34_GH3. The mass spectrum shows the [M-H]⁻ mass of glucuronic acid +1 (194.044001). The red spectrum represents the negative control without enzymes (glucuronic acid peak is missing here) while the purple spectrum represents the reaction with enzymes. The left large peak is not shown in its complete height to remain clarity of the purple reaction peak. The y-axis (intensity) is not shown in this image section.

P29_R300E	1246	19.17	0.23
P29_K303I	1075	16.54	0.23
P29_Y306F	967	15.02	0.23

Table S6. Nucleotide sequence (codon-optimized) encoded in the enzymes study

>P36_S1_25	CAGACCGTGAAGAAGGAGAAGCCGAACATCATTTTCATCCTGACCGACGAT CAGCGTTTCGACGCGATTGGTTATGCGGGTAATAAGTTCGTGAACACCCCG GAAATGGATAAGCTGGCGCAGCAAGGCACCTACTTTGACCACGCGATCGTT ACCACCCCGATTTGCGCGGCGAGCCGTGCGAGCCTGTGGACCGGCCTGC ATGAGCGTAGCCACAACCTTCAACTTTTCAGACCGGTAACGTGCGTGAGGAAT ATATGAACAACGCGTACCCGAAGCTGCTGAAAAACAACGGTTACTATACCG GTTTTCTATGGCAAATACGGTGTTCGTTATGACAACCTGGAAAGCCAATTCGA CGAGTTTGAAAGCTATGATCGTAACAACCGTTACAAAGATAAGCGTGGCTA CTATTACAAGACCATCAACAACGACACCGTGCACCTGACCCGTTACACCGG TCAGCAAGCGATCGACTTCATTGATAAAAACGCGACCAACACCCAGCCGTT CATGCTGAGCCTGAGCTTTAGCGCGCCGCACGCGCATGATGGTGCGCCGG AACAGTATTTTTGGCAAACCACCCAGCAGCGCTGCTGCAAGATACCAACC TGCCGGGTCCGGACCTGGCGGATGAGAAGTACTTCCTGGCGCAGCCGCAA GCGGTTTCGTGACGGTTTTAACCGTCTGCGTTGGACCTGGCGTTATGACGAT CCGGAGAAGTACCAGCACAGCCTGAAAGGCTATTACCGTATGATCAGCGG TATTGACCTGAAATCAAGAAAATTCGTGATAAACTGAAGGAGAAAGGTGT GGACAAAAACACCGTGATCATTGTTATGGGCGATAACGGTTATTTCTGGG CGAACGTCAACTGGCGGGCAAGTGGCTGATGTACGACAACAGCATCCGTG TGCCGCTGATTGTTTTGATCCGCGTGTTAACAAACACCAGGACATCAGCG AGATGGTGTGAACATCGACGTTACCCAAACCATTGCGGATCTGGCGGGC GTGAAGGCGCCGAAAGCTGGCAGGGCAAGAGCCTGCTGCCGCTGGTTA AACAAGAAACCAGCACCATCAGCCGTGATACCATCCTGATTGAGCACCTGT GGGACTTCGAAAACATTCCGCCGAGCGAGGGCGTGCGTACCGAGGAATGG AAGTATTTTCGTTACGTTAACGATAAAACCATCGAGGAACTGTATAACATTA AGAAAGACCCGAAAGAAATCAACAACCTGATTGGTAAGAAAAAGTACCAGA ACGTGGCGAAGGCGCTGCGTGAAAAACTGGACGAACTGATTGCGAAAAAT AGCGACGAATTCCGTAA
------------	--

Article III

RESEARCH

Open Access



Metabolic engineering enables *Bacillus licheniformis* to grow on the marine polysaccharide ulvan

Theresa Dutschei¹, Marie-Katherin Zühlke^{2,3}, Norma Welsch^{2,3}, Tom Eisenack², Maximilian Hilkmann^{2,3}, Joris Krull^{3,4,5}, Carlo Stühle¹, Stefan Brott¹, Alexandra Dürwald², Lukas Reisky¹, Jan-Hendrik Hehemann^{3,4,5}, Dörte Becher⁶, Thomas Schweder^{2,3*} and Uwe T. Bornscheuer^{1,3*}

Abstract

Background: Marine algae are responsible for half of the global primary production, converting carbon dioxide into organic compounds like carbohydrates. Particularly in eutrophic waters, they can grow into massive algal blooms. This polysaccharide rich biomass represents a cheap and abundant renewable carbon source. In nature, the diverse group of polysaccharides is decomposed by highly specialized microbial catabolic systems. We elucidated the complete degradation pathway of the green algae-specific polysaccharide ulvan in previous studies using a toolbox of enzymes discovered in the marine flavobacterium *Formosa agariphila* and recombinantly expressed in *Escherichia coli*.

Results: In this study we show that ulvan from algal biomass can be used as feedstock for a biotechnological production strain using recombinantly expressed carbohydrate-active enzymes. We demonstrate that *Bacillus licheniformis* is able to grow on ulvan-derived xylose-containing oligosaccharides. Comparative growth experiments with different ulvan hydrolysates and physiological proteogenomic analyses indicated that analogues of the *F. agariphila* ulvan lyase and an unsaturated β -glucuronidase are missing in *B. licheniformis*. We reveal that the heterologous expression of these two marine enzymes in *B. licheniformis* enables an efficient conversion of the algal polysaccharide ulvan as carbon and energy source.

Conclusion: Our data demonstrate the physiological capability of the industrially relevant bacterium *B. licheniformis* to grow on ulvan. We present a metabolic engineering strategy to enable ulvan-based biorefinery processes using this bacterial cell factory. With this study, we provide a stepping stone for the development of future bioprocesses with *Bacillus* using the abundant marine renewable carbon source ulvan.

Keywords: Ulvan, Marine polysaccharide, Green algae, Biorefinery process, *Bacillus licheniformis*

Background

Eutrophication and global warming impact frequency and extent of algal blooming events and thus accumulation of algal biomasses in coastal areas [1–3]. Despite algae or algal products being already used in food, cosmetics, biotechnology and pharmaceutical industry [4–6], washed up algae are still largely unexploited. As a consequence, interest has been raised to develop processes that convert this cheap biomass to valuable products [7] and first

*Correspondence: schweder@uni-greifswald.de;
uwe.bornscheuer@uni-greifswald.de

¹ Department of Biotechnology & Enzyme Catalysis, Institute of Biochemistry, University of Greifswald, 17487 Greifswald, Germany

² Department of Pharmaceutical Biotechnology, Institute of Pharmacy, University of Greifswald, 17487 Greifswald, Germany

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

attempts are already underway [8]. Amongst a variety of compounds that could be harnessed, polysaccharides are attractive targets. They account for up to 50% of macroalgal biomass and mostly represent cell wall or storage components [9–11]. These polysaccharides are highly diverse in their structure and composition [12]. Targeting this versatile substrate pool thus requires a multitude of enzymes which are usually encoded in highly clustered genomic regions of polysaccharide degrading bacteria. These so-called polysaccharide utilization loci (PUL) encode proteins to mediate binding, degradation and uptake of saccharides [13]. Recently, we were able to elucidate a complex enzymatic cascade to completely deconstruct polymeric ulvan to monomeric sugar compounds using enzymes from the marine flavobacterium *Formosa agariphila* KM3901^T recombinantly expressed in *Escherichia coli* [14, 15]. Ulvan is the main cell wall polysaccharide in the green seaweed *Ulva* spp. [16]. The sugar backbone is composed of L-rhamnose, D-xylose and D-glucuronic acid/L-iduronic acid and is highly branched and sulfated. Moreover, the monosaccharide composition varies between species and sampling sites [17]. In *F. agariphila*, ulvan lyases catalyze the initial degradation step, releasing several oligosaccharide species with a 5-dehydro-4-deoxy-D-glucuronate at the non-reducing end [14, 18]. This unsaturated moiety is then removed by glycoside hydrolases (GH), which allows further GH-mediated hydrolysis of oligosaccharides prior to or after their desulfation [14]. On the one hand, such enzyme cascades can be used for the production of rare (sulfated) sugar oligosaccharides that could be interesting due to their immunomodulating activities [19]. On the other hand, these hydrolysates may represent a starting material for biotechnological processes as alternative feedstock for common sugars like glucose for microbial fermentation [5]. Microbial engineering and systems biology can further help to develop such new biomass based bioprocesses [20, 21]. Consequently, in-depth characterization of the selected microbial production species is a prerequisite for strain optimization. The well-established biotechnological work horse *Bacillus licheniformis* is an attractive target to be investigated for the utilization of alternative algal derived biomasses: It produces a variety of enzymes to degrade plant materials, it is a generally recognized as safe (GRAS) strain, has a fast growth rate and is already of high industrial importance [22]. This bacterial cell factory naturally produces the extracellular protease subtilisin [23], which has been developed into industrial production due to its widespread use in detergents [24]. In addition, first processes that use *B. licheniformis* to convert plant biomass into valuable products have already been established. This includes metabolic engineering approaches which enabled the production of

acetoin, 2,3-butanediol or lactic acid from kitchen waste or corncob molasses [22, 25–27]. Furthermore, the production of extracellular proteins from algal feedstock [28] was already studied to broaden up the possible use of this bacterium in fermentation processes.

In order to develop an ulvan based bioprocess, we investigated a variety of bacterial strains for their ability to utilize ulvan and identified the industrially relevant bacterium *B. licheniformis* DSM13, which is able to grow on pre-digested ulvan. We investigated strain specific metabolic properties of this bacterium, which are required for ulvan utilization. Our study provides first insights into the development of a potential ulvan based bioprocess with *Bacillus* species.

Results and discussion

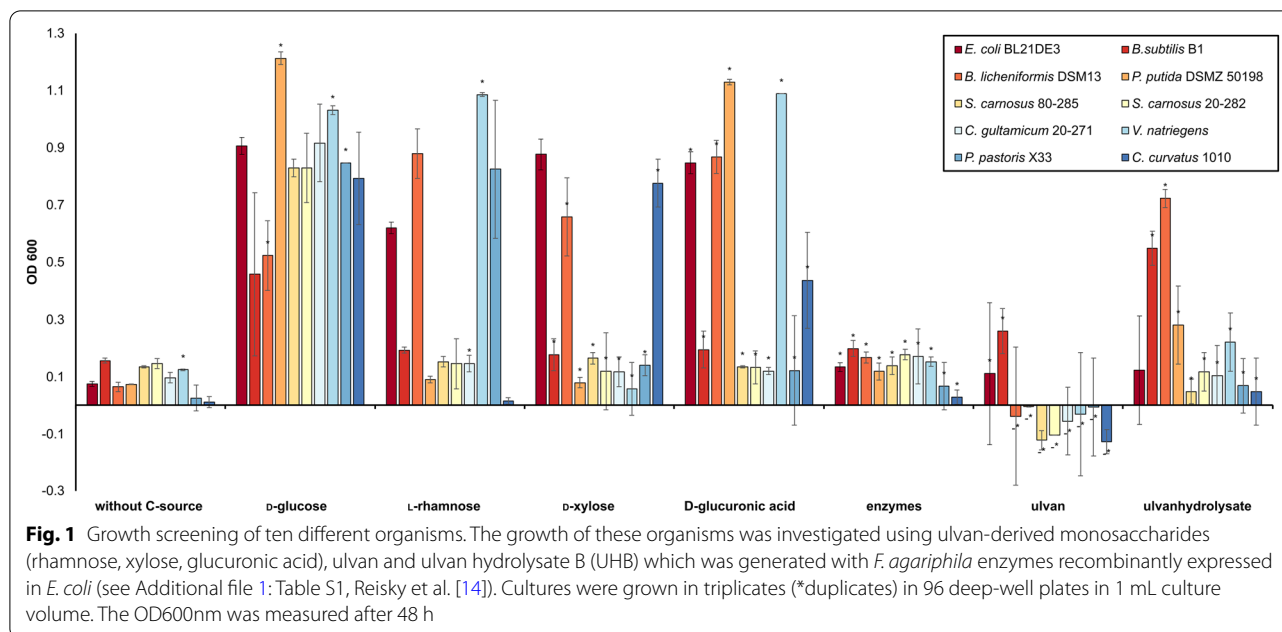
Bacillus licheniformis DSM13 efficiently consumes ulvan-derived monomers

In a first attempt, we screened 10 different strains for their ability to grow on ulvan and ulvan-derived monosaccharides, as single monomers or as monosaccharide mixture (Fig. 1, Additional file 1: Fig. S1). While none of the strains grew on raw ulvan, *B. licheniformis* DSM13, *Cryptococcus curvatus* 1010 and *Pseudomonas putida* DSMZ 50198 consumed the monomer cocktail derived from ulvan digestion using the complete enzymatic cascade of *F. agariphila* that was recombinantly expressed in *E. coli* BL21(DE3) as described previously (Additional file 1: Table S1, Fig. S2) [14, 15].

This mixture, ulvan hydrolysate B (UHB), provided L-rhamnose, D-xylose, D-glucuronic acid and 5-dehydro-4-deoxy-D-glucuronate. *B. licheniformis* DSM13 grew also well on each individual monosaccharide present in the mixture, as described before [29, 30], even better than on D-glucose. Although *P. putida* DSMZ 50198 and *Bacillus subtilis* B1 consumed UHB, they were not able to grow on L-rhamnose (Fig. 1), which is known for *P. putida* [31], but disagrees with observations reported for *C. curvatus* 1010 [32]. *P. putida* DSMZ 50198 also lacks the ability to grow on D-xylose. Growth experiments identified *B. licheniformis* DSM13 as a suitable candidate for further investigations to establish an ulvan sugar-based bioprocess.

Bacillus licheniformis DSM13 grows and accumulates proteases on fully digested ulvan

To investigate the suitability of the abundant macroalgal polysaccharide ulvan as feedstock for production processes, we quantified exemplarily protease activity via the AAPF-assay [33] during cultivation, like the alkaline serine protease (AprX, Q65IP4), subtilisin protease Apr (Q65LP7) and extracellular serine protease Vpr (Q65DN2). Following growth over time, *B. licheniformis*



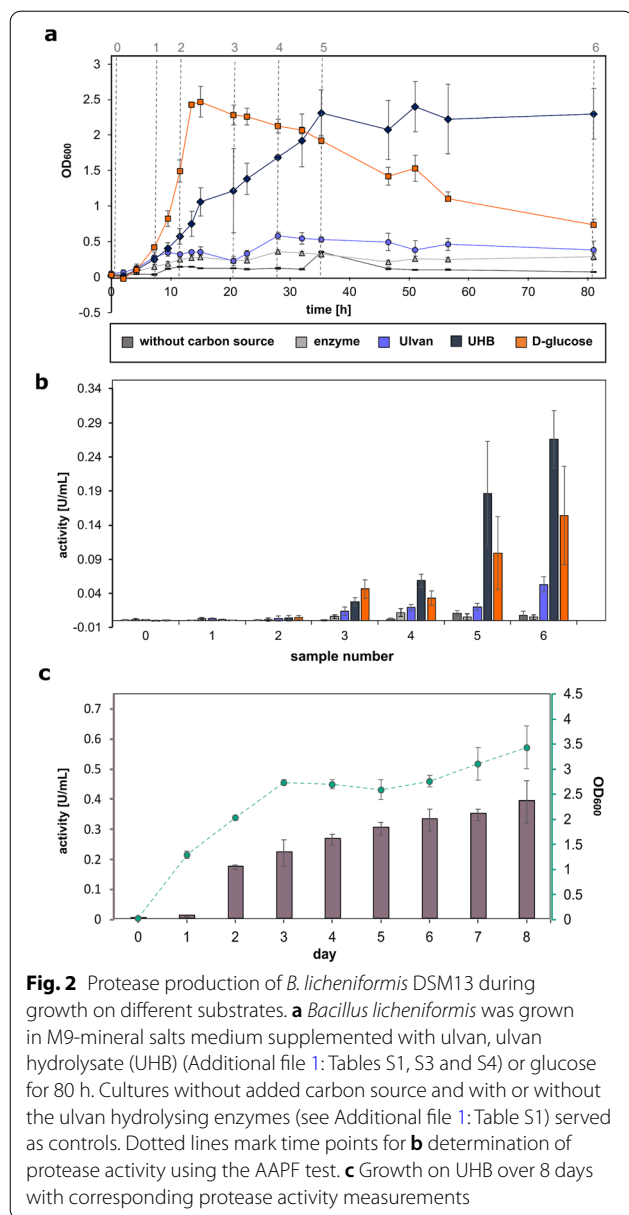
DSM13 grew more slowly on UHB compared to D-glucose, but reached a comparable maximum optical density that was stable until the end of the experiment (Fig. 2a). At the same time, protease activity increased over time (Fig. 2b) and was stable even in prolonged cultivations (Fig. 2c). These growth experiments revealed a constant stationary phase over more than 5 days for *B. licheniformis* using UHB as the sole carbon source. The resulting increased biomass until the end of the cultivation improved protease production significantly compared to the glucose-based cultivations.

Capabilities of *B. licheniformis* DSM13 to grow on ulvan-derived oligosaccharides

The initial growth experiments demonstrated the physiological capability of *B. licheniformis* DSM13 to utilize ulvan-specific monosaccharides (Fig. 2, Additional file 1: Fig. S3). Consumption of these monosaccharides as well as the fact that this bacterium is well known to degrade plant material [34, 35], gave reasons to suspect also the acceptance of ulvan-derived oligosaccharides. Therefore, 12 different ulvan hydrolysates were examined as potential substrates (Additional file 1: Fig. S4). These enzymatically digested ulvan-extracts cover different levels of ulvan depolymerisation as described in our previous studies and thus differ in their mono- and oligosaccharide composition [14, 15]. Again, each hydrolysate, including the aforementioned UHB, was produced using selected *F. agariphila* ulvan-degrading enzymes recombinantly expressed in *E. coli* BL21(DE3) (Additional file 1: Table S1, Fig. S2; Fig. 3a) [14]. The cell densities

achieved after 24 h of cultivation identified the required level of hydrolyzation to allow growth of *B. licheniformis* DSM13. At the same time, they indicated which enzymatic activities might be missing in *Bacillus* and would thus enable growth on higher degrees of polymerization or ulvan itself (Fig. 3b). The ulvan lyase-generated hydrolysates improved digestibility only to a small extent (P30_PL28 > P10_PL40), similar to P31_GH39 and P17_GH2 pre-digestion (Additional file 1: Fig. S4).

However, optical densities were considerably increased if the ulvan lyase activity of P30_PL28 was either supported by the unsaturated glucuronyl hydrolase P1_GH88 or the glycoside hydrolase P33_GH105 (UHA). This may be due to the release of smaller oligosaccharides and unsaturated uronic acids as carbon source in UHA. The P30_PL28 ulvan lyase cleaves the ulvan polymer between α -L-rhamnose-3-sulfate-(1,4)- β -D-glucuronic acid, which produces an unsaturated uronic acid at the non-reducing end of the released oligosaccharide, which is specific for lyases. This unsaturated uronic acid (4-deoxy- α -L-threo-hex-4-enopyranuronic acid) is then hydrolyzed by P33_GH105. Indeed, previous growth experiments confirmed *B. licheniformis* DSM13 to consume 4-deoxy- α -L-threo-hex-4-enopyranuronic acid (Additional file 1: Fig. S3). This way, not only easily digestible monosaccharides are released from oligosaccharides using P33_GH105, its activity also enables P30_PL28 to cleave the oligomer even further since lyase products inhibit subsequent lyase activities [14, 18]. In addition, unsaturated uronic acids in oligosaccharides might hinder their subsequent disassembling by *B. licheniformis* DSM13.



Interestingly, additional hydrolysis steps, which also included sulfatases, did not further improve growth. This led to the assumption that the two initial major enzyme activities of the ulvan degradation pathway [14], ensured by the ulvan lyase (PL28) and unsaturated glucuronyle hydrolases (GH105, GH88), provided an oligosaccharide mixture suitable for *B. licheniformis* DSM13 to degrade ulvan. This also indicated the availability of putative CAZymes in *B. licheniformis* DSM13 to utilize L-rhamnose, D-xylose and D-glucuronic acid from ulvan oligomers and to channel them into its carbon and energy metabolism.

Proteogenomic analysis of *B. licheniformis* DSM13

To further interpret our results and to explore the physiological potential of *B. licheniformis* to utilize ulvan derived sugars, we performed computational and proteome analyses. We analyzed the intracellular soluble as well as the extracellular proteomes of ulvan-, UHA- and UHB-cultivated cells (Fig. 3, Additional file 1: Fig. S5) compared to rhamnose and glucose cultures. In general, it is well known that *B. licheniformis* DSM13 secretes a variety of extracellular CAZymes to degrade polysaccharides [34, 35]. Correspondingly, computational analysis with the web server for automated CAZyme annotation, dbCAN2 [36, 37], identified 86 PLs, GHs and CEs to be encoded in its genome, 58 of which were captured by intracellular and extracellular proteomes (Additional file 2: Table S6; Additional file 1: Fig. S5).

Enzymes to cleave ulvan are lacking

Proteome and dbCAN2 analyses did not reveal suitable ulvanolytic enzyme activities of *B. licheniformis* DSM13 wild type strain, which are required for the initial digestion of ulvan, and thus confirmed our growth experiments shown in Fig. 2. The strain lacks PLs from families 24, 25, 28 and 40 [12] to cleave ulvan into oligosaccharides. Moreover, proteome analyses do not indicate PLs or GHs that depolymerize pectin and pectin components to also cleave ulvan, since corresponding proteins were either low abundant (PL11_1 Q65KY4) or quantified across all samples (PL1_5 Q65DC2, PL3_1 Q65EF5 and GH28 Q65F26, Additional file 1: Fig. S5; Additional file 2: Table S6). *B. licheniformis* DSM13 encodes two GH105s (Q65FY9, Q65KY9) as candidates to catalyze the next necessary enzymatic step in ulvan disassembling, but both of them were not detected in our proteome analyses during growth on ulvan oligosaccharides. Instead, they might be involved in rhamnogalacturonan I degradation, like in *B. subtilis* [39].

GHs that may disassemble oligosaccharides

Nevertheless, the adaptation of *B. licheniformis* DSM13 to pectin or hemicellulose usage may still allow for consumption of certain ulvan oligosaccharides as demonstrated by our growth experiments. Proteome analyses captured potentially involved GH43s, the GH43_4 (Q65D31) being highly abundant in ulvan and ulvan hydrolysate samples (1–3% of the total extracellular proteome) (Fig. 4, Additional file 2: Table S6, S7). GH43_4 (Q65D31, YxiA/Abn2) as well as GH43_5 (Q65GB9, AbnA) are both extracellular enzymes that degrade arabinans [40, 41]. By contrast, in *F. agariphila* a GH43_10 cleaved xylose moieties from ulvan-derived oligosaccharides [14]. The corresponding *B. licheniformis*

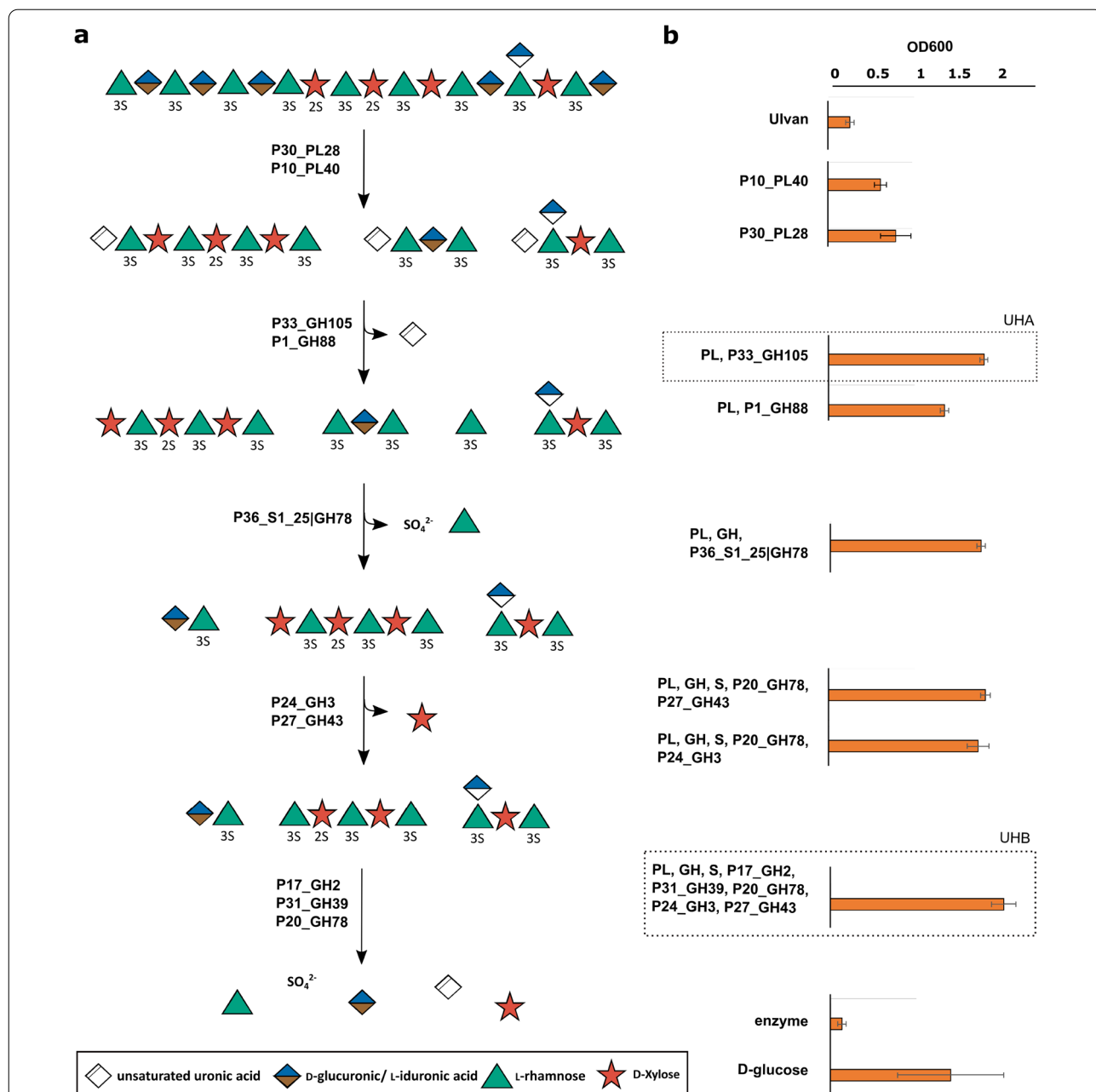


Fig. 3 In-depth analysis of the capability of *B. licheniformis* DSM13 to degrade ulvan-derived oligosaccharides. Ulvan was digested with selected enzymes or enzyme cocktails, described before [14, 15], to produce a total of 12 different ulvan hydrolysates (Additional file 1: Fig. S4). **a** These vary in their mono- and oligosaccharide content based on **b** the enzymes used and thus provide specific carbon sources for *B. licheniformis* DSM13. The DSM13 strain was cultivated in M9-mineral medium supplemented with ulvan or enzyme-generated ulvan hydrolysates and OD600 was measured after 24 h. Growth on hydrolysates UHA and UHB, which were used for further investigations, is highlighted. (PL: mix of P10_PL40; GH: mix of P33_GH105, P1_GH88; S: P36_S1_25|GH78)

DSM13 enzyme (Q65MB7) was not quantified by proteome analyses. However, several other GHs of family 1, 3 and 4 that might have xylosidase activity, as well as an unclassified (nc) GH, were quantified in the proteome of *B. licheniformis* DSM13 grown on ulvan extracts.

The involvement of sulfatases remains speculative

Ulvan degradation does not only require PLs and GHs to cleave the sugar chain, but also sulfatases to act on sulfated rhamnose or xylose units. To encounter the complexity of ulvan composition and its degree of sulfation,

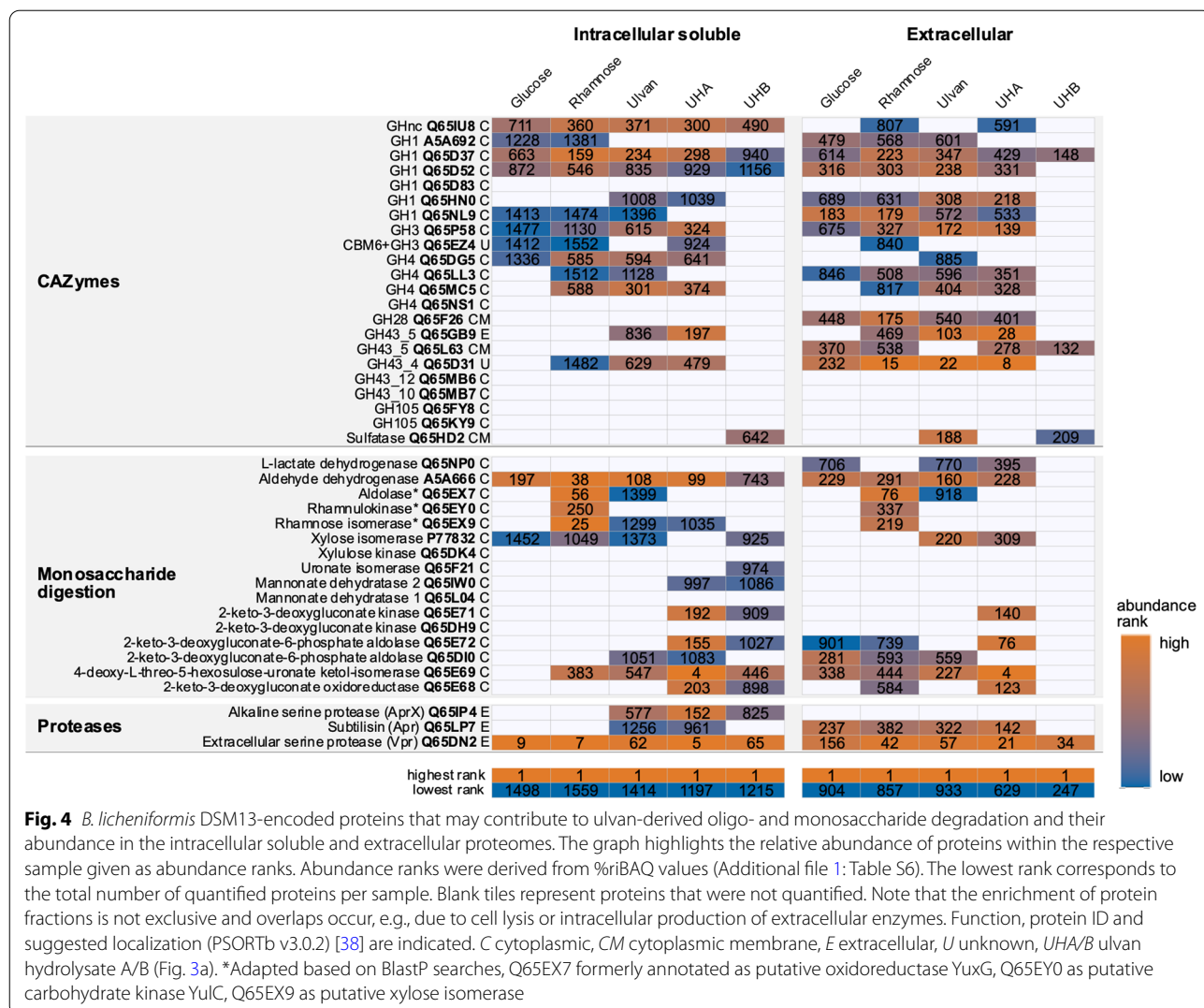


Fig. 4 *B. licheniformis* DSM13-encoded proteins that may contribute to ulvan-derived oligo- and monosaccharide degradation and their abundance in the intracellular soluble and extracellular proteomes. The graph highlights the relative abundance of proteins within the respective sample given as abundance ranks. Abundance ranks were derived from %iBAQ values (Additional file 1: Table S6). The lowest rank corresponds to the total number of quantified proteins per sample. Blank tiles represent proteins that were not quantified. Note that the enrichment of protein fractions is not exclusive and overlaps occur, e.g., due to cell lysis or intracellular production of extracellular enzymes. Function, protein ID and suggested localization (PSORTb v3.0.2) [38] are indicated. C cytoplasmic, CM cytoplasmic membrane, E extracellular, U unknown, UHA/B ulvan hydrolysate A/B (Fig. 3a). *Adapted based on BlastP searches, Q65EX7 formerly annotated as putative oxidoreductase YuxG, Q65EY0 as putative carbohydrate kinase YulC, Q65EX9 as putative xylose isomerase

marine ulvan targeting strains encode a set of sulfatases [14, 42–44], e.g., eight sulfatases from five S1 subfamilies are encoded in the *F. agariphila* ulvan PUL. In *B. licheniformis*, only three proteins are annotated as putative sulfatases. However, for two of them, YfnI (Q65D92) and YflE (Q62XX8), it has been discovered that they are involved in cell wall lipoteichoic acid synthesis in *B. subtilis* (Additional file 1: Fig. S6) [45]. Indeed, the remaining sulfatase (Q65HD2) was abundant in ulvan and UHB secretomes (Fig. 4). We therefore cloned and overexpressed the respective gene in *E. coli*, but so far, no specific sulfatase activity of this enzyme could be detected (data not shown). At the same time, since UHB hydrolysate provides desulfated monosaccharides, its role in desulfation needs to be investigated in more detail in future studies. Nevertheless, the results underline that sulfatases are largely underexplored in *B. licheniformis* and might not even be recognized as such, e.g.

alkaline phosphatases preferentially cleave phosphate monoesters, but are also active on the sulfate counterparts [46]. In another scenario, *B. licheniformis* DSM13 could just consume desulfated ulvan fragments.

Consumption of ulvan-derived monosaccharides

In case of UHA the hydrolysate does not only contain oligosaccharides, but also free unsaturated uronic acids as substrates, which was demonstrated to be consumed by our growth experiments (Additional file 1: Fig. S3). Confirming this, a 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase (Q65E69) was among the most abundant proteins in UHA samples (Fig. 4) representing 1.6% and 4.5% of the total UHA intracellular and extracellular proteome, respectively (Additional file 2: Table S6, Additional file 3: Table S7). Pathways for the other monosaccharides could also be mapped in UHA and UHB samples (Fig. 4), although they were not fully covered.

Taking multiple samples over time and comparing them to respective monosaccharide cultures could close these gaps. Monosaccharides are probably consumed successively, as glucose and xylose are not degraded simultaneously in *Bacillus* species [47, 48].

Protease expression during growth on ulvan and oligosaccharides

In addition, the detected significantly increased protease activities of *B. licheniformis* cultivations with ulvan hydrolysates (see Fig. 2) indicated an elevated protease expression under these conditions. This was supported by our proteome analyses, where the alkaline serine protease (AprX, Q65IP4) was only quantified in ulvan and ulvan hydrolysate samples. However, the subtilisin protease Apr (Q65LP7) or the extracellular serine protease Vpr (Q65DN2) were present throughout all conditions, with high levels of Vpr (Fig. 4). Whereas putative algal-derived proteins from extraction were probably negligible (ulvan, UHA and UHB samples) as inducers of these enzymatic activities, the added *F. agariphila* enzyme extracts to generate ulvan hydrolysates injected additional protein sources into our samples. However, it is worth emphasizing that this potential nutrient source did not cause a significant biomass increase in our control growth experiments (Fig. 2a, enzyme control). The observed increased protease activities thus underline the suitability of ulvan and hydrolysates thereof as potential substrates for industrial bulk protease production processes.

Functional expression of two initial ulvan-degrading CAZymes in *Bacillus*

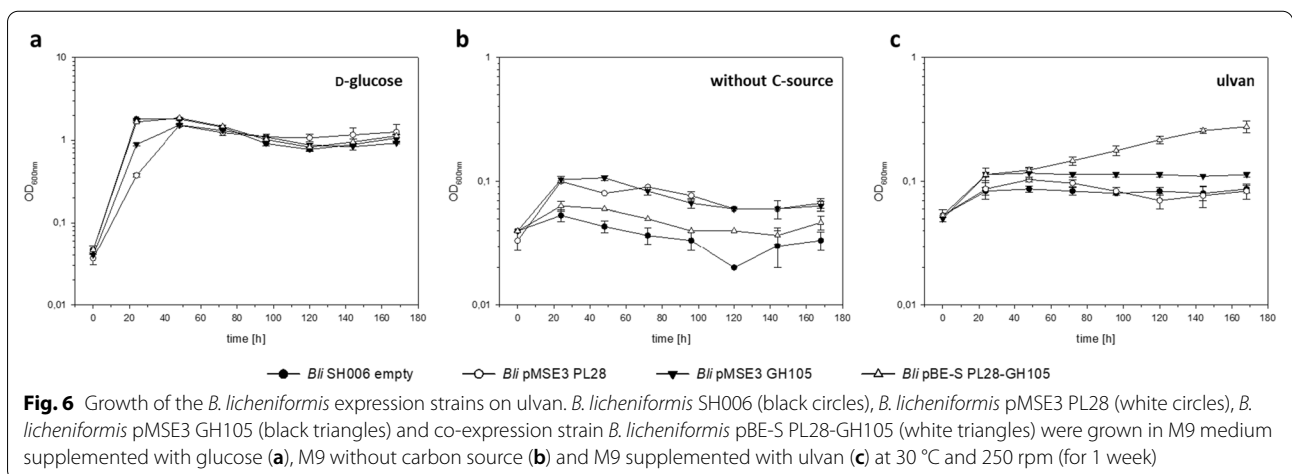
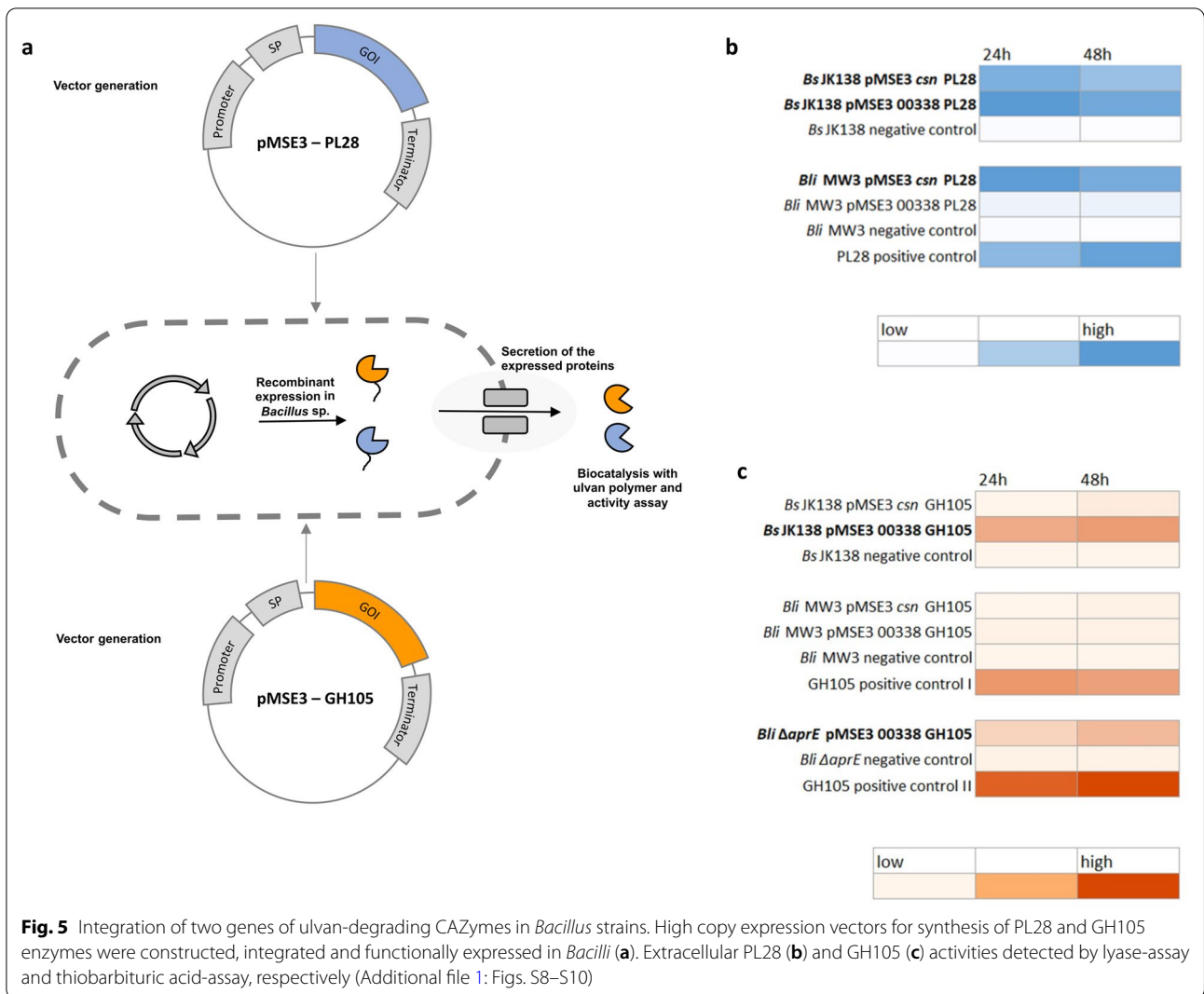
Our previous experiments have shown that *B. licheniformis* DSM13 lacks two initial enzyme activities, ulvan lyase (PL28) and unsaturated glucuronyl hydrolase (GH105, GH88), to use the ulvan polymer as sole carbon source. Therefore, we integrated the *F. agariphila* P30_PL28 and P33_GH105 into a *Bacillus* host-vector system, expressing them as secreted proteins to disassemble ulvan and thus enabling a self-sufficient *Bacillus* strain (Fig. 5a). Since *B. subtilis* and *B. licheniformis* share a similar CAZyme repertoire [35, 49] and showed similar growth in comparative experiments (Additional file 1: Fig. S7), *B. subtilis* JK138 and *B. licheniformis* MW3 were selected as first expression hosts. *B. licheniformis* MW3, a derivative of *B. licheniformis* DSM13, lacks the RM-system (restriction and modification system) which facilitates the genetic accessibility of this strain [50]. Starting with the signal peptide *csn* from *B. subtilis* and 00338 from *B. licheniformis* for both hosts, proteins were expressed in an active form extra- and intracellularly after growth in EnpressoB under simulated fed-batch

conditions (Additional file 1: Fig. S9). PL28 synthesis and activity was confirmed for both expression hosts by ulvan lyase assays (Fig. 5b) and C-PAGE analysis (Additional file 1: Figs. S9, S10).

Production of the GH105 enzyme was detected in *B. subtilis* in combination with the 00338-secretion signal but not in *B. licheniformis* MW3, which was most probably caused by protease activity in this strain. Thus, we additionally used another *B. licheniformis* expression strain (*B. licheniformis* SH006), which is similar to *B. subtilis* JK138 being deficient in the main extracellular protease Apr. Indeed, functional expression of GH105 was detected in the *B. licheniformis* protease-mutant strain using the same expression cassette as for *B. subtilis* (Fig. 5b). Corresponding carbohydrate polyacrylamide gel electrophoresis (C-PAGE) analysis for GH105 activity is shown in Additional file 1: Fig. S10. Although enzyme activity was higher within the extracellular fraction, intracellular enzyme activities for PL28 and GH105 were measured in the activity assays (data not shown) and were also detected by C-PAGE (Additional file 1: Fig. S10), indicating an incomplete protein secretion. To further improve protein secretion of PL28 and GH105, all Sec dependent signal peptides of *B. subtilis* were screened to enhance protein secretion. For this purpose, the *B. subtilis* Secretory Protein Expression System (Takara Clontech) was used, which allows the fusion of 173 Sec-dependent signal peptides of *B. subtilis* to the genes of interest. Based on *B. subtilis* JK138, it could be demonstrated, that higher enzyme activities for PL28 were measured when protein secretion was mediated by the secretion signal of *wprA*, whereas none of the investigated signal peptides mediated an increase in GH105 activity (data not shown). Taken together, we were able to establish the functional expression of the two initial ulvan degrading enzymes PL28 and GH105 in *B. subtilis* and *B. licheniformis*, which may enable both organisms for applications in bioprocess development based on the alternative biomass ulvan.

Co-expression of the PL28 and GH105 in *B. licheniformis* SH006

As soon as the functional expression of either the PL28 or the GH105 encoding gene was established, a self-sufficient strain was designed by combining both enzyme genes. In order to compare PL28 or GH105 single expression vs. co-expression of both marine enzymes in *B. licheniformis* SH006, growth experiments for protein expression (Additional file 1: Fig. S11) and ulvan utilization (Fig. 6) were carried out simultaneously for the *B. licheniformis* “empty” strain, serving as the negative control, for *B. licheniformis* pMSE3 P_{aprE} *csn*-UL, *B. licheniformis* pMSE3 P_{aprE} 00,338-GH and



B. licheniformis pBE-S PL28-GH105. Determination of enzymatic activities revealed functional expression of the PL28 enzyme in the PL28 single expression strain *B. licheniformis* pMSE3 PL28 and the co-expression strain *B. licheniformis* pBE-S PL28-GH105 (Additional file 1: Fig. S12a) whereas GH105 activity was detected in the *B. licheniformis* pMSE3 GH105 and *B. licheniformis* pBE-S PL28-GH105 strains for both investigated time points (Additional file 1: Fig. S12b). As illustrated in Additional file 1: Fig. S12a, the measured PL28 activities in *B. licheniformis* pBE-S PL28-GH105 were very low after 24 h and even in a negative range after 48 h of expression. However, this represents a strong hint for co-expression of both enzymes: the unsaturated uronic acid formed by the PL28 led to an increased absorption ($A_{235\text{nm}}$), which was then reversed by the GH105 that cleaved this moiety. Corresponding C-PAGE analysis for PL28 and GH105 showed activities in all analyzed strains as shown in Additional file 1: Fig. S13. Taken together, the data of our protein expression experiments under simulated fed-batch conditions in EnpressoB medium clearly demonstrated that both marine enzymes were actively co-expressed.

In a final experiment, we thus wished to demonstrate, that *B. licheniformis*, equipped with the pBE-S PL28-GH105 co-expression vector, is able to grow on ulvan as the sole carbon source. Therefore, the same four expression strains were grown for 7 days in: (i) M9 mineral medium without carbon source, serving as the negative control; (ii) M9 mineral medium with 0.4% D-glucose, serving as the positive control; and (iii) with 1% ulvan as the sole carbon source (Fig. 6). While the single PL28 and GH105 strains lack the ability to grow on the ulvan, the co-expression strain shows an increased growth over this period of time. A C-PAGE of the culture's supernatant showed ulvan hydrolysis (Additional file 1: Fig. S14) in the PL28 and co-expression cultivation. This demonstrated PL28 and GH105 expression and activity in the M9-mineral medium supplemented with ulvan. Additionally, this experiment confirmed that the *B. licheniformis* strain needs PL28 and GH105 to grow on ulvan biomass.

Conclusion

This study reveals the promising metabolic potential of the bacterial cell factory *B. licheniformis* to utilize the abundant and renewable marine algal polysaccharide ulvan. We demonstrated that the native *B. licheniformis* DSM13 strain can grow on ulvan-derived oligo- and monosaccharides obtained by enzymatic pre-hydrolysis. Our proteogenomic analyses indicated that *B. licheniformis* DSM13 lacks the initial ulvan degradation enzymes, but that the pre-digestion of this marine

polysaccharide with two particular enzymes suffices to generate a suitable carbon source for this bacterium. We demonstrate that a metabolically engineered *B. licheniformis* strain, equipped with two marine heterologously expressed marine enzymes for the initial breakdown of the algal polysaccharide ulvan, is able to grow on ulvan as the sole carbon and energy source. Thus, this study pinpoints a suitable metabolic engineering strategy for future strain development aiming for a cell factory for the conversion of the abundant marine polysaccharide ulvan as alternative feedstock in large-scale bioprocesses.

Methods

Genes and enzyme expression

We used the already available pET28a(+) based expression constructs, coding for the *Formosa agariphila* KMM3901^T (collection number DSM15362 at DSMZ, Braunschweig, Germany) specific ulvan enzyme cascade [14]. Enzyme overproduction in *E. coli* was performed as described previously [14]. After cell lysis the soluble fraction was filtered (0.45 μm) and the resulting crude extract was aliquoted and shock frozen. The enzyme expression was confirmed via SDS-PAGE (Additional file 1: Fig. S2).

Extraction of ulvan

Dried *Ulva* biomass from the Atlantic coast in Spain was purchased as organic sea lettuce (Kulau, Berlin, Germany). Ulvan was extracted as described before, but distilled water was used as extraction solvent [14].

Enzyme assays

The thiobarbituric acid assay [51] adapted for reduced volumes detected 5-dehydro-4-deoxy-D-glucuronate in the culture supernatant. The protease/peptidase activity in the culture supernatant was determined via the AAPF-assay, through the release of *p*-nitroanilin (410 nm, $E^M=9800$) from the substrate *N*-succinyl-Ala-Ala-Pro-Phe-*para*-nitroanilide (succinyl-AAPF-*p*-Nitroanilide). The enzyme activity was calculated from the amount of *p*-nitroanilin released per time [33].

Strains

Ten microorganisms were selected for growth screening on ulvan extracts or hydrolysate: *Escherichia coli* Top10, *E. coli* BL21(DE3), *Bacillus subtilis* B1, *Saccharomyces cerevisiae* GRF18, *Vibrio natriegens* ATCC 14048, *Pseudomonas putida* DSMZ 50198, *Pichia pastoris* X33, *Bacillus licheniformis* DSM13, *Cupriavidus necator* H16 and *Cutaneotrichosporan curvatus* DSM 101032 (Additional file 1: Table S2). All strains were maintained as glycerol stocks, stored at $-80\text{ }^\circ\text{C}$. The Bacilli strains *Bacillus subtilis* JK138, *Bacillus licheniformis* MW3 and *Bacillus licheniformis* SH006 were used for all

expression experiments in this study. *E. coli* DH10B (Invitrogen, Darmstadt, Germany) [F-endA1 recA1 galE15 galK16 nupG rpsL Δ lacX74 Φ 80lacZ Δ M15 araD139 Δ (ara,leu)7697 mcrA Δ (mrr-hsdRMS-mcrBC) λ -] was used as the host strain for all subcloning procedures. *Bacillus licheniformis* DSM13 mutant SH006 was constructed with a homologous recombination method using a pE194-derived shuttle vector pE194SV analogous to the pMAD system [52]. pE194SV consist of the temperature-sensitive ori and erythromycin resistance marker gene from pE194ts [53] cloned into the *Sma*I site of pUC18, in which the native *Bsa*I site was removed. Moreover, pE194SV carries a type-II-assembly mRFP cassette from pBSd141R [54]; GenBank accession number: KY995200) integrated into the *Bam*HI site of pUC18. The pE194SV based gene deletion procedure was conducted according to Nahrstedt et al. 2005, using 45 °C instead of 42 °C as non-permissive temperature [55].

For the deletion of the restriction endonuclease (*hsdR1*) within the restriction modification operon 1 and the adjacent *mcrA* gene, 5'- and 3'-homologous flanking regions were PCR amplified from DSM13 genomic DNA. The 5'-flanking region was amplified using primers P1-hsdR1 and P2-hsdR1 and the 3'-flanking region was amplified with primers P3-hsdR1 and P4-hsdR1. Primers P1-hsdR1 and P4-hsdR1 introduced *Bsa*I cut sites and unique overhangs for subsequent cloning. Both fragments were ligated by SOE-PCR [56] and cloned via *Bsa*I into pE194SV, resulting in plasmid pDhsdR1.

For the deletion of the restriction endonuclease (*hsdR2*) within the restriction modification operon 2, 5'- and 3'-homologous flanking regions were PCR amplified from DSM13 genomic DNA. The 5'-flanking region was amplified using primers P1-hsdR2 and P2-hsdR2 and the 3' flanking region was amplified with primers P3-hsdR2 and P4-hsdR2. Primers P1-hsdR2 and P4-hsdR2 introduced *Bsm*BI cut sites and unique overhangs for subsequent cloning. Both fragments were then ligated by SOE-PCR, digested with *Bsm*BI and cloned into the *Bsa*I digested pE194SV, resulting in plasmid pDhsdR2.

For the deletion of the poly- γ -glutamic acid (*pga*) synthesis operon (*pgsBCAE*) and the *apr* gene encoding an extracellular alkaline serine protease individual cassettes comprising the 5'- and 3'-homologous flanking regions, flanked by *Bsa*I cut sites and unique overhangs were ordered as synthetic fragments. Each cassette was cloned separately via *Bsa*I into pE194SV, resulting in plasmids pDpga (Δ *pga*) and pDapr (Δ *apr*).

Preparation of a sugar rich hydrolysate

2 mg/mL ulvan in phosphate buffer (25 mM, 50 mM NaCl, pH 7.5) was incubated with 0.5% (v/v) of the respective *F. agariphila* crude enzyme (Additional file 1:

Table S1) overnight. The ulvan hydrolysates were centrifuged for 5 min at 4500 \times g and were then filtered (0.2 μ m).

Monosaccharide composition analysis of the cultivation media

The ulvan, UHA and UHB raw media (Additional file 1: Fig. S1) were chemically hydrolysed (1 M HCl for 24 h at 100 °C). Afterwards, the samples were filtered (0.2 μ m Spin-X filter) prior to HPAEC-PAD analyses using a Dionex CarboPac PA10 column (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and monosaccharide mixtures as standards for column calibration [57].

Cultivation of different strains with various carbon sources

The M9-mineral medium with 0.2% (w/v) yeast extract was supplemented with various sugar sources (Additional file 1: Tables S3, S4). Selected monosaccharides were D-glucose, L-rhamnose, D-xylose and D-glucuronic acid, each at a final concentration of 0.4%(w/v). A final concentration of 1%(w/v) was used in case of ulvan and ulvan hydrolysates. M9 and M9 supplemented with enzyme mixtures from the preparation of the sugar rich hydrolysates were used as controls. Precultures were prepared in respective rich media for the corresponding microorganisms (LB media, YPD for yeasts) and overnight (for yeasts 1.5 days). This preculture was used to inoculate a second preculture in M9-mineral media with 0.2%(g/L) glucose as carbon source (1:100). The main culture was inoculated (1:100) with the M9-mineral media preculture and cultured up to 4 days at 30 °C and 180 rpm. The optical density was measured at 600 nm.

Proteome analyses

For proteome analyses, late logarithmic phase cells from triplicates of ulvan, UHA, UHB, L-rhamnose and D-glucose cultures were separated from supernatants by centrifugation (20 min, 4000 \times g, 4 °C). Intracellular soluble proteins were extracted by suspending cell pellets in lysis buffer (4% SDS, 1% NaDCA, 50 mM TEAB) adapted to Hinzke et al. [58]. Samples were incubated for 5 min at 600 rpm and 95 °C, then cooled on ice shortly and sonicated for 5 min. Cell debris was removed from the protein extract (intracellular soluble proteome) by centrifugation (10 min, 14,000 \times g, room temperature). Protein concentration was determined using the PierceTM BCA Protein Assay Kit (Thermo Fisher Scientific, Waltham, Massachusetts, US). Secreted and detached proteins (extracellular proteome) were extracted from cultivation supernatants using StrataClean beads (Agilent, Santa Clara, California, US) [59]. In brief, 20 μ L of bead solution, extracting approximately 20 – 30 μ g of protein, was removed. Beads were primed in 180 μ L 37% hydrochloric acid

(100 °C, 6 h) and then washed in TE buffer (50 mM Tris, 10 mM EDTA, pH 8.0) twice (5 min, 3500×g, room temperature). 0.2 µm-filtered supernatants were incubated with prepared beads overnight in a 360° rotating shaker at 8 rpm and 4 °C. The protein-loaded beads were pelleted by centrifugation (45 min, 10,000×g, 4 °C) and washed in TE buffer. In a last step, they were resuspended in 1 mL of ultrapure water and dried by vacuum centrifugation. 25 µg of protein from intracellular soluble protein extracts as well as protein-loaded beads were separated by 1D SDS PAGE (12% SDS gels) at 120 V. Proteins were in-gel digested using trypsin [59]. Peptides were separated by reversed phase chromatography and analyzed in an LTQ-Orbitrap Classic mass spectrometer equipped with a nanoelectrospray ion source [60]. MS/MS spectra were searched against a target decoy database using MaxQuant v. 1.6.10.43 [61]. The database covered all protein sequences predicted from the *B. licheniformis* DSM13 genome, selected *F. agariphila* KM3901^T ulvan PUL-encoded enzymes (Additional file 1: Table S1) and common laboratory contaminants as well as corresponding reversed sequences (decoys). The MaxQuant computed iBAQ values (intensity-based absolute quantification [62]) were used to manually calculate %riBAQ values, giving the relative protein abundance in % per sample. Quantified *F. agariphila* KM3901^T proteins were excluded from %riBAQ calculations. Proteins quantified in at least two out of the three replicates were considered for further calculations and for statistical tests. Since the total number of quantified proteins varied considerably between substrates (e.g., 933 proteins in ulvan extracellular samples compared to 247 proteins in the UHB samples), %riBAQ mean values were ranked according to their abundance (e.g., rank 1 for most abundant protein in the sample) to increase comparability between conditions. In addition, the total number of quantified proteins per sample was considered for the color code in graphs. Welch's two-sided t-test (permutation-based FDR 0.05) identified statistical significance to protein abundance differences between samples within the intracellular soluble proteome samples and within the extracellular samples using Perseus v. 1.6.0.7 [63]. Only samples with a similar number of quantified proteins were compared. CAZymes were identified using dbCAN2 [37]. The enrichment of protein fractions is not exclusive and overlaps may occur, e.g., due to cell lysis or intracellular production of extracellular enzymes. Therefore, protein localization was also predicted using PSORTb v3.0.2 [38]. Proteomic data were deposited to the ProteomeXchange Consortium via the PRIDE partner repository [64] with the dataset

identifier PXD033411. CAZymes were identified using dbCAN2 [37].

Development of a *Bacillus* host-vector system

The nucleotide sequence of both genes from *F. agariphila* KMM3901^T P30_PL28 and P33_GH105 were ordered by GenScript Biotech (Leiden, Netherlands). Both synthetic genes were codon-optimized for expression in *B. licheniformis* using the GenSmartTM Codon Optimization tool (GenScript). The algorithm utilizes a matrix for the most frequently occurring codons in *B. licheniformis*. The constructs were assembled from synthetic oligonucleotides and provided in the backbone of the pUC19 vector. Amplification of the PL28 nucleotide sequence with *csn* and 00338 signal peptides (SP) was carried out in two discrete polymerase chain reactions, first with oligonucleotides MaZu8 and MaZu9 and second with MaZu7 and MaZu9 for the *csn*-SP whereas fusion of the 00338-SP was carried out with MaZu13/MaZu9 and MaZu12/MaZu9. The nucleotide sequence of GH105 with *csn*-SP was amplified with oligonucleotides MaZu10 and MaZu11 in a first PCR and, using the purified PCR product as the template, with MaZu7 and MaZu11 in a second PCR. Fusion of the 00338-SP to the GH105 sequence occurred with oligonucleotides MaZu14/MaZu11 and MaZu12/MaZu11. The PCR products were digested with *NdeI* and *KpnI* and subsequently gel-purified. After ligation into the *NdeI* and *KpnI* sites of pMSE3 *P_{appr}* *E. coli* DH10B was transformed with the recombinant plasmids, yielding pMSE3 *P_{appr}* *csn*-UL, pMSE3 *P_{appr}* 00338-UL, pMSE3 *P_{appr}* *csn*-GH and pMSE3 *P_{appr}* 00338-GH. Sequence identity of all expression vectors was verified by sequencing (Eurofins Genomics, Ebersberg, Germany). All four expression vectors were then transferred into both *Bacillus* expression hosts, *B. subtilis* JK138 and *B. licheniformis* MW3 [50] by electroporation.

Protein expression experiments were performed under simulated fed-batch conditions in EnpressoB-Medium as recommended by the manufacturer (Biosilta) at 30 °C and 250 rpm. Samples for protein analysis (SDS-PAGE, Activity screening and carbohydrate electrophoresis [15]) were taken after 24 h and 48 h of cultivation.

Development of a co-expression host *B. licheniformis* SH006

Construction of the appropriate expression vector was achieved by Gibson assembly. To this purpose, PCRs of the PL28 and GH105 expression cassettes were carried out with oligonucleotides MaZu37/MaZu29 (PL28) and MaZu28/MaZu35 (GH105) using pMSE3 *P_{apprE}* *csn*-UL and pMSE3 *P_{apprE}* 00338-GH as the templates. The vector backbone of the medium copy-vector pBE-S was amplified with oligonucleotides MaZu38/MaZu36. The PCR

products were gel-purified and, in case of the pBE-S vector backbone, digested with *DpnI* in order to remove remaining circular plasmid DNA. All purified DNA fragments were then assembled in a vector:insert ratio of 1:2 and 3 μL of the reaction were used for transformation of *E. coli* DH10B yielding pBE-S P_{aprE} *csn*-UL - P_{aprE} 00338-GH (pBE-S PL28-GH105). Sequence identity of the PL28-GH105 co-expression vector was verified and the plasmid was subsequently integrated into *B. licheniformis* SH006 by electroporation. Protein expression experiments of the newly constructed *B. licheniformis* SH006 PL28-GH105 co-expression strain were performed as before for the single constructs. Additionally, in order to demonstrate the ability of the newly constructed *B. licheniformis* SH006 PL28-GH105 strain to grow on ulvan, cultivations in M9-mineral media supplemented with either D-glucose or ulvan as the sole carbon source and also without any carbon source were carried out as described before.

Activity measurement of ulvan lyase (PL28) and glycoside hydrolase (GH105)

The ulvan lyase activity was detected as described before [15] using the intra- or extracellular extract of *Bacillus* sp. cultivations instead of purified protein. For the detection of the glycoside hydrolase (GH105) activity the reversed ulvan lyase assay was used, while ulvan was PL28-pre-hydrolysed and heat inactivated after 16 h. The breakdown products resulting from the ulvan lyase assays were additionally analyzed via C-PAGE, the MBTH- assay and the thiobarbituric acid assay as described before [15].

Construction of the plasmid libraries for PL28 and GH105 activity screening

In order to obtain both plasmid libraries with 173 different types of signal peptides DNA sequences in the required size of at least 2000 *E. coli* clones, the “*B. subtilis* Secretory Protein Expression System” (Takara/Clontech) in combination with the “In Fusion HD Cloning Plus Kit” (Takara/Clontech) was used according to the manufacturer’s instructions. To this end, the nucleotide sequences for PL28 and GH105 were amplified from pMSE3-PL28 and pMSE3-GH105 using oligonucleotides MaZu19/MaZu20 (for PL28) and MaZu21/MaZu22 (for GH105). After restriction with *NdeI* and *XbaI*, the purified PCR products and the pBE-S vector were ligated and *E. coli* DH10B was transformed with the recombinant plasmids pBE-S-PL28 and pBE-S-GH105. After validation of sequence identity for PL28 and GH105, all different signal peptide sequences included in the provided SP library were integrated into the vector backbones of pBE-S-PL28 and pBE-S-GH105 following the manufacturer’s instructions. In brief, the *EagI* and *MluI* digested vector was ligated with the 173 SP-containing DNA mixture using

the “In Fusion Cloning” technology. Chemically competent *E. coli* Stellar cells (included in the kit) were transformed with 2 μL of the “In Fusion” reaction and selected on LB agar plates with ampicillin. All colony forming units (cfu) were rinsed from the plate to isolate the SP-plasmid library, which was subsequently integrated into the *Bacillus* expression hosts by electroporation.

Abbreviations

CAZyme: Carbohydrate active enzyme; GH: Glycoside hydrolase; UL: Ulvan lyase; U: Ulvan; UH: Ulvan hydrolysate; UHA: Ulvan hydrolysate resulting from the digestion of ulvan with the enzymes P30_PL28 and P33_GH105; UHB: Ulvan hydrolysate resulting from the digestion of ulvan with the whole ulvan specific enzyme cascade according to Reisky et al. [14].

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12934-022-01931-0>.

Additional file 1: Table S1. Proteins and accession numbers. **Table S2.** Bacterial strains. **Table S3.** M9-mineral media D-glucose composition. **Table S4.** M9-mineral media additives. **Table S5.** Primer list. **Figure S1.** Sugar composition of the cultivation media. **Figure S2.** SDS-PAGE of *F. agariphila* KMM3901^T enzymes expressed recombinantly in *E. coli*. **Figure S3.** Consumption of 5-dehydro-4-deoxy-D-glucuronate from *B. licheniformis* DSM13 during cultivation. **Figure S4.** Growth of *B. licheniformis* DSM13 on different ulvan hydrolysates. **Figure S5.** *B. licheniformis* DSM13 CAZyme repertoire and their expression. **Figure S6.** Alignment *B. licheniformis* DSM13 sulfatases with lipoteichoic acid synthases. **Figure S7.** Comparison of different *Bacillus* sp. to digest ulvan hydrolysate and ulvan derived monosaccharides. **Figure S8.** Growth curves of PL28 and GH105 *Bacillus* sp. expression strains. **Figure S9.** Activity assays from *Bacillus* sp. PL28 and GH105 expression strains. **Figure S10.** C-PAGE results from *Bacillus* sp. PL28 and GH105 expression strains. **Figure S11.** Growth of the different *B. licheniformis* expression strains. **Figure S12.** Activity assay results from *B. licheniformis* SH006, PL28 and GH105 single- and co-expression strain. **Figure S13.** C-PAGE results from *B. licheniformis* SH006 PL28, GH105 and co-expression strains. **Figure S14.** C-PAGE from the cultivation supernatant of *B. licheniformis* strains in M9-mineral media.

Additional file 2: Table S6. Summary of the proteomic results.

Additional file 3: Table S7. Results of statistical analyses.

Acknowledgements

We thank Alek Bolte for the help in the monosaccharide composition analysis and Sebastian Grund for the mass spectroscopy measurements. In addition, we thank our Bachelor student Maik Behrens for her supporting work.

Author contributions

UTB, LR and TS designed the study, supervised its execution and co-wrote the manuscript. TD wrote the main manuscript with the support of MKZ and NW. TD performed the initial screening of organism and growth and activity assays, with TE and CS. The proteome analysis was performed by TE and MKZ. DB coordinated MS measurements. The computational analysis was performed by MKZ, AD and SB. JK performed the sugar monosaccharide composition analysis in the lab of JHH. SB expressed and purified the putative sulfatases and performed the activity assays. NW developed the *Bacillus* host-vector system and the strain *Bacillus* design with MH. All authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. This study was supported by the German Federal Ministry of Education and Research (BMBF) in the frame of the Plant³ MarZucker project (03WIR2205A/C) and the German Research Foundation (DFG) for the Research Unit FOR 2406 “Proteogenomics

of Marine Polysaccharide Utilization" (POMPU) by Grants to U.T.B. (BO 1862/17-2), J.-H.H. (HE 7217/2-2), and T.S. (SCHW 595/10-2).

Availability of data and materials

Proteomic data were deposited to the ProteomeXchange Consortium via the PRIDE partner repository [64] with the dataset identifier PXD033411.

Declarations

Consent for publication

All authors have read and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Biotechnology & Enzyme Catalysis, Institute of Biochemistry, University of Greifswald, 17487 Greifswald, Germany. ²Department of Pharmaceutical Biotechnology, Institute of Pharmacy, University of Greifswald, 17487 Greifswald, Germany. ³Institute of Marine Biotechnology e.V., 17489 Greifswald, Germany. ⁴Max Planck-Institute for Marine Microbiology, 28359 Bremen, Germany. ⁵Center for Marine Environmental Sciences (MARUM), University of Bremen, 28359 Bremen, Germany. ⁶Department of Microbial Proteomics, Institute for Microbiology, University of Greifswald, 17487 Greifswald, Germany.

Received: 28 June 2022 Accepted: 26 September 2022

Published online: 10 October 2022

References

- Wang M, Hu C, Barnes BB, Mitchum G, Lapointe B, Montoya JP. The great Atlantic Sargassum belt. *Science*. 2019;365:83–7.
- Smetacek V, Zingone A. Green and golden seaweed tides on the rise. *Nature*. 2013;504:84–8.
- Field CB. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*. 1998;281:237–40.
- Hehemann J-H, Boraston AB, Czjzek M. A sweet new wave: structures and mechanisms of enzymes that digest polysaccharides from marine algae. *Curr Opin Struct Biol*. 2014;28:77–86.
- Filote C, Santos SCR, Popa VI, Botelho CMS, Volf I. Biorefinery of marine macroalgae into high-tech bioproducts: a review. *Environ Chem Lett*. 2020;19:969–1000.
- Laurens LML, Lane M, Nelson RS. Sustainable seaweed biotechnology solutions for carbon capture, composition, and deconstruction. *Trends Biotechnol*. 2020;38:1232–44.
- Cesário MT, da Fonseca MMR, Marques MM, de Almeida MCMD. Marine algal carbohydrates as carbon sources for the production of biochemicals and biomaterials. *Biotechnol Adv*. 2018;36:798–817.
- Kim HT, Lee S, Kim KH, Choi I-G. The complete enzymatic saccharification of agarose and its application to simultaneous saccharification and fermentation of agarose for ethanol production. *Bioresour Technol*. 2012;107:301–6.
- Murata M, Nakazoe J. Production and use of marine algae in Japan. *Jpn Agric Res Q JARQ*. 2001;35:281–90.
- Lahaye M, Gomez-Pinchetti J-L, del Rio MJ, Garcia-Reina G. Natural decoloration, composition and increase in dietary fibre content of an edible marine algae, *Ulva rigida* (Chlorophyta), grown under different nitrogen conditions. *J Sci Food Agric*. 1995;68:99–104.
- Kloareg B, Quatrano RS. Structure of the cell walls of marine algae and ecophysiological functions of the matrix polysaccharides. *Oceanogr Mar Biol Annu Rev*. 1988;26:259–315.
- Bäumgen M, Dutschei T, Bornscheuer UT. Marine polysaccharides: occurrence, enzymatic degradation and utilization. *ChemBioChem*. 2021;22:2247–56.
- Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The carbohydrate-active enzymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res*. 2009;37:D233–8.
- Reisky L, Préchoux A, Zühlke M-K, Bäumgen M, Robb CS, Gerlach N, et al. A marine bacterial enzymatic cascade degrades the algal polysaccharide ulvan. *Nat Chem Biol*. 2019;15:803–12.
- Bäumgen M, Dutschei T, Bartosik D, Suster C, Reisky L, Gerlach N, et al. A new carbohydrate-active oligosaccharide dehydratase is involved in the degradation of ulvan. *J Biol Chem*. 2021. <https://doi.org/10.1016/j.jbc.2021.101210>.
- Cardoso SM, Carvalho LG, Silva PJ, Rodrigues MS, Pereira ORP, Pereira L. Bioproducts from seaweeds: a review with special focus on the Iberian Peninsula. *Curr Org Chem*. 2014. <https://doi.org/10.2174/138527281807140515154116>
- Lahaye M, Robic A. Structure and functional properties of ulvan, a polysaccharide from green seaweeds. *Biomacromol*. 2007;8:1765–74.
- Reisky L, Stanetty C, Mihovilovic MD, Schweder T, Hehemann J-H, Bornscheuer UT. Biochemical characterization of an ulvan lyase from the marine flavobacterium *Formosa agariphila* KMM 3901T. *Appl Microbiol Biotechnol*. 2018;102:6987–96.
- Jose ML, Rosario C, Jon AA, Jesus L. Immunomodulating activities of acidic sulphated polysaccharides obtained from the seaweed *Ulva rigida*. *Int Immunopharmacol*. 2007;7:879–88.
- Wargacki AJ, Leonard E, Win MN, Regitsky DD, Santos CNS, Kim PB, et al. An engineered microbial platform for direct biofuel production from brown macroalgae. *Science*. 2012;335:308–13.
- Stephanopoulos G. Challenges in engineering microbes for biofuels production. *Science*. 2007;315:801–4.
- Lü C, Ge Y, Cao M, Guo X, Liu P, Gao C, et al. Metabolic engineering of *Bacillus licheniformis* for production of acetoin. *Front Bioeng Biotechnol*. 2020;8:125.
- Wells JA, Estell DA. Subtilisin—an enzyme designed to be engineered. *Trends Biochem Sci*. 1988;13:291–7.
- Huebner U, Bock U, Schuegerl K. Production of alkaline serine protease subtilisin Carlsberg by *Bacillus licheniformis* on complex medium in a stirred tank reactor. *Appl Microbiol Biotechnol*. 1993;40:182–8.
- Song CW, Chelladurai R, Park JM, Song H. Engineering a newly isolated *Bacillus licheniformis* strain for the production of (2R,3R)-butanediol. *J Ind Microbiol Biotechnol*. 2020;47:97–108.
- Nilegaonkar SS, Bhosale SB, Dandage CN, Kapadi AH. Potential of *Bacillus licheniformis* for the production of 2,3-butanediol. *J Ferment Bioeng*. 1996;82:408–10.
- Sakai K, Yamanami T. Thermotolerant *Bacillus licheniformis* TY7 produces optically active L-lactic acid from kitchen refuse under open condition. *J Biosci Bioeng*. 2006;102:132–4.
- Pervez S, Shahid F, Aman A, Qader SAU. Algal biomass: a sustainable, economical and renewable approach for microbial production of pectinolytic enzymes using submerged and solid state fermentation techniques. *Biocatal Biotransformation*. 2017;35:442–9.
- Caspi R, Billington R, Fulcher CA, Keseler IM, Kothari A, Krummenacker M, et al. The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res*. 2018;46:D633–9.
- Wang L, Zhao B, Liu B, Yu B, Ma C, Su F, et al. Efficient production of L-lactic acid from corn cob molasses, a waste by-product in xylitol production, by a newly isolated xylose utilizing *Bacillus* sp. strain. *Bioresour Technol*. 2010;101:7908–15.
- Ramos JL, Duque E, Huertas MJ, Haïdour A. Isolation and expansion of the catabolic potential of a *Pseudomonas putida* strain able to grow in the presence of high concentrations of aromatic hydrocarbons. *J Bacteriol*. 1995;177:3911–6.
- Fonseca Á, Boekhout T, Fell JW. Chapter 138—Cryptococcus Vuillemin (1901). In: Kurtzman CP, Fell JW, Boekhout T, editors. *Yeasts*. 5th ed. London: Elsevier; 2011. p. 1661–737.
- Kannan Y, Koga Y, Inoue Y, Haruki M, Takagi M, Imanaka T, et al. Active subtilisin-like protease from a hyperthermophilic archaeon in a form with a putative prosequence. *Appl Environ Microbiol*. 2001;67:2445–52.
- Voigt B, Schroeter R, Schweder T, Jürgen B, Albrecht D, van Dijl JM, et al. A proteomic view of cell physiology of the industrial workhorse *Bacillus licheniformis*. *J Biotechnol*. 2014;191:139–49.
- Veith B, Herzberg C, Steckel S, Feesche J, Maurer KH, Ehrenreich P, et al. The complete genome sequence of *Bacillus licheniformis* DSM13, an organism with great industrial potential. *J Mol Microbiol Biotechnol*. 2004;7:204–11.

36. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, et al. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 2018;46:W95–101.
37. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 2012;40:W445–51.
38. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics.* 2010;26:1608–15.
39. Silva IR, Jers C, Meyer AS, Mikkelsen JD. Rhamnogalacturonan I modifying enzymes: an update. *New Biotechnol.* 2016;33:41–54.
40. Voigt B, Antelmann H, Albrecht D, Ehrenreich A, Maurer K-H, Evers S, et al. Cell physiology and protein secretion of *Bacillus licheniformis* compared to *Bacillus subtilis*. *Microb Physiol.* 2009;16:53–68.
41. Inácio JM, de Sá-Nogueira I. Characterization of *abn2* (*xyiA*), encoding a *Bacillus subtilis* GH43 arabinanase, *Abn2*, and its role in arabino-polysaccharide degradation. *J Bacteriol.* 2008;190:4272–80.
42. Koch H, Freese HM, Hahnke RL, Simon M, Wietz M. Adaptations of *Alteromonas* sp. 76-1 to polysaccharide degradation: a CAZyme plasmid for ulvan degradation and two alginateolytic systems. *Front Microbiol.* 2019;10:504.
43. Foran E, Buravenkov V, Kopel M, Mizrahi N, Shoshani S, Helbert W, et al. Functional characterization of a novel “ulvan utilization loci” found in *Alteromonas* sp. LOR genome. *Algal Res.* 2017;25:39–46.
44. Salinas A, French CE. The enzymatic ulvan depolymerisation system from the alga-associated marine flavobacterium *Formosa agariphila*. *Algal Res.* 2017;27:335–44.
45. Matsuoka S, Hashimoto M, Kamiya Y, Miyazawa T, Ishikawa K, Hara H, et al. The *Bacillus subtilis* essential gene *dgkB* is dispensable in mutants with defective lipoteichoic acid synthesis. *Genes Genet Syst.* 2011;86:365–76.
46. Andrews LD, Zalatan JG, Herschlag D. Probing the origins of catalytic discrimination between phosphate and sulfate monoester hydrolysis: comparative analysis of alkaline phosphatase and protein tyrosine phosphatases. *Biochemistry.* 2014;53:6811–9.
47. Gärtner D, Geissendörfer M, Hillen W. Expression of the *Bacillus subtilis* *xyl* operon is repressed at the level of transcription and is induced by xylose. *J Bacteriol.* 1988;170:3102–9.
48. Sun J-D, Tang C, Zhou J, Wei P, Wang Y-J, An W, et al. Production of poly- γ -glutamic acid (γ -PGA) from xylose-glucose mixtures by *Bacillus amyloliquefaciens* C1. *3 Biotech.* 2021;11:100.
49. Sharma A, Satyanarayana T. Comparative genomics of *Bacillus* species and its relevance in industrial microbiology. *Genom Insights.* 2013;6:25–36.
50. Waschkau B, Waldeck J, Wieland S, Eichstädt R, Meinhardt F. Generation of readily transformable *Bacillus licheniformis* mutants. *Appl Microbiol Biotechnol.* 2008;78:181–8.
51. Itoh T, Ochiai A, Mikami B, Hashimoto W, Murata K. A novel glycoside hydrolase family 105: the structure of family 105 unsaturated rhamnogalacturonyl hydrolase complexed with a disaccharide in comparison with family 88 enzyme complexed with the disaccharide. *J Mol Biol.* 2006;360:573–85.
52. Arnaud M, Chastanet A, Débarbouillé M. New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Appl Environ Microbiol.* 2004;70:6887–91.
53. Villafane R, Bechhofer DH, Narayanan CS, Dubnau D. Replication control genes of plasmid pE194. *J Bacteriol.* 1987;169:4822–9.
54. Radeck J, Meyer D, Lautenschläger N, Mascher T. *Bacillus SEVA* siblings: a Golden Gate-based toolbox to create personalized integrative vectors for *Bacillus subtilis*. *Sci Rep.* 2017;7:14134.
55. Nahrstedt H, Waldeck J, Gröne M, Eichstädt R, Feesche J, Meinhardt F. Strain development in *Bacillus licheniformis*: construction of biologically contained mutants deficient in sporulation and DNA repair. *J Biotechnol.* 2005;119:245–54.
56. Heckman KL, Pease LR. Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat Protoc.* 2007;2:924–32.
57. Engel A, Händel N. A novel protocol for determining the concentration and composition of sugars in particulate and in high molecular weight dissolved organic matter (HMW-DOM) in seawater. *Mar Chem.* 2011;127:180–91.
58. Hinzke T, Markert S. Efficient protein extraction for proteomics and metaproteomics (also suitable for low biomass samples). *Protoc Io.* 2017;10:6.
59. Bonn F, Bartel J, Büttner K, Hecker M, Otto A, Becher D. Picking vanished proteins from the void: how to collect and ship/share extremely dilute proteins in a reproducible and highly efficient manner. *Anal Chem.* 2014;86:7421–7.
60. Otto A, Bernhardt J, Meyer H, Schaffer M, Herbst F-A, Siebourg J, et al. Systems-wide temporal proteomic profiling in glucose-starved *Bacillus subtilis*. *Nat Commun.* 2010;1:137.
61. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol.* 2008;26:1367–72.
62. Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global quantification of mammalian gene expression control. *Nature.* 2011;473:337–42.
63. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods.* 2016;13:731–40.
64. Vizcaino JA, Csordas A, del Toro N, Dienes JA, Griss J, Lavidas I, et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* 2016;44:D447–56.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



Additional file 1_ supplementary table and figures

Metabolic engineering enables *Bacillus licheniformis* to grow on the marine polysaccharide ulvan

Theresa Dutschei¹, Marie-Katherin Zühlke^{2,3}, Norma Welsch^{2,3}, Tom Eisenack², Maximilian Hilkmann^{2,3}, Joris Krull^{3,4,5}, Carlo Stühle¹, Stefan Brott¹, Alexandra Dürwald², Lukas Reisky¹, Jan-Hendrik Hehemann^{3,4,5}, Dörte Becher⁶, Thomas Schweder^{2,3‡} and Uwe T. Bornscheuer^{1,3‡}

¹Department of Biotechnology & Enzyme Catalysis, Institute of Biochemistry, University of Greifswald, 17487 Greifswald, Germany

²Department of Pharmaceutical Biotechnology, Institute of Pharmacy, University of Greifswald, 17487 Greifswald, Germany

³Institute of Marine Biotechnology e.V., 17489 Greifswald, Germany

⁴Max Planck-Institute for Marine Microbiology, 28359 Bremen, Germany

⁵University of Bremen, Center for Marine Environmental Sciences (MARUM), 28359 Bremen, Germany

⁶Department of Microbial Proteomics, Institute for Microbiology, University Greifswald, 17487 Greifswald, Germany

‡ Corresponding authors:

Thomas Schweder, schweder@uni-greifswald.de

Uwe T. Bornscheuer, uwe.bornscheuer@uni-greifswald.de

Additional tables

- Table S1: Proteins and accession numbers
- Table S2: Bacterial strains and plasmids
- Table S3: M9-mineral media D-glucose composition
- Table S4: M9-mineral media additives
- Table S5: Primer list
- The Supplementary Tables S6 and S7 are available as separate Excel files.

Additional figures

- Figure S1: Sugar composition of the cultivation media
- Figure S2: SDS-PAGE of *F. agariphila* KMM3901^T enzymes expressed recombinantly in *E. coli*
- Figure S3: Consumption of 5-dehydro-4-deoxy-D-glucuronate from *B. licheniformis* DSM13 during cultivation
- Figure S4: Growth of *B. licheniformis* DSM13 on different ulvan hydrolysates
- Figure S5: *B. licheniformis* DSM13 CAZyme repertoire and their expression
- Figure S6: Alignment *B. licheniformis* DSM13 sulfatases with lipoteichoic acid synthases
- Figure S7: Comparison of different *Bacillus* sp. to digest ulvan hydrolysate and ulvan derived monosaccharides
- Figure S8: Growth of the *Bacillus* sp. PL28 and GH105 expression strains
- Figure S9: Activity assays of *Bacillus* sp. PL28 and GH105 expression strains
- Figure S10: C-PAGE results from *Bacillus* sp. PL28 and GH105 expression strains.
- Figure S11: Growth of the different *B. licheniformis* expression strains.
- Figure S12: Activity assay results from *B. licheniformis* SH006, PL28 and GH105 single- and co-expression strain.
- Figure S13: C-PAGE results from *B. licheniformis* SH006 PL28, GH105 and co-expression strains.
- Figure S14: C-PAGE from the cultivation supernatant of *B. licheniformis* strains in M9-mineral media.

Table S1: List of *F. agariphila* proteins used to produce ulvan hydrolysates. (UHA) P30_PL28 and P33_GH105 and (UHB) all listed enzymes. All proteins were recombinantly expressed in *E. coli* as described previously [1].

Name	Locus tag	Uniprot ID	Functional annotation	size [kDa]
P1_GH88	*21900	T2KLZ3	Unsaturated glucuronylhydrolase (GH88)	44.1
P10_PLnc	*21990	T2KNA3	Ulvan lyase (PLnc)	92.6
P17_GH2	*22060	T2KN75	β -Galactosidase (GH2)	112.9
P18_S1_7	*22070	T2KPK5	Arylsulfatase (S1_7)	53.1
P20_GH78	*22090	T2KNB2	α -L-rhamnosidase (GH78)	100.4
P24_GH3	*22130	T2KMH0	β -Glucosidase (GH3)	79.3
P27_GH43	*22160	T2KN85	β -Xylosidase (GH43)	67.3
P30_PL28	*22190	T2KNC2	Ulvan lyase (PL28)	44.1
P31_GH39	*22200	T2KM23	Glycoside hydrolase (GH39)	54.6
P33_GH105	*22220	T2KPL9	Glycoside hydrolase (GH105)	40.7
P36_S1_25	*22250	T2KM26	α -L-rhamnosidase/-sulfatase (GH78/S1_25)	134.3

*=BN863

Table S2: Bacterial strains and plasmids

Strain	Description or genotype	Reference or source
<i>Escherichia coli</i> TOP10	<i>F- mcrA</i> Δ (<i>mrr-hsdRMS-mrcB</i>) ϕ 80 <i>lacZ</i> Δ M15 Δ <i>lacX74 nupG</i> <i>recA1 araD139</i> Δ (<i>ara-leu</i>)7697 <i>galK16 rpsL</i> (StR) <i>endA1 fhuA2</i> λ -	Originally purchased from Invitrogen by Thermo Fisher Scientific (Waltham, MS, USA)
<i>Escherichia coli</i> BL21(DE3)	B F-- <i>ompT gal dcm lon</i> <i>hsdSB</i> (rB—mB-) λ (DE3 [<i>lacI</i> <i>lacUV -T7 gene 1 ind1 sam7 nin5</i>])	Originally purchased from New England Biolabs (Ipswich, MS, USA)
<i>Bacillus subtilis</i> B1	Wildtype	Department of Biotechnology & Enzyme Catalysis (University of Greifswald)
<i>Bacillus licheniformis</i> DSM13	Wildtype	Department of Pharmaceutical Biotechnology (University of Greifswald) originating from Veith et al [2]
<i>Staphylococcus carnosus</i> 80-285	Wildtype	Kindly provided by the Enzymicals AG (Greifswald, Germany)
<i>Staphylococcus carnosus</i> 20-282	Wildtype	Kindly provided by the Enzymicals AG (Greifswald, Germany)
<i>Saccharomyces cerevisiae</i> GRF18	Wildtype	Department of Biotechnology & Enzyme Catalysis (University of Greifswald)
<i>Vibrio natriegens</i> ATCC 14048	Wildtype	Originating from the ATCC (Manassas, V, USA)
<i>Pseudomonas putida</i> DSMZ 50198	Wildtype	Originating from the DSMZ (Braunschweig, German)
<i>Pichia pastoris</i> X33	Wildtype	Originally from Invitrogen by Thermo Fisher Scientific (Waltham, MS, USA)
<i>Cutaneotrichosporan curvatus</i> DSM 101032	Wildtype	Department of Computational Synthetic Biology (TU Darmstadt) as described by Hofmeyer <i>et al.</i> [3]
<i>Cupriavidus necator</i> H16	Wildtype	Department of Microbial proteomics (University of Greifswald) as described by Pohlmann <i>et al.</i> [4]
<i>Bacillus licheniformis</i> MW3	Δ <i>hsdR1</i> , Δ <i>hsdR2</i>	Waschkau <i>et al.</i> 2008 [5]
<i>Bacillus subtilis</i> JK138	<i>sfp+</i> , Δ <i>sacA::SpecR</i> , Δ <i>lytC::lox72</i> , Δ <i>bpr-spo::lox72</i> , Δ <i>nprB::lox72</i> , Δ <i>mpr::lox72</i> , Δ <i>aprE::lox72</i> , Δ <i>nprE::lox72</i> , Δ <i>vpr::lox72</i> , Δ <i>epr::lox72</i> , Δ <i>wprA::lox72</i> , Δ <i>srfA::(comS,lox72)</i> , Δ <i>pksX::lox72</i> , Δ <i>pps::lox72</i> , Δ <i>amyE::lox72</i>	Krüger <i>et al.</i> 2022
<i>Bacillus licheniformis</i> SH006	Δ <i>hsdR1</i> , Δ <i>hsdR2</i> , Δ <i>pga</i> , Δ <i>apr</i>	this study
pE194SV	pE194 derivative; <i>E. coli</i> / <i>Bacillus</i> -shuttle vector; <i>EmR</i> ; ori pE194; <i>AmpR</i> , ori pUC18	This work
pDhsdR1	pE194SV with Δ <i>hsdR1</i> homology flanking region	This work
pDhsdR2	pE194SV with Δ <i>hsdR2</i> homology flanking region	This work
pDpga	pE194SV with Δ <i>pga</i> homology flanking region	This work
pDapr	pE194SV with Δ <i>apr</i> homology flanking region	This work

Table S3: D-glucose-supplemented M9-mineral medium

M9 - mineral media stock solution	Final concentration per 1 L
M9 salt solution	1x
20 % glucose	0.4 %
1 M MgSO ₄	1 mM
1 M CaCl ₂	0.3 mM
Biotin (1 mg/ml)	4.09 nM
Thiamine (1mg/ml)	3.77 nM
Trace element solution (100x)	1 x

Table S4: M9-mineral medium additives

Additive	Component	Concentration in stock solution
M9 salts (10x)	Na ₂ HPO ₄ *2 H ₂ O	422.6 mM
	KH ₂ PO ₄	141.4 mM
	NaCl	93.6 mM
	NH ₄ Cl	93.5 mM
	ddH ₂ O	1 L
Ulvan solution	Ulvan	2 %
	Sodium phosphate buffer	25 mM
	Sodium chloride	50 mM
	pH 7.5	
Ulvan hydrolysate	Ulvan	2 %
	Sodium phosphate buffer	25 mM
	Sodium chloride	50 mM
	pH 7.5	
	Enzyme mix	5 µL/mL per enzyme
Enzyme mix	Sodium phosphate buffer	25 mM
	Sodium chloride	50 mM
	pH 7.5	
	Enzyme mix	5 µL/mL per crude extract enzyme from table S1
Trace element solution 100 x	FeCl ₃	13.1 mM
	ZnCl ₂	0.62 mM
	0.1 M CaCl ₂ • 2H ₂ O	76 µM
	0.2 M CoCl ₂ • 6 H ₂ O	42 µM
	0.1 M HBO ₃	162 µM
	1 M MnCl ₂ • 4 H ₂ O	8.1 µM
	KI	0.5 mM
	Na ₂ MoO ₄	1 mM
EDTA	13.4 mM	
Yeast extract		2 % (w/v)
Sodium chloride		15 % (w/v)

Table S5: List and sequences of primers used in this study.

Primer	Sequence 5' → 3'
MaZu7	ACGTCATATGAAAATCAGTATGCAAAAAGCAGATTTTTGGAAAAAAGCAGCGATCTCATTACT TGTTTTACCCATGTTTTTACCCTGATGATGAGCG
MaZu8	CACCATGTTTTTACCCTGATGATGAGCGAAACGGTTTTTGCGCCAAACGGCACCGGATGAA GACACAAGCGCCATTACGAG
MaZu9	ACGTGGTACCTTACAGGCTTTCGACTGCGATGGCTTTCAGACGC
MaZu10	CACCATGTTTTTACCCTGATGATGAGCGAAACGGTTTTTGCGCCAAACGGCCTTAACCATA GCGAAATCGAAGC
MaZu11	ACGTGGTACCTTATTCTTCCAGTTTTCAAGACTTCGCTTCCCTGCCATCAG
MaZu12	ACGTCATATGTTGATCAACAAAAGCAAAAAGTTTTTCGTTTTTCTTTCATTTTTGTTATGATG CTGAGCCTCTCATTTGTGAATGGG
MaZu13	GATGCTGAGCCTCTCATTTGTGAATGGGGAAGTTGCAAAAGCCCAAACGGCACCGGATGAA GACACAAGCGCC
MaZu14	GATGCTGAGCCTCTCATTTGTGAATGGGGAAGTTGCAAAAGCCCAAACGGCCTTAACCAT AGCGAAATCGAAGCG
MaZu19	ACGTACCATATGCAAACGGCACCGGATGAAGACACAAGCGCC
MaZu20	ACGTACTCTAGACAGGCTTTCGACTGCGATGGCTTTCAGAC
MaZu21	ACGTACCATATGCAAAAAGGCCTTAACCATAGCGAAATCGAAGCG
MaZu22	ACGTACTCTAGATTCTTCCAGTTTCAAGACTTCGCTTCCCTGCCATCAG
MaZu28	TGATAAGCGTTGGTTTGGCAATCTTATCGGGCTATGCATTTATAAAATG
MaZu29	GCATAGCCCGATAAGATTGCCAAACCAACGCTTATCAATAGAAAAAGAGCATTTTTTGAAAC AAAATTC
MaZu35	TGCGTTAGCAATTTAACTGTGATAAACTACCGCATTAATAGAAAAAGAGCATTTTTTGAAACA AAAATTC
MaZu36	TAATGCGGTAGTTTATCACAGTTAAATTGCTAACGCAGTCAGGCACCGT
MaZu37	GTTTTTAAAGGCTTTTAAAGCCGTCTGTACGTTTCTTAACATCTGATGTCTTTGCTTGCGAATG TTCATCT
MaZu38	TTAGGAACGTACAGACGGCTTAAAAGCCTTTAAAACGTTTTTAAAGGGTTTGTAGACAAGG TAAAGGATAAAACAG
P1-hsdR1	CTGCAGGGTCTCAACCCGAACAGCGTAAGGCTGATG
P2-hsdR1	CCGTAATTTGAATCTATTAGACAAACATCTTTTGTAGGAATG
P3-hsdR1	GATGTTTGTCTAATAGATTCAAATTACGGGCCTTG
P4-hsdR1	TCTAGAGGTCTCATGAGATCGGTTTTATGAAAGCGTC
P1-hsdR2	CTGCAGCGTCTCAACCCGATAAAAGGATTACTGTGCG
P2-hsdR2	TCCATGTTGTCACAACCTATTGTTGAGAATAAAGGAAAAGGAG
P3-hsdR2	CCTTTATTCTCAACAATAGGTTGTGACAACATGGAAAG
P4-hsdR2	TCTAGACGTCTCATGAGGTGCTTTCATCAATCGTAAATC

Table S6: Summary of the proteomic results (available as separate Excel file)**Table S7:** Results of statistical analyses (Welch's T-test, FDR 0.05) (available as separate Excel file)

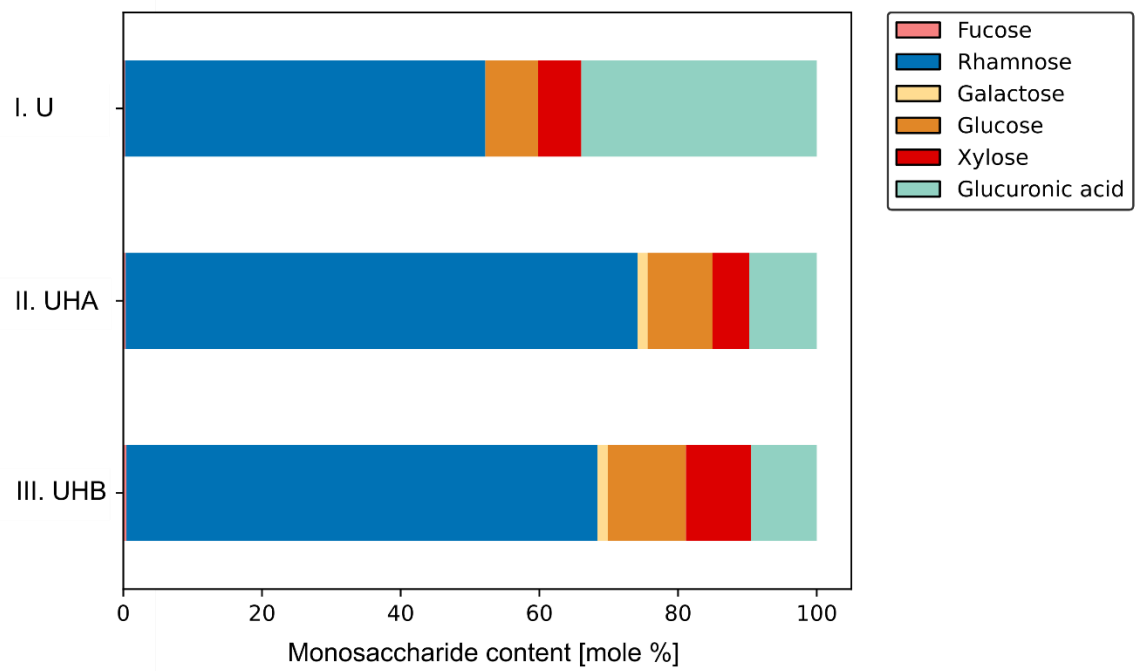


Fig. S1 Sugar composition of the cultivation media. The complex carbon sources were analysed for their detailed monosaccharide composition via acid hydrolysis and HPAEC-PAD. Ulvan (U), the partially P30_PL28 and P33_GH105-hydrolyzed ulvan (UHA) and the completely digested ulvan (UHB) using the whole cascade of ulvan degrading enzymes [1].

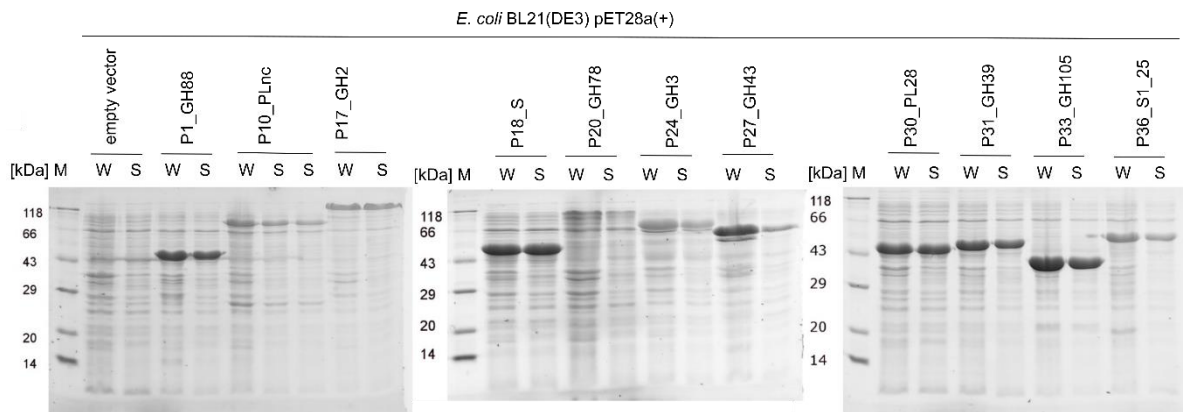


Fig. S2 SDS-PAGE of *F. agariphila* KMM3901^T enzymes expressed recombinantly in *E. coli* BL21(DE3) as described by Reisky *et al.* [1]. The cells were normed (7/OD) and the whole cell (W) extract and soluble protein (S) fraction were analyzed by SDS-PAGE containing 1% (V/V) trichloroethanol (TCE). The protein bands were visualised under UV and the pictures were colour-inverted and decolorized. The protein marker Roti[®]-Mark from Carl Roth (Karlsruhe, Germany) was used.

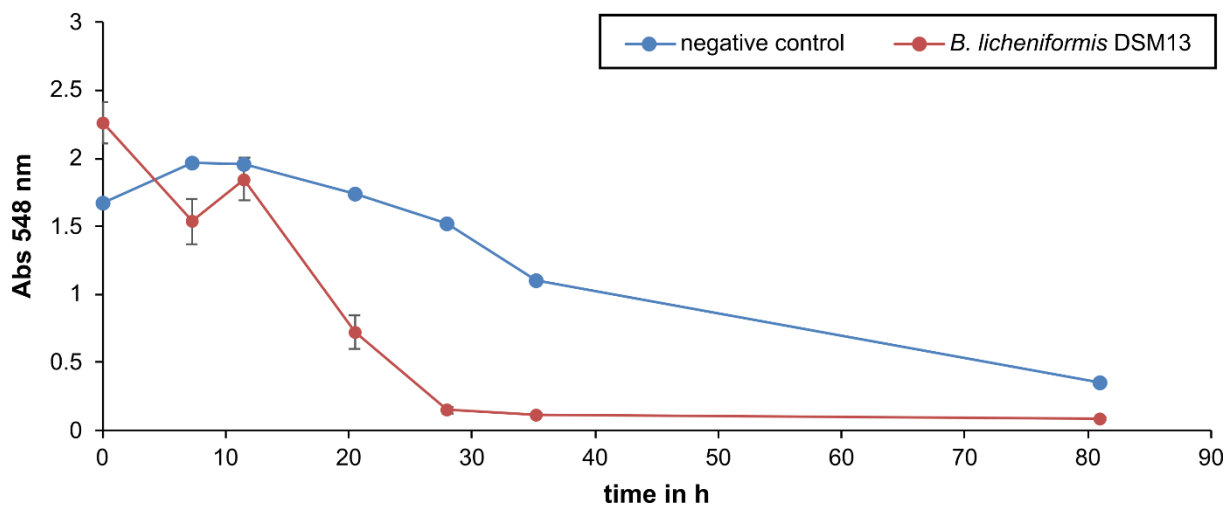


Fig. S3 Consumption of 5-dehydro-4-deoxy-D-glucuronate by *B. licheniformis* DSM13 during cultivation. Thiobarbituric acid assay determined 5-dehydro-4-deoxy-D-glucuronate in cultivation supernatants (M9-ulvan hydrolysate UHB, see Fig. 3a). Cell-free medium served as negative control.

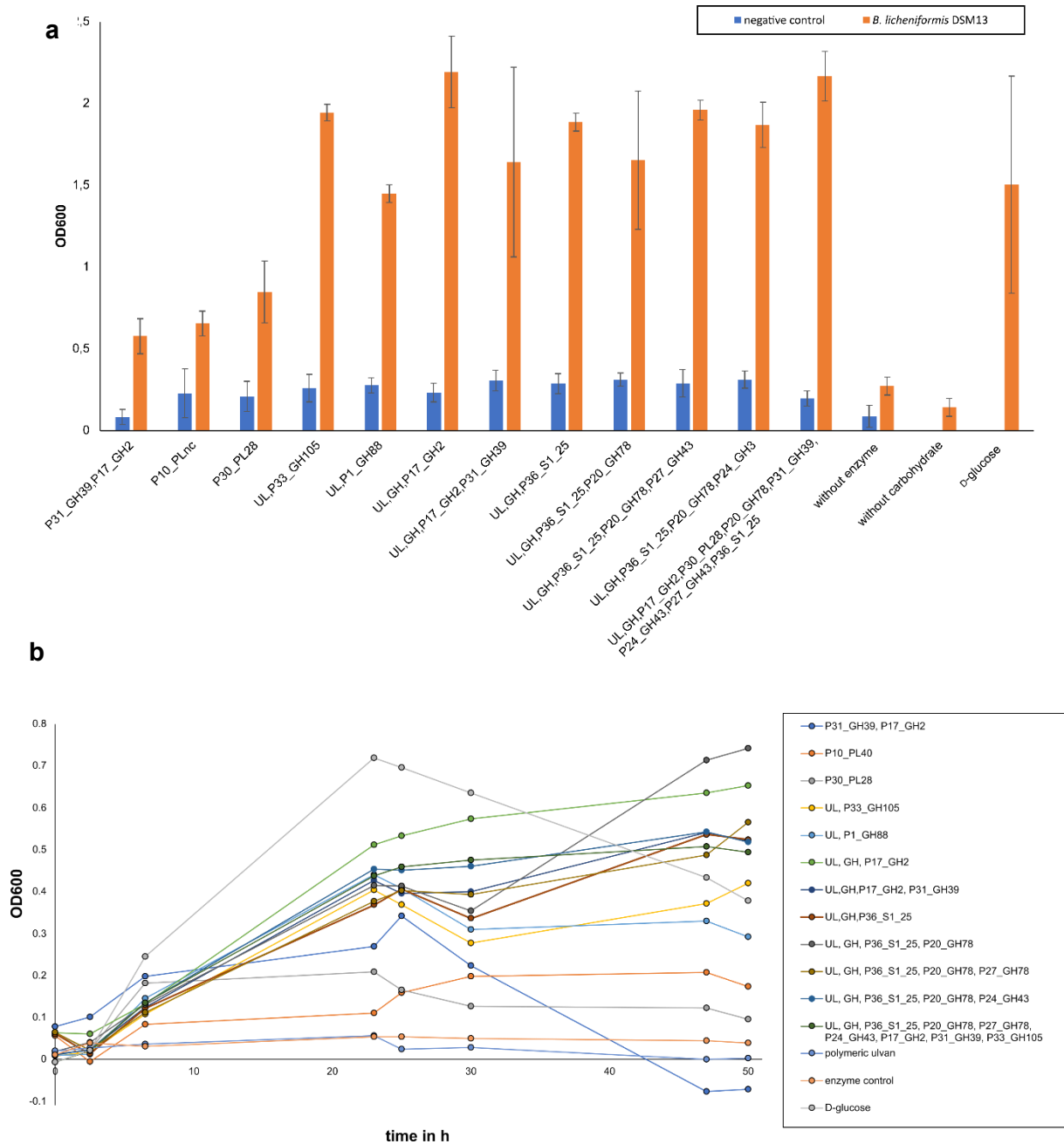


Fig. S4 Growth of *B. licheniformis* DSM13 on ulvan (without enzyme) and ulvan hydrolysates representing different levels of degradation. For hydrolysis, different *F. agariphila* enzymes, enzyme combinations or all enzymes (recombinantly expressed in *E. coli*) were used (Table S1). *B. licheniformis* DSM13 was cultivated in 1 mL in 96 deep-well plates and OD600 was measured after 48 h (a). This graph represents the full dataset of the Figure 2 in the main text. For investigation of the growth behaviour the culture was cultivated in 200 μ L scale in a low-well plate and measured for 48 h (b).

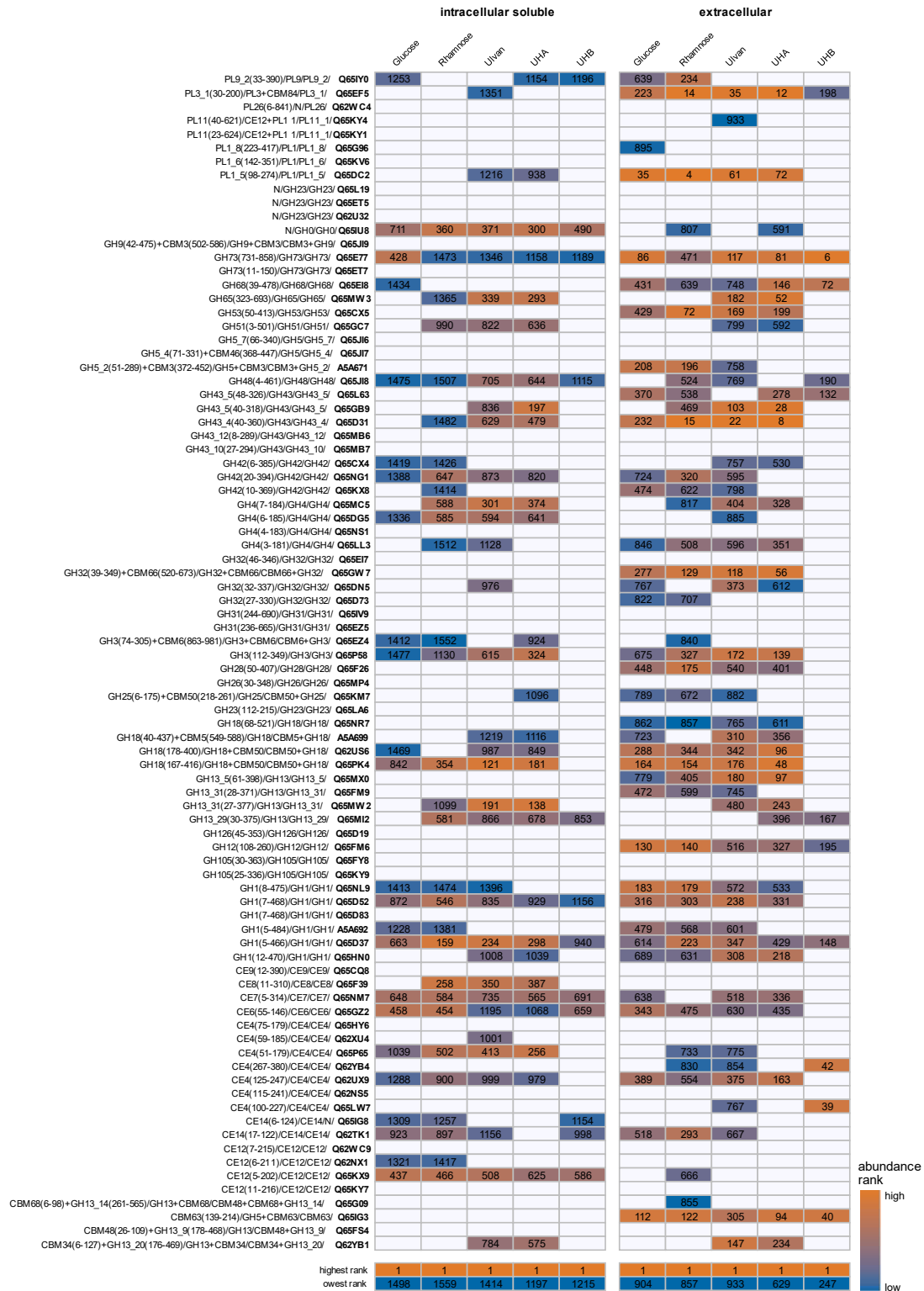


Fig. S5 List of PLs, GHs and CEs identified by dbCAN2 [6] and their abundance in the intracellular soluble and extracellular proteomes. The graph indicates the relative abundance of proteins within the respective sample given as abundance ranks. Abundance ranks were derived from %riBAQ values (Table S6). The lowest rank corresponds to the total number of quantified proteins per sample. Blank tiles represent proteins that were not quantified. UHA/B: Ulvan hydrolysate A/B (see Fig. 3a). Protein IDs are highlighted in bold together with the full output of dbCAN2 analyses (HMMER/Hotpep/DIAMOND/Protein ID).

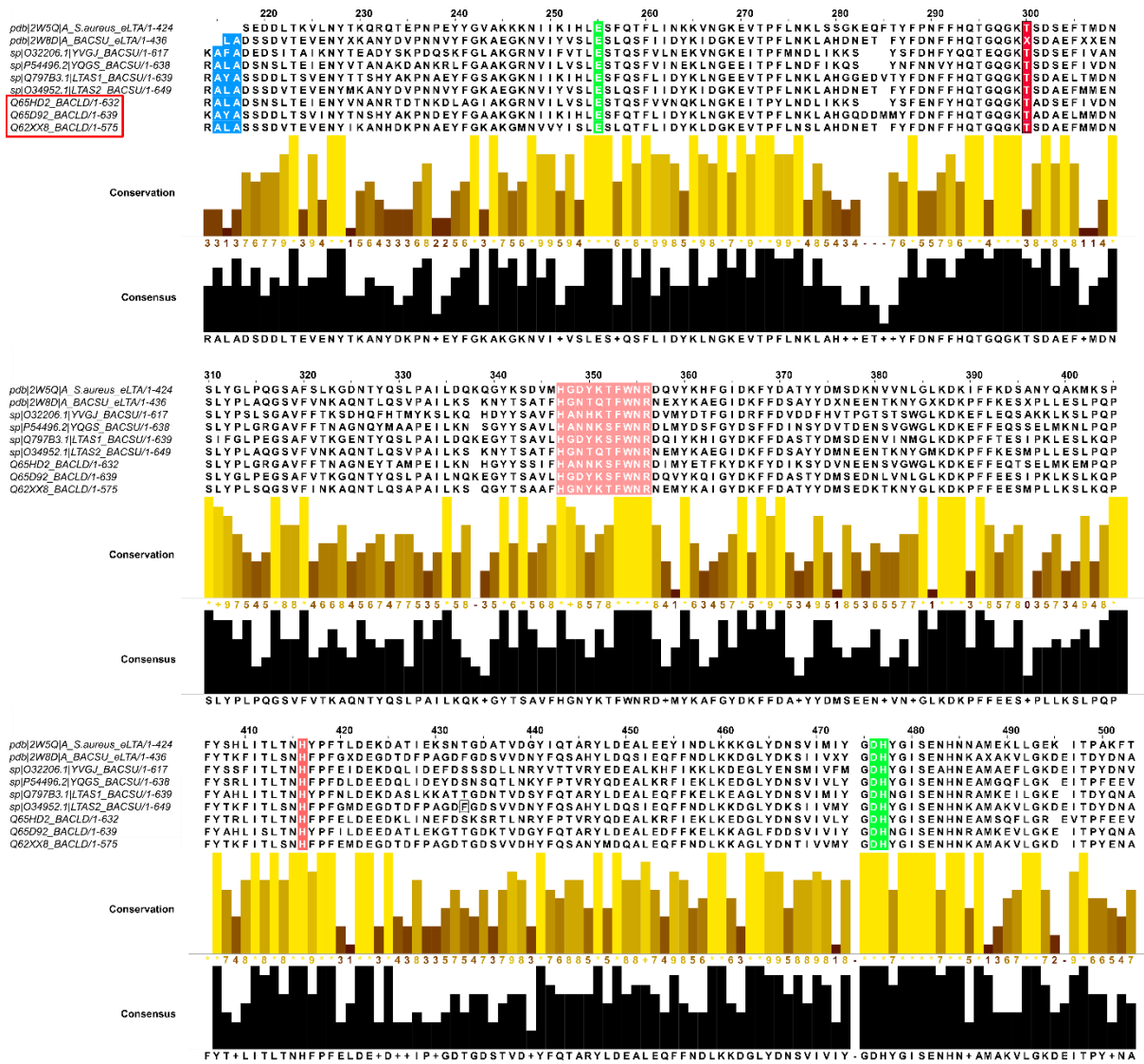


Fig. S6 Alignment of sulfatases from *B. licheniformis* DSM13 with lipoteichoic acid synthases of other Gram-positive bacteria. The lipoteichoic acid synthases (LTA) were the highest hit in the BlastP search. LTAs consist of a transmembrane and an extracellular domain. Amino acid residues in the region of the extracellular LTA (eLTA) domain (210 to 585 As) are displayed. LTAs are synthesized as membrane proteins and are cut at a cleavage sequence AXA (blue). T300 (marked in red) is the catalytic residue in the eLTA of *S. aureus*, while H416 is involved in the reaction mechanism (protonation of the leaving group). Residues marked in green are important for the binding of Mn²⁺. Residues 347-356 (HxD/NxxFW/YNR) are important for substrate binding [7].

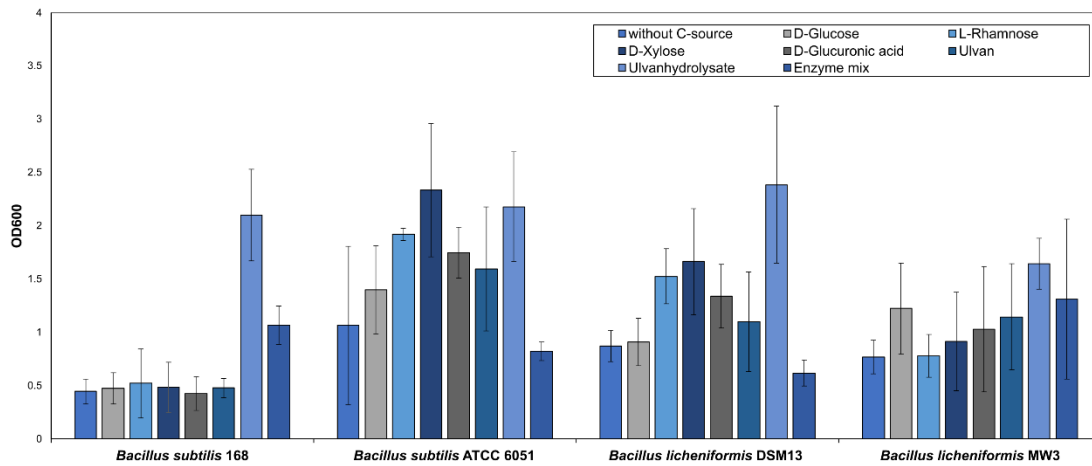


Fig. S7 Comparison of different *Bacillus* sp. for the conversion of ulvan hydrolysate and ulvan derived monosaccharides. The strains were cultivated in 1 mL Belitzky-Minimal media [8] with variation of the carbon source in a 96-deep well-plate. General growth in OD600nm was measured after 24 h, showing a general acceptance of the ulvan hydrolysate for all four *Bacillus* sp. in the chosen Belitzky-Minimal medium conditions.

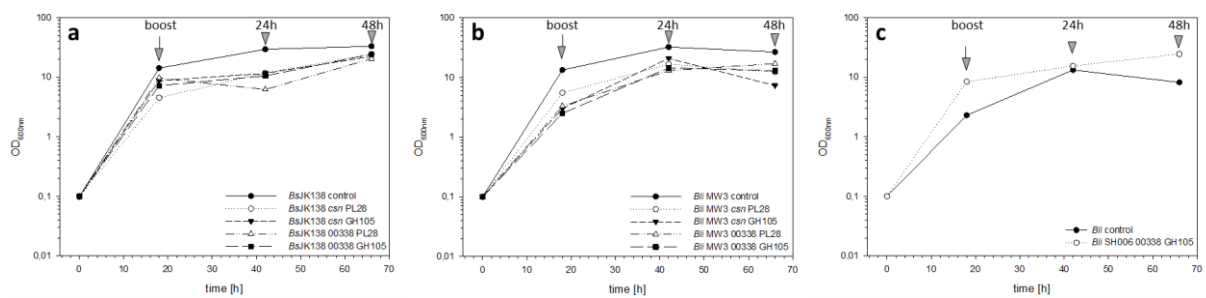


Fig. S8 Growth of the PL28 and GH105 *Bacillus* expression strains. *B. subtilis* JK138. **(a)**, *B. licheniformis* MW3 **(b)** and *B. licheniformis* SH006 (Δapr) **(c)** were grown under simulated fed-batch conditions in Espresso-B medium at 30°C and 250 rpm. Sampling points for activity measurements, 24 h and 48 h after the boost are indicated.

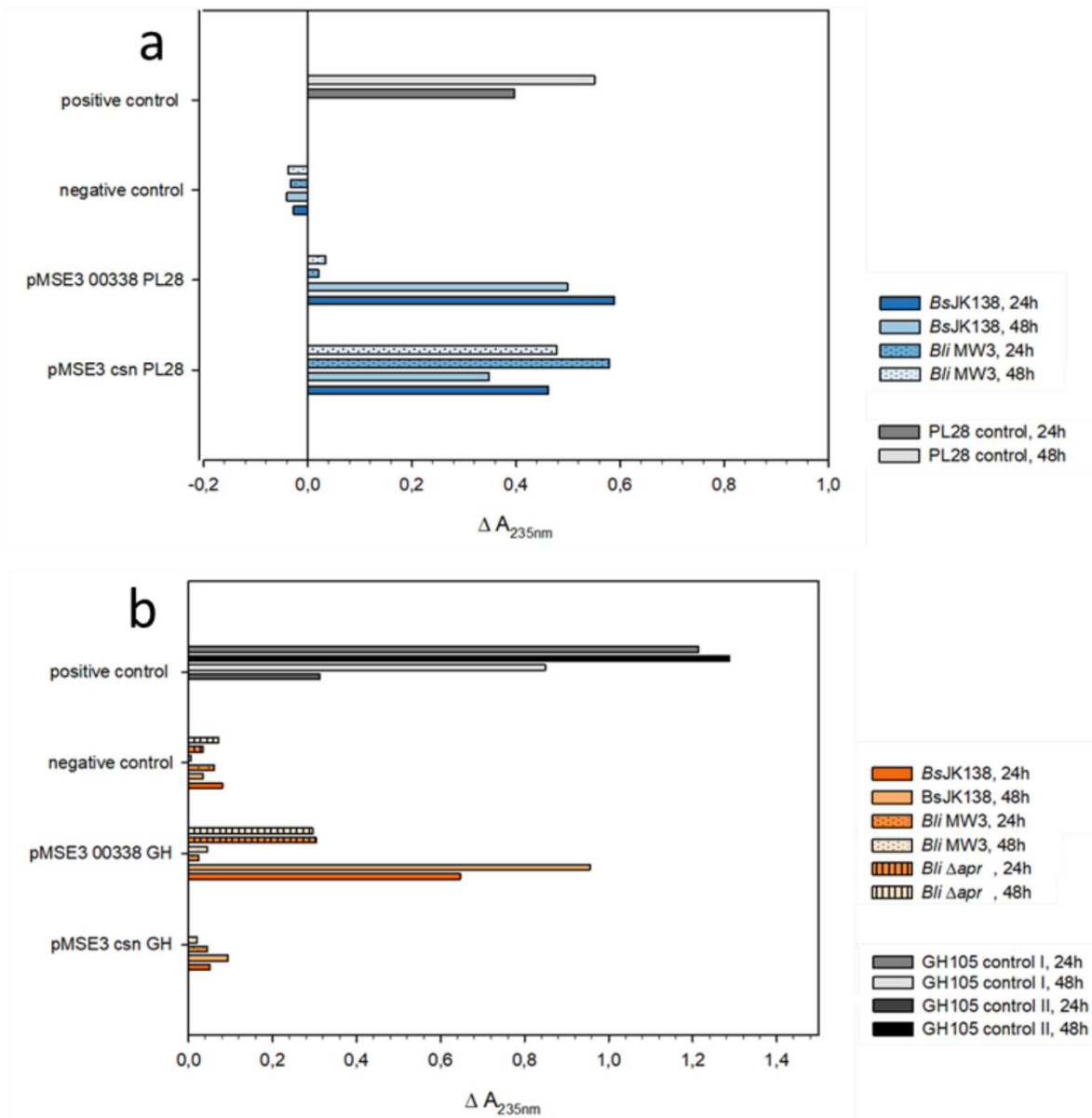


Fig. S9 Activity assay results from *Bacillus* sp. PL28 and GH105 expression strains. (a) Results of the ulvan lyase (UL) assay measurement of the lyase product formation at 235 nm over 60 min. The deviation of the absorption of the end-start reveals the lyase product formation (lyase activity), while the deviation of the absorption start-end shows the reversed reaction of the GH105 activity (b) cleaving of the lyase moiety using prehydrolyzed ulvan from the recombinantly expressed ulvan lyase PL28.

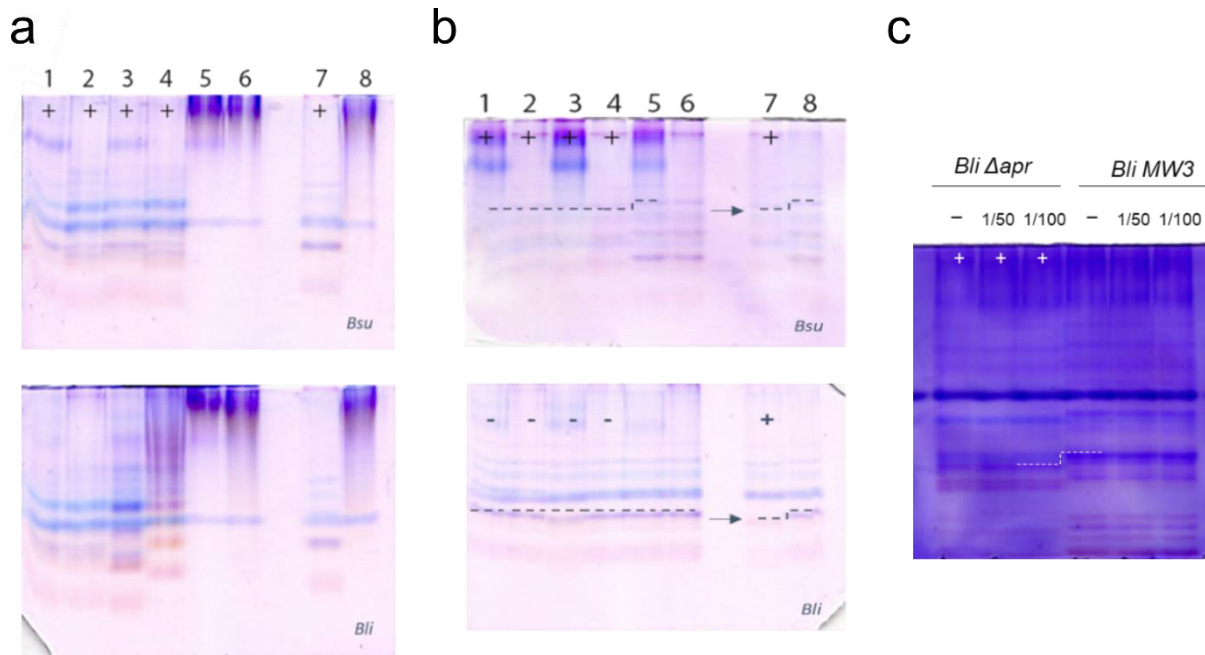


Fig. S10 C-PAGE results from *Bacillus* sp. PL28 and GH105 expression strains. **(a)** The C-PAGE (carbohydrate polyacrylamide gel electrophoresis) corresponds to the ulvan lyase assay shown in Fig. S10 to the results of the ulvan lyase (UL) expression and **(b)** reaction of the GH105 activity cleaving of the lyase moiety using prehydrolyzed ulvan from the recombinantly expressed ulvan lyase PL28. Intracellular (in) and extracellular (ex) fractions of *B. subtilis* JK138 (*Bsu*) and *B. licheniformis* DSM13 (*Bli*) were analyzed. Lane 1: TB1 csn-UL ex, Lane 2: TB1 csn-UL in, Lane 3: TB1 00338-UL ex, Lane 4: TB1 00338-UL in, Lane 5: Ko (empty vector) ex, Lane 6: Ko (empty vector), Lane 7: positive control, Lane 8: negative control. **(c)** C-PAGE after 24 h of GH105 expression in *B. licheniformis* SH006 (*Bli* Δapr) and *B. licheniformis* MW3 (*Bli* MW3) with the addition of the protease inhibitor in 1/100, 1/50 and no protease (-) addition.

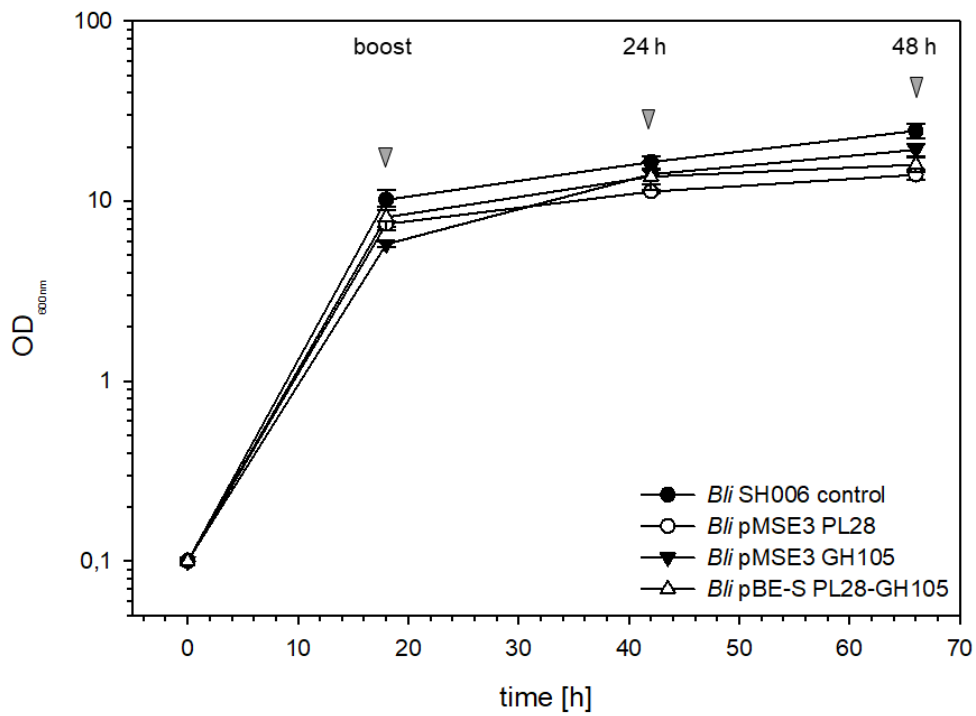


Fig. S11 Growth of the different *B. licheniformis* expression strains. *B. licheniformis* SH006, *B. licheniformis* pMSE3 PL28, *B. licheniformis* pMSE3 GH105 and *B. licheniformis* pBE-S PL28-GH105 were grown under simulated fed-batch conditions in Espresso-B medium at 30°C and 250 rpm. Sampling points for activity measurements, 24 h and 48 h after the boost are indicated.

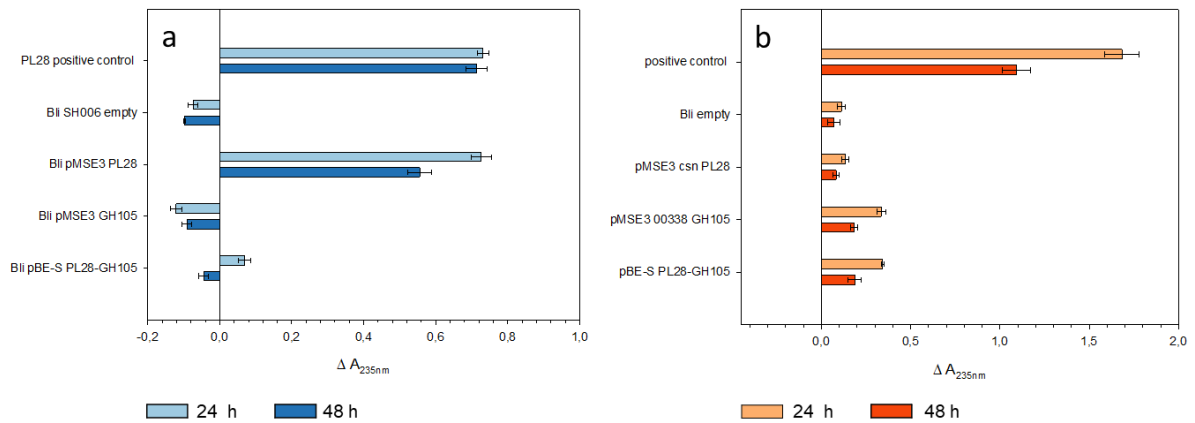


Fig. S12 Activity assay results from *B. licheniformis* SH006, PL28 and GH105 single- and co-expression strain. **(a)** Results of the ulvan lyase (PL28) assay measurement of the lyase product formation at 235 nm over 60 min. The deviation of the absorption of the end-start shows the lyase product formation (lyase activity). **(b)** The deviation of the absorption start-end shows the reversed reaction of the GH105 activity by cleaving of the lyase moiety using prehydrolyzed ulvan from the recombinantly expressed ulvan lyase PL28. Corresponding C-PAGE gels are shown in Fig. S13.

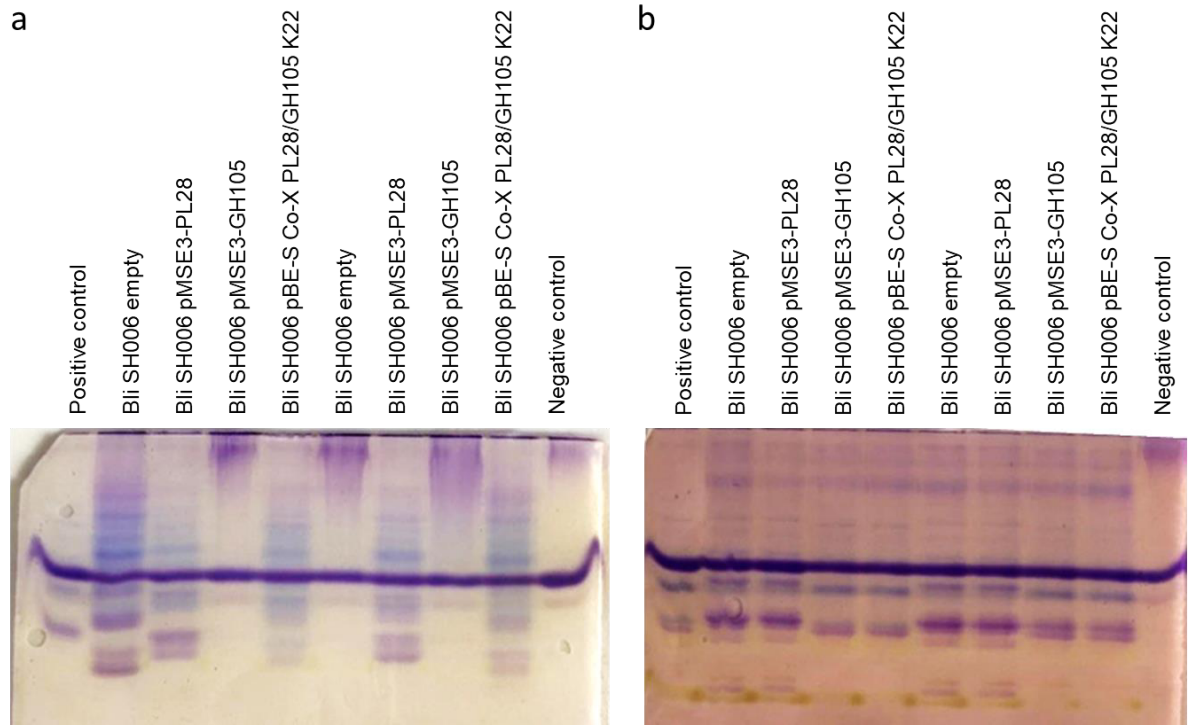


Fig. S13 C-PAGE results from *B. licheniformis* SH006 PL28, GH105 and co-expression strains. The results correspond to the ulvan lyase assay shown in Fig. S12. (a) Shows the C-PAGE result of the PL28 ulvan lyase activity (extracellular fraction) of the strain constructs *B. licheniformis* SH006 (Bli SH006) empty, pMSE3_PL28, pMSE3- GH105 or pBE-S Co-X PL28/GH105 K22 (b) shows the reaction of the GH105 cleaving of the lyase moiety using prehydrolyzed ulvan.



Fig. S14 C-PAGE from the cultivation supernatant of *B. licheniformis* strains in M9-mineral media. *B. licheniformis* (Bli) SH006 PL28, GH105 and coexpression strains were cultivated in M9-mineral media (see Fig. 6) for seven days. A supernatant sample was loaded, as negative control the ulvan polymer and as a positive control the predigested ulvan with PL28 was loaded on the gel.

References

1. Reisky L, Précoux A, Zühlke MK, Baumgen M, Robb CS, Gerlach N, et al. A complex enzyme cascade degrades the polysaccharide ulvan from green algae. *Nat Chem Biol.* 2019;803–12.
2. Veith B, Herzberg C, Steckel S, Feesche J, Maurer KH, Ehrenreich P, et al. The complete genome sequence of *Bacillus licheniformis* DSM13, an organism with great industrial potential. *J Mol Microbiol Biotechnol.* 2004;7:204–11.
3. Hofmeyer T, Hackenschmidt S, Nadler F, Thürmer A, Daniel R, Kabisch J. Draft genome sequence of *Cutaneotrichosporon curvatus* DSM 101032 (Formerly *Cryptococcus curvatus*), an oleaginous yeast producing polyunsaturated fatty acids. *Genome Announc.* 2016;4:e00362-16.
4. Pohlmann A, Fricke WF, Reinecke F, Kusian B, Liesegang H, Cramm R, et al. Genome sequence of the bioplastic-producing “Knallgas” bacterium *Ralstonia eutropha* H16. *Nat Biotechnol.* 2006;24:1257–62.
5. Waschkau B, Waldeck J, Wieland S, Eichstädt R, Meinhardt F. Generation of readily transformable *Bacillus licheniformis* mutants. *Appl Microbiol Biotechnol.* 2008;78:181–8.
6. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 2012;40:W445–51.
7. Lu D, Wörmann ME, Zhang X, Schneewind O, Gründling A, Freemont PS. Structure-based mechanism of lipoteichoic acid synthesis by *Staphylococcus aureus* LtaS. *Proc Natl Acad Sci. Proceedings of the National Academy of Sciences;* 2009;106:1584–9.
8. Stülke J, Hillen W. Regulation of Carbon Catabolism in *Bacillus* Species. *Annu Rev Microbiol.* 2000;54:849–80.

Article IV

1 **Utilization of a diverse range of xylan structures by marine *Bacteroidetes***

2 Theresa Dutschei¹, Irena Beidler², Daniel Bartosik², Julia-Maria Seeßelberg³, Michelle Teune¹,
3 Marcus Bäumgen¹, Soraia Querido Ferreira¹, Julia Heldmann¹, Felix Nagel⁴, Joris Krull^{5,6},
4 Leona Berndt⁷, Karen Methling⁸, Martin Hein⁹, Dörte Becher¹⁰, Peter Langer⁹, Mihaela Delcea⁴,
5 Michael Lalk⁷, Michael Lammers⁷, Matthias Höhne³, Jan-Hendrik Hehemann^{5,6}, Thomas
6 Schweder^{2,5*}, Uwe T. Bornscheuer^{1*}

7

8 ¹Department of Biotechnology & Enzyme Catalysis, Institute of Biochemistry, University
9 Greifswald, 17487 Greifswald, Germany

10 ²Department of Pharmaceutical Biotechnology, Institute of Pharmacy, University of Greifswald,
11 17487 Greifswald, Germany

12 ³Department of Protein Biochemistry, Institute of Biochemistry, University of Greifswald, 17487
13 Greifswald, Germany

14 ⁴ Department of Biophysical Chemistry, Institute of Biochemistry, University of Greifswald,
15 17487 Greifswald, Germany

16 ⁵ Institute of Marine Biotechnology e.V., 17489 Greifswald, Germany

17 ⁶ University of Bremen, Center for Marine Environmental Sciences, 28359 Bremen, Germany

18 ⁷ Department of Synthetic and Structural Biochemistry, Institute of Biochemistry, University of
19 Greifswald, 17487 Greifswald, Germany

20 ⁸ Department of Cellular Biochemistry and Metabolomics, Institute of Biochemistry, University
21 of Greifswald, 17487 Greifswald, Germany

22 ⁹ Department of Organic Chemistry, Institute of Chemistry, University of Rostock, 18059
23 Rostock, Germany

24 ¹⁰ Department of Microbial Proteomics, Institute of Microbiology, University of Greifswald,
25 17487 Greifswald, Germany

26

27

28 *Corresponding authors:

29 Thomas Schweder, schweder@uni-greifswald.de;

30 Uwe Bornscheuer, uwe.bornscheuer@uni-greifswald.de

31

32

33 **Running title:** Xylan utilization of marine *Bacteroidetes*

34

35

36

37

38

39 **Originality-Significance Statement**

40 Xylose-containing polysaccharides are abundant sugar polymers in nature and have been
41 detected in macro- and microalgae. However, their exact individual structures are poorly
42 understood, and degradation pathways that decompose these complex algal sugars are
43 unknown. The proteogenomic and biochemical characterizations in this study elucidated for
44 the first time how different xylose-containing substrates are enzymatically degraded by marine
45 bacteria. Two xylan-specific enzymatic pathways could be detected in the marine
46 Bacteroidetes strain *Flavimarina* sp. Hel_I_48, encoded in dedicated polysaccharide utilization
47 loci (PULs). Moreover, global genome analysis revealed that conserved sets of genes
48 encoding xylan-degrading enzymes are abundant in the genomes of many polysaccharide-
49 utilizing marine Bacteroidetes. This suggests that enzyme modularity and substrate flexibility
50 may enable marine bacteria to consume a diverse range of xylan structures in their dynamic
51 marine habitats, in which polysaccharide compositions can be highly variable.

52

53 **Summary**

54 Members of the phylum *Bacteroidetes* are primary degraders of algal polysaccharides and are
55 therefore key players in marine carbon cycling. However, several underlying enzymatic
56 pathways that drive these utilization processes remain obscure. We identified a marine
57 *Bacteroidetes* strain, *Flavimarina* sp. Hel_I_48, which encodes two separate polysaccharide
58 utilization loci (PULs) that target different xylose-containing polysaccharides. We determined
59 the substrate specificity of these PULs by using proteogenomic and biochemical analyses of
60 the encoded carbohydrate-active enzymes. Proteomics indicated that these genomic regions
61 are specific for glucuronoxylans and arabinoxylans. The substrate specificities of these two
62 PULs were confirmed by biochemical analyses of the encoded key enzymes, which allowed
63 us to deduce the respective metabolic pathways for xylan utilization. The investigation of
64 different xylans revealed that the encoded xylanases show a promiscuous activity towards the
65 β -1,4-linked xylan backbone, while other encoded glycoside hydrolases and carbohydrate
66 esterases are specialized towards a specific xylan structure. We show that genes encoding

67 xylan-degrading enzymes are abundant in the genomes of polysaccharide-utilizing marine
68 *Bacteroidetes*. The observed enzyme modularity and their substrate flexibility may enable
69 marine bacteria to consume a diverse range of xylan structures, and presumably poses an
70 adaptation to the highly dynamic marine habitat with varying polysaccharide compositions.

71

72 **Keywords:** xylan; xylanase; *Bacteroidetes*; Flavobacteriia; GH43; carbon cycle; glycoside
73 hydrolase; proteome; algae; marine polysaccharides

74

75 **Introduction**

76 Marine algae catalyze half of the global photosynthetic production of carbohydrates [1]. The
77 major components of algal biomass are polysaccharides, which can represent more than 50%
78 of algae dry mass [2, 3]. Many organisms use polysaccharides as intracellular energy storage
79 compounds as well as structural cell wall components [4] or secrete them as extracellular
80 polymeric substances (EPS) with various functions [5]. While different microbial communities
81 have developed strategies to degrade particular polysaccharides [6], especially *Bacteroidetes*
82 are a dominant phylum in glycan-rich environments and employ special polysaccharide
83 degradation principles [7]. These organisms harbor gene clusters that encode for
84 carbohydrate-active enzymes (CAZymes), often organized with transporters and regulatory
85 proteins in so-called polysaccharide utilization loci (PULs) [8]. Previous studies showed that
86 CAZymes are highly specific towards their target polysaccharide and its decorations [9].
87 However, the exact enzymatic degradation mechanisms of algal polysaccharides are far less
88 understood than their terrestrial counterparts such as starch or cellulose [10, 11]. Interestingly,
89 some polysaccharides in the marine ecosystem are high similarity in composition to their
90 terrestrial relatives, like the hemicellulose xylan [12], which is known to occur in cell walls of
91 grains and wood in form of arabinoxylan, galactoarabinoxylan and glucuronoxylan with a β -
92 1,4-linked D-xylopyranose backbone [13]. Marine xylan can be found in the cell wall of green
93 algae (Chlorophyta/ Charophyta) and red algae (Rhodophyta) [14]. Its backbone is composed

94 of β -1,4- or β -1,3-linked D-xylopyranose, depending on the algal species and source.
95 Substituted β -1,4-xylan was found in species of charophyte green algae [12, 15]. In
96 chlorophyte green algae β -1,3-xylan is part of the cell wall [16, 17] and is reported to form triple
97 helix microfibrils. Red algae contain mostly β -1,3-linked xylans [18–20] and β -1,3:1,4-linked
98 xylans, in which one β -1,3-linkage follows four β -1,4-linkages [17, 21]. Furthermore, marine
99 xylans can be sulfated or phosphorylated, further demonstrating the high variability in
100 polysaccharide composition [21]. Xylose-containing polysaccharides can also be found in
101 microalgal biomasses [22–24] but their exact individual polysaccharide structures remain
102 obscure.

103 The depolymerization of complex xylans to their monomers D-xylose, L-arabinose and D-
104 glucuronic acid requires several enzymes. In previous studies mainly the enzymatic
105 degradation of terrestrial xylans was described [25–29]. While characterizations of single
106 enzymes provide a first insight into marine xylan degradation [30, 31], studies exploring full
107 enzymatic cascades present in a marine organism – as shown in this work – are essential to
108 our understanding of the multitude of substrates available in marine environments. As the
109 marine *Bacteroidetes* strain *Flavimarina* sp. Hel_I_48 contains two distinct putative xylan PULs
110 (PUL I: P162_RS02310-P162_RS02395; PUL II: P162_RS04015-P162_RS04080), we chose
111 it as a model organism for the degradation of xylan-rich marine substrates [32]. While both
112 genomic regions encode for similar proteins like xylanases and xylosidases, they differ in
113 further encoded CAZymes. The aforementioned PULs also contain genes for an unusually
114 high number of transporters and receptors for polysaccharide uptake (SusC/D-like pairs).

115 In this study we analyzed xylan utilization pathways facilitated by proteins encoded by both
116 PULs to further understand xylan metabolism of marine *Bacteroidetes*. CAZymes and SusD-
117 like binding proteins of these PULs were investigated with diverse marine and terrestrial xylans,
118 covering different motifs of xylan structures and common motifs of L-arabinose and D-
119 glucuronic acid decorations. This allowed us to gain an extensive picture of how marine
120 bacteria are adapted to the wide variety of marine xylose-containing polysaccharides.

121 Results

122 *Flavimarina sp. Hel_I_48 can utilize different xylan types*

123 To elucidate the preferred xylan substrates of the putative xylan PULs of *Flavimarina sp.*
124 *Hel_I_48* (Fig. 1), we employed growth experiments and proteomics with different xyans as
125 sole carbon source (Fig. S1, S2). This revealed a divergent expression pattern with PUL I
126 (P162_RS02310-RS02390) being highly expressed during growth on beechwood xylan (BX),
127 while PUL II (P162_RS04015-RS04080) showed almost no induction on this substrate (Fig.
128 1). The polysaccharide-binding protein SusD can be used as indicator for substrate
129 recognition. In this respect it is interesting to note that especially SusD_I_2 (P162_RS02355)
130 and SusD_I_3 (P162_RS02370) of PUL I showed a massive induction on xylan compared to
131 growth on all other investigated carbon sources, with SusD_I_3 (P162_RS02370) making up
132 over 1% of the entire proteome observed. This indicates a specificity of this PUL towards
133 glucuronoxyylan, a polysaccharide recently identified in samples of marine particulate organic
134 matter [24]. Higher protein levels of SusD_I_1 (P162_RS02310) from PUL I were observed on
135 all investigated xyans, most notably on the β -1,4/1,3-mix linked xylan of *Palmaria palmata*
136 (PPX). The broad expression response indicates a potential role of this SusD-like protein in
137 the general recognition of the xylan backbone. In contrast, SusD_II_1 (P162_RS4065) and
138 SusD_II_2 (P162_RS4075) of PUL II were not upregulated either on PPX, BX or *Caulerpa*
139 *prolifera* xylan (CPX), but were highly expressed during growth on arabinoxylyans from rye
140 (RAX) and wheat xylan (WAX) (Fig. 1). These substrates consist of a main β -1,4-xylan chain
141 with side chains of L-arabinose at the C3 or C2 position [25], indicating a specificity of this PUL
142 region towards such xylan substrates. The arabinoxyylan-type carbohydrate was also detected
143 in diatoms [23] by using a wheat arabinoxyylan-based antibody, which identified xylan in
144 samples of algal blooms. Additionally, a separate PUL containing multiple putative
145 arabinofuranosidases (P162_RS00625-RS00655) was significantly upregulated during growth
146 on arabinoxylyans compared to all other investigated substrates (Fig. 1, Table S5), indicating
147 further adaptation of *Flavimarina sp.* towards arabinose-containing polysaccharides.

148 *Multiple SusD proteins indicate highly diverse xylan substrates*

149 The two investigated xylan PULs contain several SusD-like proteins that were divergently
150 expressed during growth on different xylan substrates (Fig. 1). Therefore, we investigated the
151 general interaction of the recombinantly expressed and purified SusD-like proteins with our
152 chosen model substrates (BX, PPX, WAX-M, RAX, laminarin and no substrate as controls).
153 SusD-like proteins assist in the transport of glycan oligosaccharides through SusC-like pores
154 into the periplasm. We therefore hypothesized that different SusD-proteins might be required
155 for capturing xylan polysaccharides as partially hydrolyzed oligosaccharides with different
156 molecular architectures. Analysis via affinity gel electrophoresis (Fig. S3) showed that the
157 SusD_I_2 (P162_RS02355) protein does interact with a glucuronoxytan-like substrate, as it is
158 retained in a gel matrix [59] that included this substrate. This correlates well with the
159 upregulation pattern observed in the proteome analysis (Fig. 1). SusD_I_2 also showed
160 interaction with the other xylans, reducing the band pattern from three bands to one in the
161 presence of PPX, RAX and WAX-M (Fig. S3). The multiple band pattern might indicate different
162 oligomeric states of the SusD-like proteins, as it is known for bovine serum albumin (BSA) due
163 to multimerization [60]. SusD-like proteins are also known to form multimers in e. g. dimeric,
164 tetrameric and octameric form [61, 62]. The reduced number of formed bands can be indicative
165 of a preferred conformation state of the protein towards the ligand. As there was no interaction
166 with the laminarin control, the general motif and secondary structure of the polymeric β -1,4-
167 linked xylan is likely to be relevant in the recognition and the oligomeric state change of
168 SusD_I_2 (Fig. S3). Additionally, the overall ligand size seems to be important for a preferred
169 oligomeric state [61, 62]. The increased size of the polymer (BX<WAX-M<RAX<PPX, Table
170 S2) led to the reduction of the band number towards a single band (Fig. S3). In general, ligand
171 sizes play a crucial role in the binding recognition of SusD-like proteins, with interactions
172 towards more common oligosaccharides and the recognition of polysaccharides being rather
173 an exception [61, 62]. Overall, the diversity of these sugar-binding units in the PULs indicate

174 that they are adapted towards the multitude of xylan-like substrates that have previously been
175 proven to exist in the ocean [22, 24].

176

177 *Flavimarina sp. xylanases are specific for β -1,4-linked xylans*

178 Extracellular enzymes play a crucial role in the initial degradation of polysaccharides to shorter
179 fragments. Subsequently, these oligomers can be internalized by transporters for downstream
180 processing through other specialized CAZymes [6]. We studied the hydrolytic activity of all
181 PUL-encoded enzymes from *Flavimarina sp.* towards arabinoxylans from different sources
182 (RAX, WAX, BX, PPX and β -1,3-linked CPX) using a DNS-reducing end assay and ANTS-
183 FACE (Figs. S4, S5, S6, S7). These analyses revealed three xylanases, all of them containing
184 a SecII lipoprotein signal peptide that suggests an extracellular localization on the cell surface
185 [63]. FI4_GH10 (P162_RS02345) and FI8_GH10 (P162_RS04060) both showed
186 promiscuous *endo*-activity towards the β -1,4-xylan polymers RAX, WAX and BX as well as the
187 β -1,4/1,3 linked PPX [17, 21], releasing several oligosaccharides of varying length but they
188 had no activity towards the purely β -1,3-linked CPX (Fig. 2, Figs. S4, S5). The putative *endo*-
189 xylanase FI7_GH10 (P162_RS04050) showed no activity towards any of the chosen
190 substrates. *Exo*-xylanase activity was detected for FI2_GH10 (P162_RS032335) with
191 promiscuous activity towards RAX, WAX, BX and PPX (Fig. 2, Figs. S4, S5) at the non-
192 reducing end of the poly- and oligosaccharides (Fig. S8). Again, no activity towards CPX could
193 be detected. This xylanase promiscuity points to a general specificity towards 1,4- β -xylan
194 linked saccharides, which does, however, also include decorated substrates. This allows the
195 enzymes to degrade a multitude of different xylan structures, such as mixed β -1,4/1,3-xylans
196 or substrates branched with D-glucuronic acid or L-arabinose moieties, as seen in BX or
197 different arabinoxylans (RAX, WAX), respectively. Their extracellular localization attached to
198 the outer membrane, in combination with this generalist degradation pattern suggests that
199 these enzymes present an adaptation of *Flavimarina sp.* towards a multitude of different β -1,4-
200 xylan structures.

201

202 *Intracellular glycoside hydrolases facilitate xylan PUL specificity*

203 As xylanases from both PULs are proven to be active on a multitude of different xylose-
204 containing polysaccharides, we postulated that the remaining CAZymes from both PULs could
205 be responsible for their specificity. PUL I additionally encodes for two putative α -
206 glucuronidases (F11_GH67 (P162_RS02330) and F13_GH115 (P162_RS02340)), indicating
207 specialization towards a glucuronoxylan type substrate. Enzymatic assays with the synthetic
208 substrate *p*NP- α -D-glucuronide indeed revealed glucuronidase activity (Fig. 3) for F11_GH67.
209 The analysis of all GHs on polymers pre-treated with the *endo*-xylanase F14_GH10 did not
210 reveal any further activity (Fig. S6). Screening of the PUL II enzymes on *endo*-xylanase
211 F118_GH10 hydrolysate (Fig. S7) showed a xylosidase activity for F115_GH8 (P162_RS04035),
212 while also showing a shift of the entire ANTS-pattern to a lower molecular size, indicating D-
213 xylose release from BX and PPX hydrolysate. As F115_GH8 did not show activity towards
214 arabinoxylan, the L-arabinose side chains seem to hinder the activity of the xylosidase. Further
215 investigation via HPLC showed F115_GH8 to be able to cleave the arabinooligosaccharide
216 A2XX to A3X and D-xylose, specifying its activity to be a reducing-end xylosidase (Fig. S8).
217 In accordance with the specific upregulation of PUL II on arabinoxylans, F113_GH43
218 (P162_RS04025) could be shown to remove L-arabinose from xylooligosaccharides and WAX,
219 confirming the annotated α -L-arabinofuranosidase function (Fig. 3, Fig. S6). Additionally, the
220 enzyme was able to cleave off arabinose from a 2³- α -L-arabinofuranosyl-xylotriose (A2XX)
221 oligomer, as measured *via* HPLC (Fig. S8) and showed hydrolytic activity towards *p*NP- α -L-
222 arabinose (Fig. S9) proving it to be an α -L-arabinofuranosidase. This enlarges the capability of
223 the *Flavimarina sp.* PUL II enzymes towards further xylan-containing biomasses and shows
224 their high potential to convert structurally diverse xylans. Based on these findings, we postulate
225 the xylan degradation scheme presented in Figure 3.
226 Enzymatic assays with synthetic substrates revealed further activities of the PUL II-encoded
227 enzyme F112_GH97 which is annotated as an α -galactosidase (EC.3.2.1.131). F112_GH97

228 shows activity towards *p*NP- α -D-galactose (Fig. S9), indicating that the natural target
229 polysaccharide might also contain D-galactose units. Terrestrial hemicelluloses are mostly
230 described to have a 1,4- β -linked D-xylose backbone with an additional amount of different
231 sugar modifications which vary depending on the source [25]. Galactose is a common
232 modification, but D-galactose did not occur in high amounts in our investigated xylans (Fig.
233 S2), likely explaining to be the reason for FII2_GH97 inactivity on the chosen model substrates.
234 Possibly, galactosidase activity can be detected on galactose-rich substrates like
235 galactoarabinoxylan from oat [64] or the marine substrate arabinogalactan recently detected
236 in diatom blooms [24]. This overall diversity of the *Flavimarina sp.* CAZymes indicates an even
237 higher diversity of xylan structures present in marine ecosystems.

238

239 *Activities of the carbohydrate esterases underline PUL specificity*

240 Both xylan PULs encode carbohydrate esterases (CE) whose activities are known to increase
241 the solubility of hemicellulose polymers by cleavage of acetyl groups or phenolic acids. By this,
242 the polysaccharide backbone becomes more accessible to GHs, as side groups are removed
243 that cross-link different cell wall polymers of terrestrial origin [65, 66]. However, the role of the
244 CEs of marine origin is less explored.

245 PUL I contains the FI6_CE15 (P162_RS02365) esterase which is annotated as a CE15, a
246 family described as 4-*O*-methyl-glucuronyl methylesterase activity [67, 68]. The enzyme family
247 is known from wood degradation, where it removes the interlinkages between hemicelluloses
248 and lignin-like moieties [69, 70] indicating potential side groups and interlinkages within the
249 marine xylan substrates. This activity was proven for FI6_CE15 via hydrolytic activity towards
250 the model substrates benzyl-D-glucuronic acid and allyl-D-glucuronic acid (Fig. S10), as well
251 as *p*NP-acetate (Fig. S9). Therefore, this enzyme might be relevant in cleaving of the methyl
252 esters of xylan structures from other components of the cell (e.g., cell walls of macroalgae).
253 The additional CBM9 domain of the FI6_CE15 gene (Table S1), which is described as xylan-
254 targeting, can aid the binding of the polysaccharide [67, 68]. Especially the CBM9 family can

255 bind insoluble xylan polysaccharides and amorphous or crystalline cellulose-like compounds
256 [71]. PUL I contains one additional carbohydrate esterase, F18_CE6 (P162_RS02385), which
257 has a CE6 acetyl xylan esterase domain [67, 68]. The removal of *O*-acetylation was verified
258 via NMR-spectroscopy and an acetate assay (Fig. S10), releasing acetate from the model
259 substrate 6-*O*-acetyl-D-glucose. These two esterase activities occurring in PUL I seem to be
260 relevant in the degradation of glucuronoxylan-like polysaccharides, whose structures are
261 known from hardwood, which contains D-glucuronic acid side groups with methyl and
262 acetylation of the 1,4- β -linked D-xylose backbone [72]. The esterases should enable
263 *Flavimarina sp.* to access and degrade the acetylated glucuronoxylan for complete
264 saccharification.

265 In addition to PUL I, PUL II contains two multimodular esterases. The enzyme F111_GH43_10
266 (P162_RS04015) has a multimodular structure consisting of a CEnc domain, which is a
267 putative acetyl xylan esterase CE3 and an additional GH43_10 module, which is annotated as
268 a xylan 1,4- β -xylosidase/ α -L-arabinofuranosidase [67, 68]. This domain combination is similar
269 to an enzyme described in *Bacteroidetes eggerthii*, containing an esterase and a GH43 domain
270 relevant in arabinoxylan degradation [73]. The putative CE3 domain and GH43_10 domain
271 were separately analyzed here as F111A_CEnc and F111B_GH43_10 constructs. Only the
272 F111B_GH43_10 showed xylosidase activity in an ANTS-FACE analysis (Fig. S6) and by the
273 DNS-assay (Fig. 2) and no esterase activity was found. The encoded F114_CE6
274 (P162_RS04030) is an enzyme with two CE domains, a CE6 /acetyl xylan esterase and a non
275 characterized carbohydrate esterase module CEnc, which is likely to be a CE1/feruloyl
276 esterase (Table S1) [67, 68]. Ferulic acid xylan esterases target phenolic groups bound to the
277 L-arabinose moieties in arabinoxylans, which then enables α -L-arabinofuranosidases to further
278 degrade this polysaccharide type [73]. This activity was proven with a substantial amount of
279 ferulic acid released from the WAX-I substrate (Fig. S9). F114_CE6 has also an activity towards
280 *p*NP-acetate and released acetate from partially acetylated birchwood xylan (Fig. S10). A
281 related multimodular CE6|CE1 protein was found in *Bacteroides intestinalis* and was described
282 with a similar activity profile [73]. Both CE activities indicate specificity towards arabinoxylan-

283 like biomasses in removing phenolic esters and acetylation from xylans, thereby enabling PUL
284 II to complete the saccharification of arabinoxylan-like substrates.

285 It can be concluded from these esterase functions that the marine xylan targeted by
286 *Flavimarina* sp. is highly complex and owns diverse modifications similar to those found in
287 xylans from terrestrial sources. In macroalgal polysaccharides, acetylation patterns were
288 already found [74]. But the overall enzyme activities and modularity in our model organism
289 lead us to assume that it is prepared for a much higher xylose-containing polysaccharide
290 diversity.

291

292 *Ecological relevance of the xylan degradation pathways and PUL-architecture*

293 The biochemical analyses in this study proved *Flavimarina* sp. to be a proficient degrader of a
294 multitude of xylans. This ability is largely facilitated by the promiscuous activity of multiple PUL-
295 associated GH10s on xylan backbones containing various modifications. We therefore conclude
296 that the presence of a GH10 with additional processing enzymes is essential to enable the
297 degradation of complex xylans. Following this premise, we identified 226 PULs in marine,
298 terrestrial and human digestive system-associated databases, sharing at least one GH10 and
299 two additional CAZymes with either of the *Flavimarina* sp. PULs, indicating complex xylans to
300 be important to diverse habitats (Fig. S11). A modularity search revealed 81 different xylan-
301 PUL architectures, most commonly containing only a few shared enzymes with *Flavimarina*
302 sp. (Fig. 4a). The most widely distributed PUL type contains only the enzymes known from the
303 *Flavimarina* sp. PUL I, which are GH43 and GH67 alongside a GH10 hydrolase. This can
304 therefore be classified as the core set of CAZymes required for xylan degradation. In some
305 cases, bacteria containing such a PUL also contain a second, more complex PUL, such as the
306 terrestrial bacterium *Flavobacterium johnsonia*, which likely facilitates niche-specific
307 adaptations (Fig. 4b).

308 The high overall PUL diversity points towards the possibility of degrading a multitude of
309 different xylan-containing substrates that require specific CAZyme repertoires. As most

310 identified PULs showed no distinct separation between CAZymes homologous to *Flavimarina*
311 sp. PUL I and PUL II, it can be assumed that this system is a niche adaptation towards varying
312 complex xylans in marine habitats. In fact, *Fibrisoma limi* and *Fibrella aestuarina*, the only other
313 species shown to have two enlarged *Flavimarina*-like PULs, are also associated with marine
314 ecosystems (Fig. 4b). Some marine bacteria as well as human gut symbionts possess the
315 entire CAZyme spectrum of the *Flavimarina* sp. PULs consolidated into a single PUL (Fig. 4b),
316 underlining that complex xylans are important to diverse ecosystems. It has to be noted that
317 many of the similar identified PULs also encode for additional CAZymes associated with
318 arabinose-containing polysaccharide degradation, such as GH39, GH51 or GH146.
319 *Flavimarina* sp. does not encode for CAZymes of these families within its two main xylan-
320 targeting PULs. However, some of these GHs are encoded in the genome and associated with
321 the arabinofuranosidase-containing PUL that was upregulated during growth on RAX and
322 WAX. This further indicates that *Flavimarina* sp. possesses the genetic potential to degrade a
323 multitude of diverse xylans.

324

325 **Discussion**

326 This study offers comprehensive insights into xylan-specific metabolic pathways of a marine
327 *Bacteroidetes* strain. The model organism *Flavimarina* sp. Hel_I_48 contains two separate
328 xylan PULs, which target glucuronoxylans (PUL I) and arabinoxylans (PUL II). This is the first
329 study describing comprehensive marine enzymatic cascades for xylan utilization and the
330 modularity of xylan degradation by different genomic regions in marine bacteria. The general
331 occurrence and abundance of multiple xylan PULs in *Bacteroidetes* was previously described
332 for the gut microbiota [28]. *Bacteroidetes ovatus* contains numerous PULs with two xylan-
333 targeting PULs showing xylan-type specialization [28]. These PULs consist of multiple
334 CAZymes showing different activities towards the chosen model substrates, indicating a more
335 general activity of the initial xylanases to degrade β -1,4-linked xylan structures generating
336 oligomers which can be imported into the cell. These distinct oligosaccharides can then be

337 further degraded by more specialized intracellular glycoside hydrolases and carbohydrate
338 esterases, removing motifs of acetylation, phenolic esters, L-arabinose and D-glucuronic acid
339 decorations from the xylan backbones. The substrates used for the characterization of
340 *Flavimarina* sp. Hel_I_48 PUL-encoded enzymes were of terrestrial and macroalgal origin,
341 serving as model substrates for different xylan types covering the complexity of the potential
342 target polysaccharides. Even though there are some microbes adapted to the degradation of
343 terrestrial biomasses in the ocean, it is unlikely that the enzymes from *Flavimarina* sp. mainly
344 target these polysaccharides, as only small amounts of terrestrial organic carbon can be found
345 in marine dissolved organic carbon (DOC) [75]. This bacterium, which was isolated during a
346 phytoplankton bloom [32], likely rather targets marine carbohydrates from macro- and
347 microalgal origin as well as microbially produced EPS matrices [76]. It was shown that marine
348 algae and terrestrial plants share a convergent evolutionary history towards their cell wall
349 structures [77], illustrated by lignin-like compounds found in red and green seaweed [77]. Thus,
350 polysaccharides orthologous to terrestrial biomass occur in the marine environment [14, 15,
351 78]. Furthermore, it was shown that glucuronoxylans and 1,4- β -xylan occur in diatom blooms
352 around the north-sea island of Helgoland [24]. Thus, it is most likely that those microalgae
353 contain xylan substrates which are targeted by our model organism *Flavimarina* sp. and other
354 related marine *Bacteroidetes* with a similar enzyme repertoire. The ecological relevance of
355 these xylan utilization pathways is further reflected by the frequently detectable conserved sets
356 of genes encoding related xylan-degrading enzymes of other polysaccharide-utilizing marine
357 *Bacteroidetes*. Our in-depth functional analysis of xylan utilization strategies from the
358 phytoplankton bloom-associated isolate *Flavimarina* sp. further indicates that marine algal
359 xylans might be heterogenous, with glucurono- and arabino-side groups and acetylation
360 patterns. This assumption is further supported by the multiple SusD proteins encoded in the
361 xylan PULs which indicate a versatile xylan backbone of marine algae. We suggest that the
362 observed CAZyme modularity and substrate flexibility may enable these marine bacteria to
363 consume this diverse range of algal xylan structures abundantly available in marine habitats.
364

365 **Experimental procedures**

366 *Bioinformatics and comparative genomics*

367 Databases were created using NCBI RefSeq assemblies [33] of prokaryotic genomes stored
368 in the RefSoilv1 database [34], MarRefv1.7 and MarDBv1.6 [35] as well as the NIH Human
369 Microbiome Project [36, 37] catalog with isolation body site „gastrointestinal_tract“. Genomes
370 were screened for *Flavimarina* sp. Hel_I_48 PUL I- and PUL II-like gene clusters with the
371 “hmm” search function of cblaster [38] v1.3.14 (default settings) using the HMM profiles
372 “GH67.hmm”, “GH115.hmm”, “CE15.hmm”, “GH43_1.hmm” (all PUL I), “GH43_10.hmm”,
373 “GH97.hmm”, “GH43_12.hmm”, “CE6.hmm”, “GH8.hmm”, “GH95.hmm” (all PUL II) and
374 “GH10.hmm” as marker profile from the dbCAN-HMMdb-V10.hmm database [39]. CAZyme
375 context of clusters encoding a GH10 with at least two other glycoside hydrolases from PUL I
376 and/or PUL II was predicted using the hmmscan function of HMMer v3.3.2 [40] against the
377 dbCAN-HMMdb-V10.hmm database. Hits were filtered using the hmmscan-parser.sh script
378 from dbCAN and validated using Protein-Protein BLAST (v2.11.0+) [41] against CAZyDB
379 (release 09242021) with an e-value threshold of E-20, a minimum sequence identity of 30%
380 and a query coverage of at least 40% [42]. The resulting gene clusters were visualized with
381 UpSetR [43, 44] and RIdeogram [45]. For the phylogenetic tree, *rpoB* genes were aligned using
382 the ClustalW [46] web service in “slow/accurate” mode and default settings. Maximum-
383 likelihood phylogenies were estimated by the PhyML 3.0 web server [47] with default settings
384 and visualized with iTOL [48].

385 The *Flavimarina* sp. Hel_I_48 PUL repertoire was annotated as described above using
386 additionally the TIGRFAM profile “TIGR04056.hmm” for prediction of SusC-like proteins and
387 the PFAM models “PF07980.11.hmm”, “PF12741.7.hmm”, “PF14322.6.hmm” or
388 “PF12771.7.hmm” for prediction of SusD-like proteins. Final PULs were predicted as described
389 previously [22], excluding sulfatase-encoding genes and visualized using Circos [49].

390

391

392

393 *Proteome analysis*

394 *Flavimarina* sp. Hel_I_48 was grown to the late exponential phase in modified MPM medium
395 [50] containing 0.1% beechwood xylan (BX), *Palmaria palmata* xylan (PPX), *Caulerpa prolifera*
396 xylan (CPX), rye arabinoxylan (RAX), and wheat arabinoxylan (WAX), or apple pectin (Pec) as
397 sole carbon sources. Triplicates of 50 mL cultures were harvested at 4,000 x *g*, 20 min and 4
398 °C. Cells were resuspended in 50 mM TEAB buffer containing 4% SDS and incubated at 95
399 °C and 600 rpm for 5 min (Thermomixer C, Eppendorf, Hamburg, Deutschland). Samples were
400 cooled to room temperature before being placed in an ultrasonic bath for 5 min. Cell debris
401 was removed *via* centrifugation at 14.000 x *g* and 4 °C for 10 min.

402 Protein concentration was measured using the BCA Pierce Protein assay kit (Thermo Fisher
403 Scientific Inc., Waltham, MA, USA). 25 µg protein per sample was loaded on a 10% 1D-SDS
404 polyacrylamide gel and separated at 120 V for 90 min. Gels were fixed with a 40% ethanol/
405 10% acetic acid solution and stained overnight using Coomassie G-250 [51]. Each sample was
406 divided into 10 subsamples, de-stained using 30% acetonitrile in 200 mM (NH₄)₂CO₃ and
407 digested for 16 h using trypsin.

408 Peptides were separated as described previously [52] by reverse phase C18 column
409 chromatography on a nano ACQUITY-UPLC (Waters Corporation, Milford, MA, USA) online-
410 coupled to an LTQ-Orbitrap Classic mass spectrometer (Thermo Fisher Scientific Inc.,
411 Waltham, MA, USA). Spectra were searched against a target-decoy protein sequence
412 database including sequences and reverse sequences of *Flavimarina* sp. Hel_I_48 and
413 common laboratory contaminants using MaxQuant [53]. Only proteins that could be detected
414 in at least two out of three replicates were considered identified. Relative protein abundance
415 values in % of all proteins in the same sample were manually calculated from iBAQ values as
416 %riBAQ (relative intensity based absolute quantification). Data and results are available via
417 the PRIDE partner repository [54] with identifier PXD033600. Reviewer access:
418 reviewer_pxd033600@ebi.ac.uk; password: QUOG4Lca.

419

420

421 *Gene cloning and enzyme production*

422 Expression constructs of FI1_GH67, FI2_GH10, FI4_GH10, FI5_hyp, FI7_GH43_1, FI8_CE6
423 and FI9_hyp (Table S1) were prepared using the FastCloning strategy (Table S2) [55] with
424 genomic DNA from *Flavimarina* sp. Hel_I_48 as template for the amplification of the inserts.
425 The genomic DNA was extracted as described previously [56]. The pET28 constructs were
426 prepared as described recently [9].

427 *Escherichia coli* BL21(DE3) was transformed with pET28-based plasmids (expression
428 constructs and gene cloning described in the supplementary information) harboring the
429 required genes. For the overexpression, 50 mL LB or TB medium with 100 $\mu\text{g mL}^{-1}$ kanamycin
430 were inoculated from an overnight culture in LB containing 50 $\mu\text{g mL}^{-1}$ kanamycin. The culture
431 was grown at 37 °C and 180 rpm until optical density at 600 nm reached 0.8. Expression was
432 then induced by adding 0.5 or 1 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) and the
433 culture was cooled to 20 °C and incubated for 24 h.

434 Plasmids containing the genes encoding for SusD-like proteins were introduced into chemo-
435 competent *E. coli* BL21(DE3) cells. 4 mL of overnight culture in LB media was used to inoculate
436 600 mL TB media both containing 50 $\mu\text{g mL}^{-1}$ kanamycin. The cells were grown to an optical
437 density at 600 nm (OD_{600}) between 1-1.5 at 37 °C and 180 rpm. Temperature was lowered to
438 20 °C and IPTG was added to a final concentration of 1 mM as the OD_{600} reached 2-3. Cells
439 were harvested after 16 h by centrifugation at 4,000 $\times g$ and 4 °C for 20 min including washing
440 with 20 mM sodium phosphate, 500 mM sodium chloride at pH 8. Washed pellets were flash
441 frozen in liquid nitrogen and stored at -20 °C until purification. Enzyme purification is described
442 in the supplementary information.

443

444 *Purification of xylan*

445 *P. palmata* dulse was purchased at Algenladen (Gießen, Germany). After milling the dry algae
446 (25 g) biomass, it was extracted two times with dH_2O (1 L) for 2 h at 70 °C [21]. Afterwards,
447 the solid particles were removed and the water content reduced to a viscous consistency. The
448 polysaccharide was precipitated by adding four volumes of cold ethanol. The precipitate was

449 then separated from the ethanol fraction. Afterwards the alcohol insoluble fraction was
450 dissolved in deionized water and freeze-dried. *C. prolifera* was extracted with the same
451 protocol, this algal material was provided by the Ozeaneum (Stralsund, Germany).
452 Monosaccharide composition of the self-extracted polysaccharides, charge and size of the
453 used polymers were analyzed (Tables S3 & S4). The xylan degradation products from
454 enzymatic reactions and the conversion of purified oligomers were analyzed by FACE.
455 Untreated xylan was used generally at a concentration of 1 g L⁻¹, while purified sugar oligomers
456 were used at 0.25 mg mL⁻¹. Incubation with the enzymes was performed overnight at room
457 temperature. Additional enzymatic assays are described in the supplementary information.

458

459 *Determination of reducing ends (DNS-assay)*

460 The dinitrosalicylic acid-assay (DNS-assay) from Bernfeld et al. was used to determine the
461 reducing ends of the carbohydrates [57]. A 20 µL reaction sample of the biocatalysis and 20
462 µL of the color reagent were combined and incubated at 100 °C for 15 min. After the samples
463 were cooled down to room temperature, 180 µL of water was added, and the 200 µL were
464 transferred to a microtiter plate to measure the absorption at 540 nm in a plate reader (Infinite
465 M200 Pro, Tecan Group, Swiss).

466

467 *Fluorophore-assisted carbohydrate electrophoresis*

468 Fluorophore-assisted carbohydrate electrophoresis (FACE) was performed with 8-
469 aminonaphthalene-1,3,6-trisulfonic acid (ANTS) as fluorophore adapted from Reisky et al. [9].
470 10 µL aliquots of the biocatalysis reaction were lyophilized and dissolved in 4 µL of ANTS (0.05
471 M in DMSO with 15% acetic acid) solution and 4 µL of NaCNBH₃ (1 M in DMSO) solution. In
472 case of the xylan experiment, AMAC was replaced by ANTS (0.2 M in water). After incubation
473 at 37 °C overnight in the dark, the samples were mixed with 20 µL loading buffer and 4 µL was
474 loaded to a FACE-gel [58].

475

476

477 *HPLC determination of oligosaccharide degradation products*

478 HPLC analysis for the determination of xylan oligosaccharide standards was performed using
479 a Chrommaster HPLC system from Hitachi (equipped with a Hitachi Chrommaster 5310
480 column oven) and a detector (Hitachi Chrommaster 5450 RI detector). Analyses were
481 performed with a mobile phase consisting of H₂O with 0.005 M H₂SO₄ on a styrene/polyvinyl
482 benzene copolymer column (SugarSep-H 10 µm 300 x 8 mm) at 70 °C for 20 min. The flow
483 rate was set to 0.5 mL/min.

484 **Acknowledgments**

485 We thank the German Research Foundation (DFG) for funding through the Research Unit FOR
486 2406 “Proteogenomics of Marine Polysaccharide Utilization” (POMPU) by grants to U.T.B. (BO
487 1862/17-2), J.-H.H. (HE 7217/2-2), M. H. (HO 4754/5-2) and T.S. (SCHW 595/10-2). Especially
488 we wish to thank the Ozeaneum Stralsund, which kindly provided the algae *C. prolifera* sp.
489 cultures.

490

491 **Conflict of interest**

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

493 **References**

- 494 1. Field CB. Primary production of the biosphere: Integrating terrestrial and oceanic
495 components. *Science* 1998; **281**: 237–240.
- 496 2. Kraan S. Algal polysaccharides, novel applications and outlook. *Carbohydr - Compr Stud*
497 *Glycobiol Glycotechnol* 2012; 1–44.
- 498 3. Becker S, Tebben J, Coffinet S, Wiltshire K, Iversen MH, Harder T, et al. Laminarin is a
499 major molecule in the marine carbon cycle. *Proc Natl Acad Sci* 2020; **117**: 6599–6607.

- 500 4. Kloareg B, Quatrano RS. Structure of the cell walls of marine algae and ecophysiological
501 functions of the matrix polysaccharides. *Oceanogr Mar Biol Annu Rev* 1988; **26**: 259–
502 315.
- 503 5. Hoagland KD, Rosowski JR, Gretz MR, Roemer SC. Diatom extracellular polymeric
504 substances: functions, fine structure, chemistry, and physiology. *J Phycol* 1993; **29**: 537–
505 566.
- 506 6. Arnosti C. Microbial extracellular enzymes and the marine carbon cycle. *Annu Rev Mar*
507 *Sci* 2011; **3**: 401–425.
- 508 7. McKee LS, La Rosa SL, Westereng B, Eijssink VG, Pope PB, Larsbrink J. Polysaccharide
509 degradation by the *Bacteroidetes*: mechanisms and nomenclature. *Environ Microbiol Rep*
510 2021; **13**: 559–581.
- 511 8. Martens EC, Koropatkin NM, Smith TJ, Gordon JI. Complex glycan catabolism by the
512 human gut microbiota: the *Bacteroidetes* Sus-like paradigm. *J Biol Chem* 2009; **284**:
513 24673–24677.
- 514 9. Reisky L, Préchoux A, Zühlke M-K, Bäumgen M, Robb CS, Gerlach N, et al. A marine
515 bacterial enzymatic cascade degrades the algal polysaccharide ulvan. *Nat Chem Biol*
516 2019; **15**: 803–812.
- 517 10. Reisky L, Stanetty C, Mihovilovic MD, Schweder T, Hehemann J-H, Bornscheuer UT.
518 Biochemical characterization of an ulvan lyase from the marine flavobacterium *Formosa*
519 *agariphila* KMM 3901T. *Appl Microbiol Biotechnol* 2018; **102**: 6987–6996.
- 520 11. Hehemann J-H, Boraston AB, Czjzek M. A sweet new wave: structures and mechanisms
521 of enzymes that digest polysaccharides from marine algae. *Curr Opin Struct Biol* 2014;
522 **28**: 77–86.
- 523 12. Hsieh Y, J. Harris P. Xylans of red and green algae: What is known about their structures
524 and how they are synthesised? *Polymers* 2019; **11**: 354.
- 525 13. Ebringerová A. Structural diversity and application potential of hemicelluloses. *Macromol*
526 *Symp* 2005; **232**: 1–12.

- 527 14. Popper ZA, Michel G, Hervé C, Domozych DS, Willats WGT, Tuohy MG, et al. Evolution
528 and diversity of plant cell walls: From algae to flowering plants. *Annu Rev Plant Biol* 2011;
529 **62**: 567–590.
- 530 15. Jensen JK, Busse-Wicher M, Poulsen CP, Fangel JU, Smith PJ, Yang J-Y, et al.
531 Identification of an algal xylan synthase indicates that there is functional orthology
532 between algal and plant cell wall biosynthesis. *New Phytol* 2018; **218**: 1049–1060.
- 533 16. Mackie IM, Percival E. The constitution of xylan from the green seaweed *Caulerpa*
534 *filiformis*. *J Chem Soc Resumed* 1959; 1151–1156.
- 535 17. Lahaye M, Rondeau-Mouro C, Deniaud E, Buléon A. Solid-state ¹³C NMR spectroscopy
536 studies of xylans in the cell wall of *Palmaria palmata*. *Carbohydr Res* 2003; **338**: 1559–
537 1569.
- 538 18. Percival EGV, Chanda SK. The Xylan of *Rhododymenia palmata*. *Nature* 1950; **166**: 787–
539 788.
- 540 19. Turvey JR, Williams EL. The structures of some xylans from red algae. *Phytochemistry*
541 1970; **9**: 2383–2388.
- 542 20. Cerezo AS. The fine structure of *Chaetangium fastigiatum* xylan: studies of the sequence
543 and configuration of the (1→3)-linkages. *Carbohydr Res* 1972; **22**: 209–211.
- 544 21. Deniaud E, Quemener B, Fleurence J, Lahaye M. Structural studies of the mix-linked β-
545 (1→3)/β-(1→4)-d-xylans from the cell wall of *Palmaria palmata* (Rhodophyta). *Int J Biol*
546 *Macromol* 2003; **33**: 9–18.
- 547 22. Francis TB, Bartosik D, Sura T, Sichert A, Hehemann J-H, Markert S, et al. Changing
548 expression patterns of TonB-dependent transporters suggest shifts in polysaccharide
549 consumption over the course of a spring phytoplankton bloom. *ISME J* 2021; **15**: 2336–
550 2350.
- 551 23. Huang G, Vidal-Melgosa S, Sichert A, Becker S, Fang Y, Niggemann J, et al. Secretion
552 of sulfated fucans by diatoms may contribute to marine aggregate formation. *Limnol*
553 *Oceanogr* 2021; **66**: 3768–3782.

- 554 24. Vidal-Melgosa S, Sichert A, Francis TB, Bartosik D, Niggemann J, Wichels A, et al.
555 Diatom fucan polysaccharide precipitates carbon during algal blooms. *Nat Commun*
556 2021; **12**: 1150.
- 557 25. Bastawde KB. Xylan structure, microbial xylanases, and their mode of action. *World J*
558 *Microbiol Biotechnol* 1992; **8**: 353–368.
- 559 26. Uffen RL. Xylan degradation: a glimpse at microbial diversity. *J Ind Microbiol Biotechnol*
560 1997; **19**: 1–6.
- 561 27. Dodd D, Cann IKO. Enzymatic deconstruction of xylan for biofuel production. *GCB*
562 *Bioenergy* 2009; **1**: 2–17.
- 563 28. Rogowski A, Briggs JA, Mortimer JC, Tryfona T, Terrapon N, Lowe EC, et al. Glycan
564 complexity dictates microbial resource allocation in the large intestine. *Nat Commun*
565 2015; **6**: 7481.
- 566 29. Malgas S, Mafa MS, Mkabayi L, Pletschke BI. A mini review of xylanolytic enzymes with
567 regards to their synergistic interactions during hetero-xylan degradation. *World J*
568 *Microbiol Biotechnol* 2019; **35**: 187.
- 569 30. Ruile P, Winterhalter C, Liebl W. Isolation and analysis of a gene encoding α -
570 glucuronidase, an enzyme with a novel primary structure involved in the breakdown of
571 xylan. *Mol Microbiol* 1997; **23**: 267–279.
- 572 31. Araki T, Inoue N, Morishita T. Purification and characterization of β -1,3-xylanase from a
573 marine bacterium, *Alcaligenes* sp. XY-234. *J Gen Appl Microbiol* 1998; **44**: 269–274.
- 574 32. Kappelmann L, Krüger K, Hehemann J-H, Harder J, Markert S, Unfried F, et al.
575 Polysaccharide utilization loci of North Sea Flavobacteriia as basis for using SusC/D-
576 protein expression for predicting major phytoplankton glycans. *ISME J* 2019; **13**: 76–91.
- 577 33. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database
578 resources of the national center for biotechnology information. *Nucleic Acids Res* 2022;
579 **50**: D20–D26.
- 580 34. Jinlyung Choi. RefSoil Database. 2017. figshare. , 4477985690 Bytes

- 581 35. Klemetsen T, Raknes IA, Fu J, Agafonov A, Balasundaram SV, Tartari G, et al. The MAR
582 databases: development and implementation of databases specific for marine
583 metagenomics. *Nucleic Acids Res* 2018; **46**: D692–D699.
- 584 36. The Human Microbiome Project Consortium. A framework for human microbiome
585 research. *Nature* 2012; **486**: 215–221.
- 586 37. The Human Microbiome Project Consortium. Structure, function and diversity of the
587 healthy human microbiome. *Nature* 2012; **486**: 207–214.
- 588 38. Gilchrist CLM, Booth TJ, van Wersch B, van Grieken L, Medema MH, Chooi Y-H. cblaster:
589 a remote search tool for rapid identification and visualization of homologous gene
590 clusters. *Bioinforma Adv* 2021; **1**: vbab016.
- 591 39. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, et al. dbCAN2: a meta server for
592 automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 2018; **46**: W95–
593 W101.
- 594 40. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity
595 searching. *Nucleic Acids Res* 2011; **39**: W29–W37.
- 596 41. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST
597 and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids*
598 *Res* 1997; **25**: 3389–3402.
- 599 42. Krüger K, Chafee M, Ben Francis T, Glavina del Rio T, Becher D, Schweder T, et al. In
600 marine *Bacteroidetes* the bulk of glycan degradation during algae blooms is mediated by
601 few clades using a restricted set of genes. *ISME J* 2019; **13**: 2800–2816.
- 602 43. Lex A, Gehlenborg N, Strobel H, Vuillemot R, Pfister H. UpSet: Visualization of
603 intersecting sets. *IEEE Trans Vis Comput Graph* 2014; **20**: 1983–1992.
- 604 44. Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of
605 intersecting sets and their properties. *Bioinformatics* 2017; **33**: 2938–2940.
- 606 45. Hao Z, Lv D, Ge Y, Shi J, Weijers D, Yu G, et al. *Rldeogram*: drawing SVG graphics to
607 visualize and map genome-wide data on the idiograms. *PeerJ Comput Sci* 2020; **6**: e251.

- 608 46. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of
609 progressive multiple sequence alignment through sequence weighting, position-specific
610 gap penalties and weight matrix choice. *Nucleic Acids Res* 1994; **22**: 4673–4680.
- 611 47. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms
612 and methods to estimate maximum-likelihood phylogenies: assessing the performance
613 of PhyML 3.0. *Syst Biol* 2010; **59**: 307–321.
- 614 48. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree
615 display and annotation. *Nucleic Acids Res* 2021; **49**: W293–W296.
- 616 49. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an
617 information aesthetic for comparative genomics. *Genome Res* 2009; **19**: 1639–1645.
- 618 50. Schut F, de Vries EJ, Gottschal JC, Robertson BR, Harder W, Prins RA, et al. Isolation
619 of typical marine bacteria by dilution culture: growth, maintenance, and characteristics of
620 isolates under laboratory conditions. *Appl Environ Microbiol* 1993; **59**: 2150–2160.
- 621 51. Candiano G, Bruschi M, Musante L, Santucci L, Ghiggeri GM, Carnemolla B, et al. Blue
622 silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis.
623 *Electrophoresis* 2004; **25**: 1327–1333.
- 624 52. Otto A, Bernhardt J, Meyer H, Schaffer M, Herbst F-A, Siebourg J, et al. Systems-wide
625 temporal proteomic profiling in glucose-starved *Bacillus subtilis*. *Nat Commun* 2010; **1**:
626 137.
- 627 53. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-
628 range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 2008;
629 **26**: 1367–1372.
- 630 54. Perez-Riverol Y, Bai J, Bandla C, García-Seisdedos D, Hewapathirana S,
631 Kamatchinathan S, et al. The PRIDE database resources in 2022: a hub for mass
632 spectrometry-based proteomics evidences. *Nucleic Acids Res* 2021; **50**: D543–D552.
- 633 55. Li C, Wen A, Shen B, Lu J, Huang Y, Chang Y. FastCloning: a highly simplified,
634 purification-free, sequence- and ligation-independent PCR cloning method. *BMC*
635 *Biotechnol* 2011; **11**: 92.

- 636 56. Chen WP, Kuo TT. Genomic DNA isolation from EPS-producing gram negative bacteria.
637 *Nucleic Acids Res* 1993; **21**: 2260.
- 638 57. Bernfeld P. Amylases, α and β . *Methods in enzymology*. 1955. Academic Press, pp 149–
639 158.
- 640 58. Hehemann J-H, Correc G, Barbeyron T, Helbert W, Czjzek M, Michel G. Transfer of
641 carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature*
642 2010; **464**: 908–912.
- 643 59. Mystkowska AA, Robb C, Vidal-Melgosa S, Vanni C, Fernandez-Guerra A, Höhne M, et
644 al. Molecular recognition of the β -glucans laminarin and pustulan by a SusD-like glycan-
645 binding protein of a marine *Bacteroidetes*. *FEBS J* 2018; **285**: 4465–4481.
- 646 60. Bryan JK. Molecular weights of protein multimers from polyacrylamide gel
647 electrophoresis. *Anal Biochem* 1977; **78**: 513–519.
- 648 61. Glenwright AJ, Pothula KR, Bhamidimarri SP, Chorev DS, Baslé A, Firbank SJ, et al.
649 Structural basis for nutrient acquisition by dominant members of the human gut
650 microbiota. *Nature* 2017; **541**: 407–411.
- 651 62. Gray DA, White JBR, Oluwole AO, Rath P, Glenwright AJ, Mazur A, et al. Insights into
652 SusCD-mediated glycan import by a prominent gut symbiont. *Nat Commun* 2021; **12**: 44.
- 653 63. Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, et
654 al. SignalP 6.0 predicts all five types of signal peptides using protein language models.
655 *Nat Biotechnol* 2022; 1023–1025.
- 656 64. Buchala AJ, Fraser CG, Wilkie KCB. An acidic galactoarabinoxylan from the stem of
657 *Avena sativa*. *Phytochemistry* 1972; **11**: 2803–2814.
- 658 65. Williamson G, Kroon P, Faulds C. Hairy plant polysaccharides: A close shave with
659 microbial esterases. *Microbiol Read Engl* 1998; **144(Pt 8)**: 2011–23.
- 660 66. Hettiarachchi SA, Kwon Y-K, Lee Y, Jo E, Eom T-Y, Kang Y-H, et al. Characterization of
661 an acetyl xylan esterase from the marine bacterium *Ochrovirga pacifica* and its synergism
662 with xylanase on beechwood xylan. *Microb Cell Factories* 2019; **18**: 122.

- 663 67. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The
664 carbohydrate-active enzymes database (CAZy): an expert resource for glycogenomics.
665 *Nucleic Acids Res* 2009; **37**: D233–D238.
- 666 68. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-
667 active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 2014; **42**: D490-495.
- 668 69. De Santi C, Gani OA, Helland R, Williamson A. Structural insight into a CE15 esterase
669 from the marine bacterial metagenome. *Sci Rep* 2017; **7**: 17278.
- 670 70. Charavgi MD, Dimarogona M, Topakas E, Christakopoulos P, Chrysinia ED. The structure
671 of a novel glucuronoyl esterase from *Myceliophthora thermophila* gives new insights into
672 its role as a potential biocatalyst. *Acta Crystallogr D Biol Crystallogr* 2013; **69**: 63–73.
- 673 71. Boraston AB, Creagh AL, Alam MdM, Kormos JM, Tomme P, Haynes CA, et al. Binding
674 specificity and thermodynamics of a family 9 carbohydrate-binding module from
675 *Thermotoga maritima* xylanase 10A. *Biochemistry* 2001; **40**: 6240–6247.
- 676 72. Kmezik C, Krska D, Mazurkewich S, Larsbrink J. Characterization of a novel multidomain
677 CE15-GH8 enzyme encoded by a polysaccharide utilization locus in the human gut
678 bacterium *Bacteroides eggerthii*. *Sci Rep* 2021; **11**: 17662.
- 679 73. Pereira GV, Abdel-Hamid AM, Dutta S, D'Alessandro-Gabazza CN, Wefers D, Farris JA,
680 et al. Degradation of complex arabinoxylans by human colonic *Bacteroidetes*. *Nat*
681 *Commun* 2021; **12**: 459.
- 682 74. Wong T, Brault L, Gasparotto E, Vallée R, Morvan P-Y, Ferrières V, et al. Formation of
683 amphiphilic molecules from the most common marine polysaccharides, toward a
684 sustainable alternative? *Molecules* 2021; **26**: 4445.
- 685 75. Bianchi TS. The role of terrestrially derived organic carbon in the coastal ocean: A
686 changing paradigm and the priming effect. *Proc Natl Acad Sci* 2011; **108**: 19473–19481.
- 687 76. Bäumgen M, Dutschei T, Bornscheuer UT. Marine polysaccharides: Occurrence,
688 enzymatic degradation and utilization. *ChemBioChem* 2021; **22**: 2247–2256.

- 689 77. Martone PT, Estevez JM, Lu F, Ruel K, Denny MW, Somerville C, et al. Discovery of
690 lignin in seaweed reveals convergent evolution of cell-wall architecture. *Curr Biol* 2009;
691 **19**: 169–175.
- 692 78. Popper ZA, Tuohy MG. Beyond the green: Understanding the evolutionary puzzle of plant
693 and algal cell walls. *Plant Physiol* 2010; **153**: 373–383.

694
695

696 **Figure legends**

697 **Fig. 1.** PULs and proteomic profiles of *Flavimarina* sp. Hel_I_48. The bacterium was cultivated
698 on different xylose-containing substrates as sole carbon source and with pectin (apple) as
699 control (see Fig. S1). Shown are the abundances (% riBAQ) of all PUL-encoded proteins on
700 each substrate (see Table S1) with darker blue color indicating a higher abundance. Proteins
701 encoded in either of the two xylan-PULs are annotated on the outside of the ring.

702

703 **Fig. 2.** PUL organization and initial xylan degradation of *Flavimarina* sp. Hel_I_48. The xylan
704 PULs of *Flavimarina* sp. Hel_I_48 were defined by Kappelmann et al. (2018). Genome loci
705 refer to the RefSeq assembly (GCF_000733945.1) of PUL I (P162_RS02310-RS02395) and
706 PUL II (P162_RS04015-RS04080). Enzyme activity was determined on the following
707 polysaccharides: beechwood xylan (BX), *P. palmata* xylan (PPX), rye arabinoxylan (RAX),
708 wheat arabinoxylan of medium viscosity (WAX-M) and insoluble wheat arabinoxylan (WAX-I).
709 Analyses were performed with the DNS reducing end assay (Fig. S5), after incubation with 15
710 µg/mL enzyme overnight.

711

712 **Fig. 3.** The specific xylan degradation targets of PUL I and PUL II. The carbohydrate structure
713 in PUL I and PUL II is assumed to be similar to glucuronoxylan and arabinoxylan as known
714 from grains respectively [28]. Summary of the initial degradation activity (Fig. 2, Figs. S4, S5,
715 S10) (a) from the GH10 xylanases (FI2_GH10, FI4_GH10 and FII8_GH10) and the esterases

716 FI8_CE6 cleaving of acetate moieties, FI4_CE6 cleaving acetate and ferulic acid modifications
717 and FI6_CE15 cleaving of 4-O-methyl-glucuronyl methylesters. The resulting
718 xylooligosaccharides can be further degraded by PULs specific enzyme activities (b)
719 glucuronidases for PUL I and arabinases and galactosidases for PUL II. The undecorated
720 xylooligosaccharides can then be further decomposed by xylosidases (c). Exemplary
721 enzymatic proof of each degradation step is either shown via HPLC results (Fig. S8) or
722 hydrolytic activity on *p*-NP-sugar substrates (Fig. S9).

723

724 **Fig. 4.** Xylan PUL modularity of bacteria from different environments. (a) Upset plot showing
725 different xylan PUL compositions as well as their prevalence in different habitats. (b)
726 Comparison of *Flavimarina* sp. PULs to other significant cluster types, such as those (1)
727 containing a second, more complex PUL, (2) all CAZymes also encoded for by *Flavimarina* sp.
728 or (3) all CAZymes but encoded in a single larger PUL. CAZymes of *Flavimarina* sp. PUL I are
729 depicted in purple, those of PUL II in green. CAZymes of families or subfamilies not encoded
730 by either of the *Flavimarina* sp. PULs are marked in orange. See also the complementary
731 phylogenetic tree in Fig. S11 of the supplementary information.

Figure 1

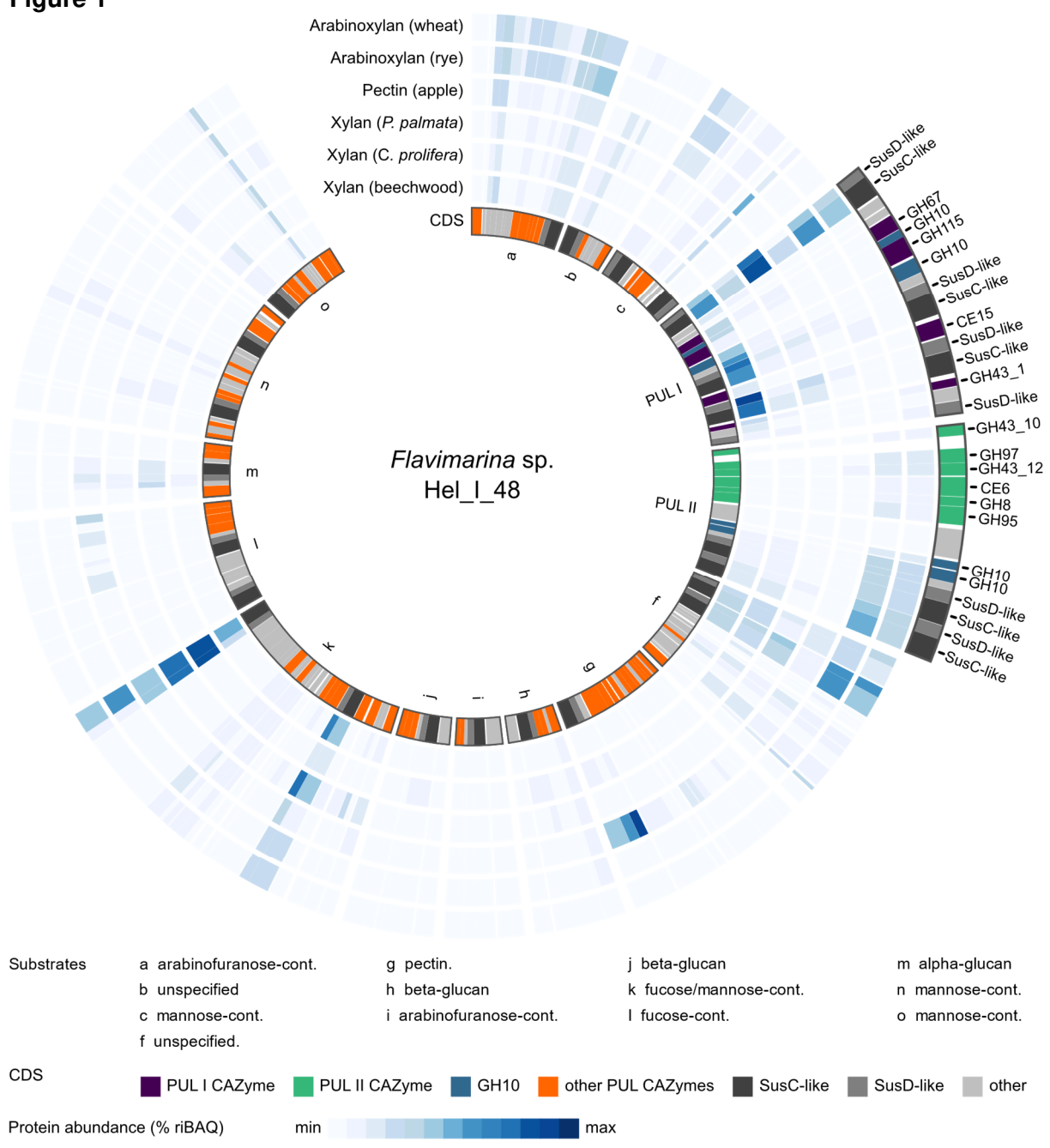


Figure 2

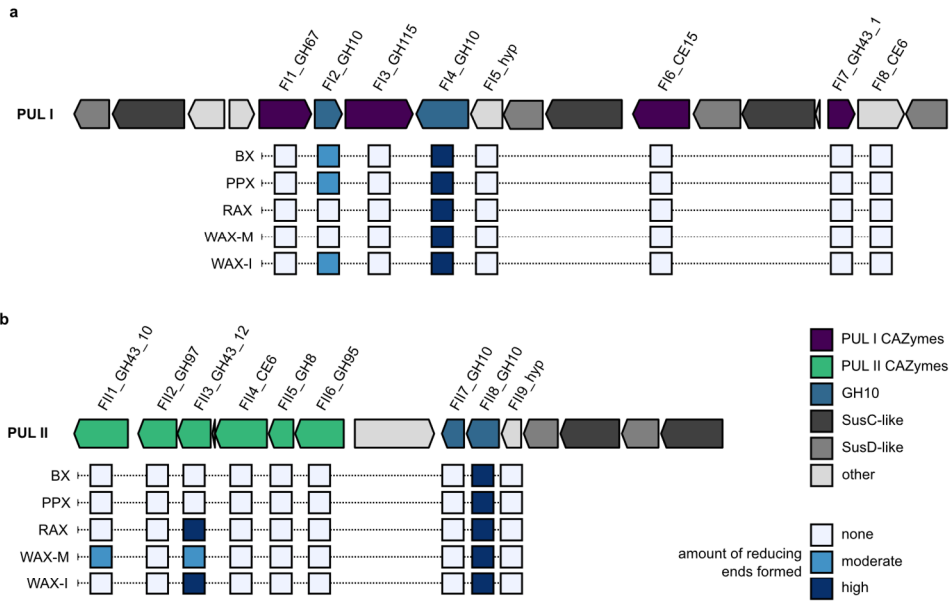
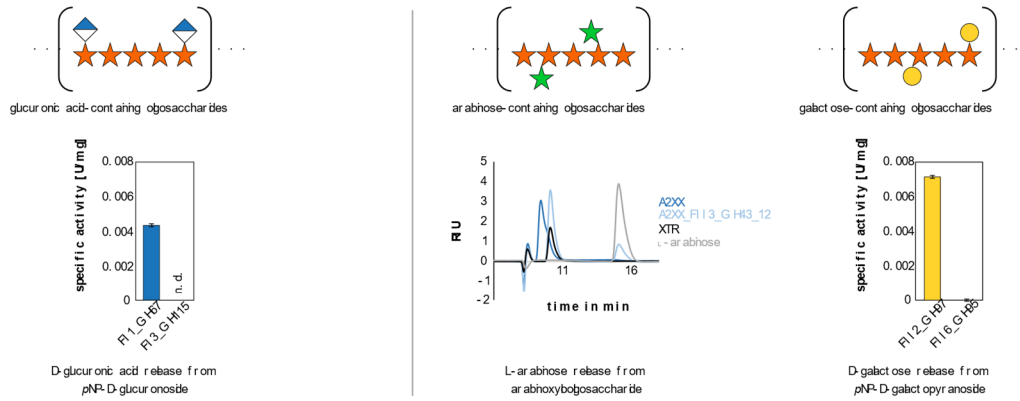


Figure 3

a Initial degradation by promiscuous activity of FUL-encoded GH10s towards β -1,4-linked xylin backbone & esterases



b FUL-specific degradation of xylin fragments



c Degradation of undecorated Xylan

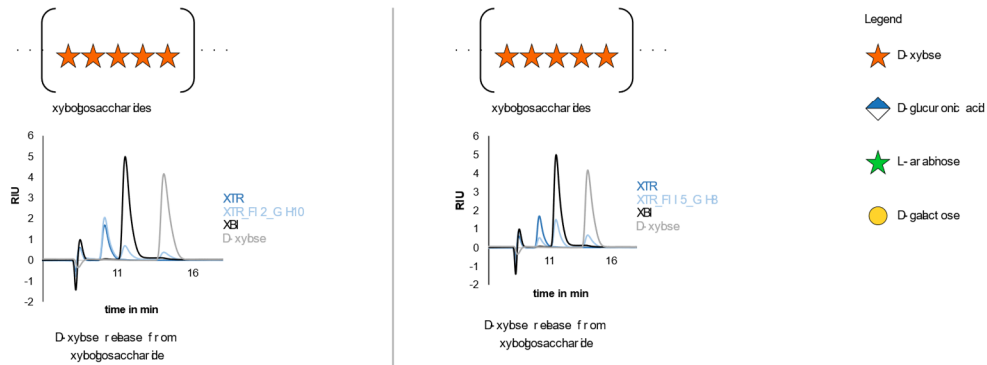
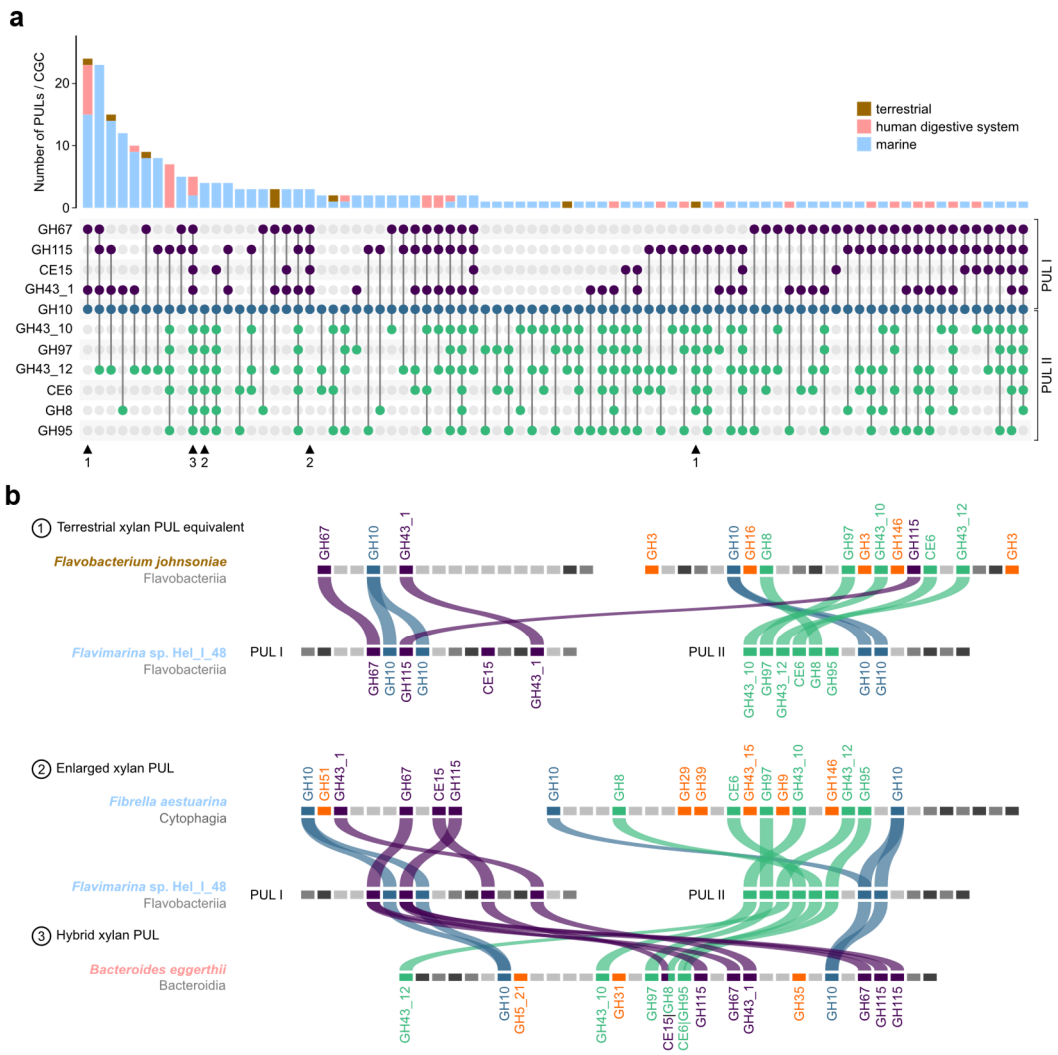


Figure 4



Supplementary online material

Utilization of a diverse range of xylan structures by marine *Bacteroidetes*

Theresa Dutschei¹, Irena Beidler², Daniel Bartosik², Julia-Maria Seeßelberg³, Michelle Teune¹, Marcus Bäumgen¹, Soraia Querido Ferreira¹, Julia Heldmann¹, Felix Nagel⁴, Joris Krull^{5,6}, Leona Berndt⁷, Karen Methling⁸, Martin Hein⁹, Dörte Becher¹⁰, Peter Langer⁹, Mihaela Delcea⁴, Michael Lalk⁷, Michael Lammers⁷, Matthias Höhne³, Jan-Hendrik Hehemann^{5,6}, Thomas Schweder^{2,5*}, Uwe T. Bornscheuer^{1*}

¹Department of Biotechnology & Enzyme Catalysis, Institute of Biochemistry, University Greifswald, 17487 Greifswald, Germany

²Department of Pharmaceutical Biotechnology, Institute of Pharmacy, University of Greifswald, 17487 Greifswald, Germany

³Department of Protein Biochemistry, Institute of Biochemistry, University of Greifswald, 17487 Greifswald, Germany

⁴Department of Biophysical Chemistry, Institute of Biochemistry, University of Greifswald, 17487 Greifswald, Germany

⁵Institute of Marine Biotechnology e.V., 17489 Greifswald, Germany

⁶University of Bremen, Center for Marine Environmental Sciences, 28359 Bremen, Germany

⁷Department of Synthetic and Structural Biochemistry, Institute of Biochemistry, University of Greifswald, 17487 Greifswald, Germany

⁸Department of Cellular Biochemistry and Metabolomics, Institute of Biochemistry, University of Greifswald, 17487 Greifswald, Germany

⁹Department of Organic Chemistry, Institute of Chemistry, University of Rostock, 18059 Rostock, Germany

¹⁰Department of Microbial Proteomics, Institute of Microbiology, University of Greifswald, 17487 Greifswald, Germany

*Corresponding authors:

Thomas Schweder, schweder@uni-greifswald.de;

Uwe Bornscheuer, uwe.bornscheuer@uni-greifswald.de

Additional methods

Gene constructs for SusD-like proteins

Enzyme purification: CAZymes and SusD-like proteins

Additional carbohydrate analyses

Monosaccharide composition analysis

Dynamic light scattering

Additional enzymatic assays

*p*NP-assay for chromogenic substrates screening

Affinity gel electrophoresis

Activity assays for carbohydrate esterases

- Acetate release
- Ferulic acid release
- D-glucuronic acid release
- Preparation of 6-*O*-acetylated derivatives of D-glucose and D-galactose (6-*O*-acetyl- α,β -D-glucopyranose and 6-*O*-acetyl- α,β -D-galactopyranose):
- ¹H-NMR spectroscopic based acetate quantification

Additional tables

Table S1: Protein summary

Table S2: Primers used in this study

Table S3: Elementary analysis of the self-extracted polysaccharides

Table S4: Dynamic light scattering results: size and zeta potential measurement

Table S5: Proteomics results (available as separate Excel file)

Additional figures

Figure S1: Growth pattern of *Flavimarina* sp. Hel_I_48 on different substrates

Figure S2: Monosaccharide composition analysis.

Figure S3: Tris-acetate based affinity gel electrophoresis of SusD-like proteins from PUL I and II on different xylan polysaccharides and laminarin from *Eisenia bicyclis*.

Figure S4: ANTS-FACE from the initial degradation of polymeric xylans.

Figure S5: DNS-reducing end assay screening of the *Flavimarina* PUL proteins.

Figure S6: ANTS-FACE from the biocatalysis of different xylan substrates pre-digested with the PUL I encoded FI4_GH10 with further PUL I CAZymes.

Figure S7: ANTS-FACE of the xylan pre-digested with the PUL II enzyme FII8_GH10.

Figure S8: HPLC result for the biocatalysis reaction with standard oligosaccharides.

Figure S9: Screening enzymatic activity with chromogenic substrates (*p*NP).

Figure S10: Analysis of the carbohydrate esterases.

Figure S11: Variety of the enzyme modularity of the two distinct flavobacterial xylan PULs.

Additional methods

Gene constructs for SusD-like proteins

Codon-optimized gene fragments encoding for SusD_I_1, SusD_I_2, SusD_I_3, SusD_I_4, SusD_II_1, SusD_II_2 and GM_SusD were ordered as geneparts (GenScript Biotech (Netherlands) B.V., Leiden, Netherlands) containing a *N*-terminal hexahistidine-tag, a tobacco etch virus cleavage site, and *N*- and *C*-terminal overhangs for a modified pET28-based vector (Table S1). The amino acids leading up to the S_{PII} cleavage site were excluded from the constructs as they just serve as signal peptide and are likely to be removed upon integration of SusD in the outer membrane [1]. Cloning was performed using the SLiCE method [2] with 125 ng of insert and a vector to insert ratio of 1:7. For preparation of the modified pET28 vector, the isolated plasmid was linearized with the restriction endonucleases FastDigest™ *NcoI* and *XbaI* (Thermo Scientific, Thermo Fisher Scientific Inc., Waltham, USA) followed by heat inactivation of the restriction enzymes at 65°C for 15 min. Due to observation of no expression of SusD_I_1, the tobacco etch virus cleavage site was substituted with the small ubiquitin-like modifier (SUMO) fusion protein SMT3 from *S. cerevisiae* [3]. Therefore, a hexahistidine-tag containing the SMT3-fusion protein coding gene fragment was amplified from a pBAD vector with primers listed in table S2. The PCR was conducted using Pfu Plus! DNA polymerase and its related buffer (EURx Ltd., Gdansk, Poland), 0.2 mM dNTPs, 0.5 μM primers and 10 ng SMT3 containing plasmid DNA under the following conditions: 5 min of initial denaturation at 95 °C, followed by 30 cycles of 30 s of denaturation at 95 °C, 30 s of annealing at 56.5 °C and 24 s of extension at 72 °C. Final extension was conducted at 72 °C for 7 min. Subsequently, residual plasmid was digested with *DpnI* in CutSmart™ Buffer (NEB, New England Biolabs Inc., Ipswich, USA) at 37 °C for one hour. 2 μg of the amplified gene fragment and SusD_I_1 containing pET28 plasmid, respectively, were incubated with 10 U *SaI* and 20 U *XbaI* (NEB, New England Biolabs Inc., Ipswich, USA) at 37 °C for 2 hours and enzymes were heat inactivated at 80 °C for 2 h. The restriction digested plasmid was mixed with Gel Loading Dye, Purple (6X) (NEB®, New England Biolabs Inc., Ipswich, USA) and separated from potential circular plasmid DNA by gel electrophoresis in a 0.8% (w/v) agarose gel containing 0.005% (v/v) ROTISafe (Carl Roth, Karlsruhe, Germany). The agarose gel electrophoresis was conducted at 90 V for one hour and the linearized plasmid was processed with the Monarch DNA gel extraction kit (NEB, New England Biolabs Inc., Ipswich, USA). The insert was purified with a NucleoSpin gel and PCR Clean-up kit (MACHEREY-NAGEL GmbH & Co. KG, Düren, Germany). Hexahistidine-tagged SMT-3 was introduced into the SusD_I_1 containing pET28a plasmid using 50 ng of vector and 200 U T4 DNA Ligase (NEB, New England Biolabs Inc., Ipswich, USA) at 24 °C for two hours and a 1:5 ratio of vector (0.02 pmol) to insert (0.1 pmol).

Enzyme purification

CAZymes

Cell pellets for the purification of CAZymes from a 50 mL culture were thawed on ice and resuspended in 10 mL of ice-cold resuspension buffer (Tris-HCl 50 mM, pH 8 + 300 mM NaCl + 10 mM imidazole). The cells were lysed by ultrasonication on ice (2 x 3 min, 50% power, 50% cycle time) and the cell debris was removed by centrifugation (15 min at 10,000 x *g*). Rotigarose-His/Ni beads (Carl Roth, Karlsruhe, Germany) incubated with the clarified lysate were used in gravity flow columns. After washing, the protein was eluted with elution buffer (Tris-HCl-50 mM, pH 8 + 100 mM NaCl + 300 mM imidazole). Fractions containing the protein of interest were pooled and desalted using PD-10 columns (GE Healthcare, Freiburg, Germany) equilibrated with 50 mM Tris-HCl (pH 8 + 10 mM NaCl). The desalted enzymes were aliquoted in tubes, flash frozen in liquid nitrogen and stored at -20 °C. The protein concentration was determined with the Roti[®]-Nanoquant kit with an albumin standard (0-100 µg/mL).

The F18A_CEnc and F11A_CEnc was additionally purified via size exclusion chromatography for the NMR-analysis. The gravity flow purified enzyme was loaded to a Superdex 200 16/600 (Cytiva, Marlborough, MS, USA) (50 mM TRIS-HCl pH 7.5, 100 mM NaCl, 2 mM β-Mercaptoethanol) and the resulting protein fraction was then dialysed overnight in 20 mM K₂HPO₄/KH₂PO₄ buffer pH 7.7 with 100 mM NaCl.

SusD-like proteins

The SusD-expression cell pellets of 600 mL culture were thawed on ice and resuspended in 20 mL of cooled lysis buffer (20 mM sodium phosphate, 500 mM sodium chloride, 20 mM imidazole, 5% (v/v) glycerol, pH 8). Cell lysis was performed at 1,000 atm with a Maximator HPL6 device (MAXIMATOR GmbH, Nordhausen, Germany) and the cell suspension was lysed twice. Cell debris were removed by centrifugation at 10,000 x *g* and 4 °C for 1 h. The supernatant was filtered using a 0.45 µm filter. Protein purification was performed at room temperature using a Cytiva HisTrapHP 1 mL column (Cytiva Europe GmbH, Freiburg, Germany) at an Aekta Pure device (Cytiva Europe GmbH, Freiburg, Germany) with a flow rate of 1 mL min⁻¹ and an imidazole gradient ranging from 20 mM to 250 mM imidazole within 7 min. Purification buffers contained 20 mM sodium phosphate, 500 mM sodium chloride and no or 500 mM imidazole respectively at pH 8. SusD-like proteins eluted between 65-86 mM imidazole. SDS-PAGE using 10% (v/v) polyacrylamide gels was performed to identify fractions containing the protein of interest. Proteins were stained with ROTIBLue (Carl Roth, Karlsruhe, Germany). Those fractions containing the desired protein were united and the volume was adjusted to 1 mL with Vivaspin 10,000 MWCO ultrafiltration units (Sartorius Stedim Lab Ltd., Stonehouse, UK) at 4,000 x *g* and 4 °C for the required time. Subsequently, the samples were centrifuged at 17,000 x *g* at 4 °C for 10 minutes in a table top centrifuge to obtain aggregate

free samples for injection. Size exclusion chromatography was performed at room temperature with a flow rate of 0.75 mL/min using a Cytiva Superdex 200 Increase 10/300 GL column (Cytiva Europe GmbH, Freiburg, Germany) in 20 mM sodium phosphate, 250 mM sodium chloride at pH 8. The proteins were concentrated again with Vivaspin 10,000 MWCO ultrafiltration units (Sartorius Stedim Lab Ltd., Stonehouse, UK). Protein concentrations were measured under consideration of their particular molecular weight and extinction calculated with the Expasy ProtParam tool [4] using a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific Inc., Waltham, USA). Aliquots of 1 mg mL⁻¹ were flash frozen in liquid nitrogen and stored at -20 °C.

Additional carbohydrate analyses

Monosaccharide composition analysis

The self-extracted polysaccharides were chemically hydrolysed (1 M HCl for 24 h at 100 °C). Afterwards, the samples were filtered using a 0.2 µm Spin-X filter prior to HPAEC-PAD analyses using a Dionex CarboPac PA10 column (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and monosaccharide mixtures as standards for column calibration [5].

Dynamic light scattering

Hydrodynamic diameters of higher order polysaccharide structures and zeta potentials were determined using dynamic light scattering (DLS). Backscattering was recorded at 173°nm and samples were equilibrated at 25°C. Zeta potential measurements were carried out using a maximum voltage of 10 V. Data represent mean and standard deviation of at least three independent experiments (Table S4).

Additional enzymatic assays

pNP-assay chromogenic substrates screening

1 mM solutions of 4-nitrophenyl-/D-galacturonide/L-arabinose/acetate were prepared in 50 mM HEPES buffer, pH 7.4 with 100 mM NaCl. 10 µL of purified enzyme solution were added to 200 µL of the substrate solutions and the absorbance at 410 nm was measured over 30 min to detect the formation of *p*-nitrophenolate. The latter was taken from Bowers et al. [6](18.3 mM). The autohydrolysis was determined by no addition of enzyme.

Affinity gel electrophoresis

Affinity gel electrophoresis (AGE) was performed using Tris-acetate based gels with 10% (v/v) acrylamide [7] casted as native gels without SDS but containing 0.5% (w/v) of polysaccharide (PPX, BX, RAX, WAX-M and Laminarin from *Eisenia bicyclis* from TCI Europe N.V., Zwijndrecht, Belgium) or no polysaccharide as reference, respectively. 5 µg protein in native charge buffer (63 mM Tris-HCl pH 6.8, 10% (v/v) glycerol, 0.01% (w/v) bromophenole blue) were applied and AGE was conducted on ice at 80 V for 6 hours in cooled buffer

containing 25 mM Tris, 192 mM glycine at pH 8.3. Proteins were stained with ROTIBLue (Carl Roth, Karlsruhe, Germany).

Activity assays for carbohydrate esterases

Acetate release

The release of acetate from polymeric xylan and from the partially acetylated beechwood xylan from Megazymes (Bray, Ireland) was measured after the biocatalytic reactions with the esterases performed at room temperature for 16 h. The acetic acid detection was performed with the acetic acid kit from R-Biopharm (Darmstadt, Germany). The release of acetate from the 6-*O*-acetylated D-glucose and the 6-*O*-acetylated D-galactose were measured with the acetate detection kit from Megazymes (Bray, Ireland). 2 μ M of the enzyme were incubated with 10 mM of the substrate and incubated at 25 °C overnight. The enzyme was removed via Ni-NTA-beads (Cube Biotech, Monheim, Germany) before the sample was analyzed.

Ferulic acid release

The release of ferulic acid from WAX-I was measured by incubating the FII4_CE6 with 10 mg/mL of the substrate, in a 50 mM sodium phosphate 100 mM buffer pH 6,5 taking time samples (2 h, 4 h, 8 h, 1 d, 2 d) incubating at 37 °C and 1,000 rpm. After incubation the enzymes were heat inactivated at 90 °C for 10 min. The same volume of methanol was added to the cooled down samples, vortexed for 30 s and centrifuged at 13,000 \times *g* for 5 min. The resulting supernatant was analyzed *via* ultra-high-performance liquid chromatography with pulse amperometric detector (U-HPLC, Agilent) injected with 5 μ L, flow rate 0.8 mL/min with the liquid phase water / acetonitrile / formic acid. UV detection carried out at 325 nm. A standard curve using ferulic acid was used for determining its concentration in the hydrolysate.

D-glucuronic acid release

The measurement of CE-activity towards D-glucuronic acid-derivatives (methyl-, benzyl-, allyl-D-glucuronic acid) was performed in a discontinuous attempt using the K-URONIC Assay Kit (Megazyme) as described previously [8]. The enzyme reactions were performed for 30 min at 25 °C in 25 mM sodium phosphate buffer, measuring 3-4 different dilutions of each enzyme. To quantify the release of GlcA, the reaction solutions were transferred to a 96-well plate (200 μ L each) and detection-solution (50 μ L, 40 % NAD⁺, 6% UDH) was added to measure the release of NADH at 340 nm for 30-60 min.

Preparation of 6-O-acetylated derivatives of D-glucose and D-galactose (6-O-acetyl- α,β -D-glucopyranose and 6-O-acetyl- α,β -D-galactopyranose)

The preparation of the 6-*O*-acetylated sugars was carried out according to a modified procedure described by Duff *et al.* [9]. 10 g of the sugar (D-glucose or D-galactose) were mixed with 67% aqueous acetic acid (20 ml) and stirred at 100 °C for 16-18 hours. Then silica gel (30 g) was added and the solvent was removed at 40 °C in the RotaVap. The dried silica gel support was directly used for column chromatographic purification (eluent: toluene/methanol

v:v = 2.5:1) yielding the corresponding 6-*O*-acetylated sugar as a mixture which mainly consisted of the anomeric pyranoses (yield: 27% for 6-*O*-acetyl-D-glucose and 24% for 6-*O*-acetyl-D-galactose). The products were identified by comparison of their NMR spectroscopic data with those from the literature. The glucose derivative can be further purified by crystallization from methanol-ethyl acetate. All used chemicals are commercially available and were used without further purification. Analytical TLC on Merck silica gel 60 F₂₅₄ plates was visualized by using anisaldehyde-sulfuric acid colouring reagent in methanol. Column chromatography was performed on Merck Geduran Si 60 (0.063-0.200 mm). ¹H NMR and ¹³C NMR spectra were recorded on a Bruker AVANCE DRX-500 or AVANCE 300 III. Chemical shifts in ppm were calibrated by residual solvent signals methanol-d₄ (¹H, 3.31 ppm, ¹³C, 49.00 ppm), D₂O (¹H, 4.79 ppm) or DMSO-d₆ (¹H, 2.50 ppm, ¹³C, 39.52 ppm).

¹H-NMR spectroscopic based acetate quantification

¹H-NMR analysis was performed as previously described with some modifications [10]. In brief, 400 µl of samples and 200 µl of 0.2 mol/L sodium hydrogen phosphate buffer solution, containing 30 % D₂O (Euriso-Top, St-Aubin Cedex, France) and 1.5 mmol/L 3-trimethylsilyl-[2,2,3,3-D₄]-1-propionic acid (Sigma-Aldrich, St. Louis, USA) were mixed in 5 mm glass tubes (103.5 mm length, Bruker Biospin GmbH, Rheinstetten, Germany). The Bruker AVANCE-NEO 600 NMR spectrometer equipped with a SampleJet autosampler and a 5mm QCI cryo probe was operated by TOPSPIN 4.0.9 software (Bruker Biospin GmbH, Rheinstetten, Germany). Metabolite quantification was done using AMIX Viewer 3.9.15 software (Bruker Biospin GmbH, Rheinstetten, Germany). Integrals of the acetate peak were compared to the integral of the ERETIC signal for absolute quantification. The ERETIC signal was generated by external calibration with the ERETIC quantification tool based on PULCON [11].

Table S1: Summary of the investigated carbohydrate active enzymes. The annotation was performed using DBCan, Pfam, Interpro, and Hmmer. Only annotations provided by at least two tools were used. The genes were either ordered as synthetic genes at Biocat (Heidelberg, Germany) (1) or GenScript (NJ, USA) (2) or cloned via FastCloning technique (3) *: FI8A and FI8B differentiate between two optional lengths, FII1A and FII1B are separate modules of the gene P162_RS04015 and were ordered as separate codon optimized genes.

Name	Gene Locus	Taq	Annotation and modularity	Functional annotation	Gene origin
FI1_GH67	P162_RS02330		GH67	n.d.	(3)
FI2_GH10	P162_RS02335		GH10	<i>Exo</i> -1,4-xylanase	(3)
FI3_GH115	P162_RS02340		GH115 GH115	Alpha-(4- <i>O</i> -methyl)-glucuronidase (EC3.2.1.)	(2)
FI4_GH10	P162_RS02345		CBM4 GH10	<i>Endo</i> -xylanase	(3)
FI5_hyp	P162_RS02350		hyp	n.d.	(3)
FI6_CE15	P162_RS02365		CE15 CBM9	4- <i>O</i> -Methyl-glucuronyl methylesterase	(2)
FI7_GH43_1	P162_RS02380		GH43_1	n.d.	(3)
FI8A_CEnc*	P162_RS02385		Putative CE6	Acetyl-xylan esterase	(3)
FI8B_CEnc*	P162_RS02385		Putative CE6	Acetyl-xylan esterase	(3)
FII1A_CEnc	P162_RS04015		CE3 GH43_10	Acetyl-xylan esterase/ (xylosidase/arabinose)	(1)
FII1B_GH43_10	P162_RS04015		CE3 GH43_10	n. d.	(1)
FII2_GH97	P162_RS04020		GH97	α -D-galactosidase	(1)
FII3_GH43_12	P162_RS04025		GH43	α -L-arabinofuranosidase	(1)
FII4_CE6	P162_RS04030		CE6 CEnc CEnc (CEnc putative CE1)	Feruloyl xylan esterase /acetyl xylan esterase	(1)
FII5_GH8	P162_RS04035		GH8	<i>Exo</i> -xylanase	(1)
FII6_GH95	P162_RS04040		GH95	n.d.	(1)
FII7_GH10	P162_RS04050		GH10	n.d.	(1)
FII8_GH10	P162_RS04055		CBM4 GH10	<i>Endo</i> -1,4-xylanase	(1)
FII9_hyp	P162_RS04060		Hyp (DUF1735)	n.d.	(1)
GM SusD	BLT93_RS06685		RagB/SusD family nutrient uptake outer membrane protein	Target substrate laminarin	(2)
SusD_I_1	P162_RS02310		RagB/SusD family nutrient uptake outer membrane protein	Target substrate xylan	(2)

SusD_I_2	P162_RS02355	RagB/SusD family nutrient uptake outer membrane protein	Target substrate xylan	(2)
SusD_I_3	P162_RS02370	RagB/SusD family nutrient uptake outer membrane protein	Target substrate xylan	(2)
SusD_II_1	P162_RS04065	RagB/SusD family nutrient uptake outer membrane protein	Target substrate xylan	(2)
SusD_II_2	P162_RS04075	RagB/SusD family nutrient uptake outer membrane protein	Target substrate xylan	(2)

Table S2: Primers used in this study. The primers were purchased from Invitrogen Life Technologies (CA, USA).

Primer name	Sequence 5' → 3'	Purpose
T7 pET mode	CCCGCGAAATTAATACGACTCAC	Sequencing
T7_term	CTAGTTATTGCTCAGCGGT	Sequencing
F11 seq fw	GCGCTGGCTGATGTGTTTCGACC	Sequencing
F11 seq fw2	CCAAAAAGGAAAAGGATCCACC	Sequencing
F14 seq fw	GTAAAGGCCGAATTTCTTTTG	Sequencing
F18A seq fw	GATACTGATCAGGGAATCATTAACT	Sequencing
F19 seq fw	GATCATGGTGGTGGTGGTGTATAT	Sequencing
F19 seq fw2	CGAGATGATTGATGGTACATTA	Sequencing
F19 seq fw3	CTAGCGCTAACGGTAATTTTGCTG	Sequencing
F11A fw	CTGCTTCACATAACAGCTTCTGATGGC	Fast Cloning
F11A rv	TTAATCCCATTGTGGCCGTATTCCC	Fast Cloning
F11B fw	TTGAAACCGCATCCTACATTCTTTTATACTC	Fast Cloning
F11B rv	ATGCTTGGGCAAGAAATAGAACTGCTACC	Fast Cloning
F12 fw	TGCAAAAACGAGACAAAACCAC	Fast Cloning
F12 rv	TTATTCGTTGATGTCGGTTACTTTATAG	Fast Cloning
F13 fw	CAGAAATCTGGTGATTATGTATCAAAAACAC	Fast Cloning
F13 rv	CTATTCACAACCTTACTTTCAGGAGGTCC	Fast Cloning
F14 fw	TGTGAAGACGATATTATGGAGTGGCAGG	Fast Cloning
F14 rv	CTAATCTCAAGCTCTCCAGTGAAATCCTC	Fast Cloning
F15 fw	TGTTCCAACGATGATGATGCTG	Fast Cloning
F15 rv	TTATTCAGGAAAATCGGTAACGGTAGG	Fast Cloning
F16 fw	CAACTCCGTTGGTCTATAACTCTGAAAATACGG	Fast Cloning
F16 rv	TTAAAGTGTCTCCTGCCAGCACCAC	Fast Cloning
F17 fw	TGCAAAAATAACACAGATAAAGATTCCG	Fast Cloning
F17 rv	TTATGGATTTTCTACCTTGGCATCAATAG	Fast Cloning

FI8A rv	CTACCAGTTATCTGTCCTAAAAGGTGAGGCAG	Fast Cloning
FI8B fw	CAGATCAAACCTGCCAAAATTAGTTTCTGACG	Fast Cloning
FI9 fw	CAGGTAGTGACCAGCGGGGCAG	Fast Cloning
FI9 rv	CTAATTCAGTTGAACCGTTCCTCCTCTTGATG	Fast Cloning
SMT3_amplification_fwd	GCTAGCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGA GATATACGATGGGTGCATCATCATCATCACGGCAGC G	Amplification of SMT3 with overhangs
SMT3_amplification_rev	GCACTACCATGGAACCACCAATCTGTTCTCTGTGAGC	Amplification of SMT3 with overhangs

Table S3: Elementary analysis of the self-extracted polysaccharides.

	N [%]	C [%]	S [%]	C/N ratio
<i>C. prolifera</i> (CPX)	0.83 ± 0.10	35.05 ± 1.19	1.18 ± 0.15	42.31 ± 5.30
<i>P. palmata</i> (PPX)	1.09 ± 0.06	35.43 ± 0.50	2.04 ± 0.13	32.43 ± 1.77
Beechwood (BX)	0.10 ± 0.02	40.76 ± 0.39	0.16 ± 0.07	410.34 ± 96.70

Table S4: Hydrodynamic diameters and ζ -potentials of polysaccharides determined by dynamic light scattering.

Sample	Hydrodynamic diameter (nm)	Zeta potential (mV)
Beechwood xylan (BX)	189 ± 29	-14.5 ± 4.3
<i>Palmaria palmata</i> xylan (PPX)	510 ± 37	-33 ± 5.75
<i>Caulerpa prolifera</i> xylan (CPX)	172 ± 4	-20.9 ± 1.9
Wheat arabinoxylan medium viscosity (WAX-M)	300 ± 21	-4.7 ± 0.3
Wheat arabinoxylan insoluble fraction (WAX-I)	699 ± 509	-4.3 ± 0.5
Rye arabinoxylan (RAX)	459 ± 8.7	-5.25 ± 0.5

Table S5: Proteomics results. *Flavimarina* sp. Hel_I_48 was grown on different xylans and pectin from apple (control) as sole carbon source. Automatically calculated iBAQ values were used to determine relative % riBAQ values for semiquantitative comparison. Only proteins identified in at least two of three replicates were classed as identified. Available as separate *exl* file

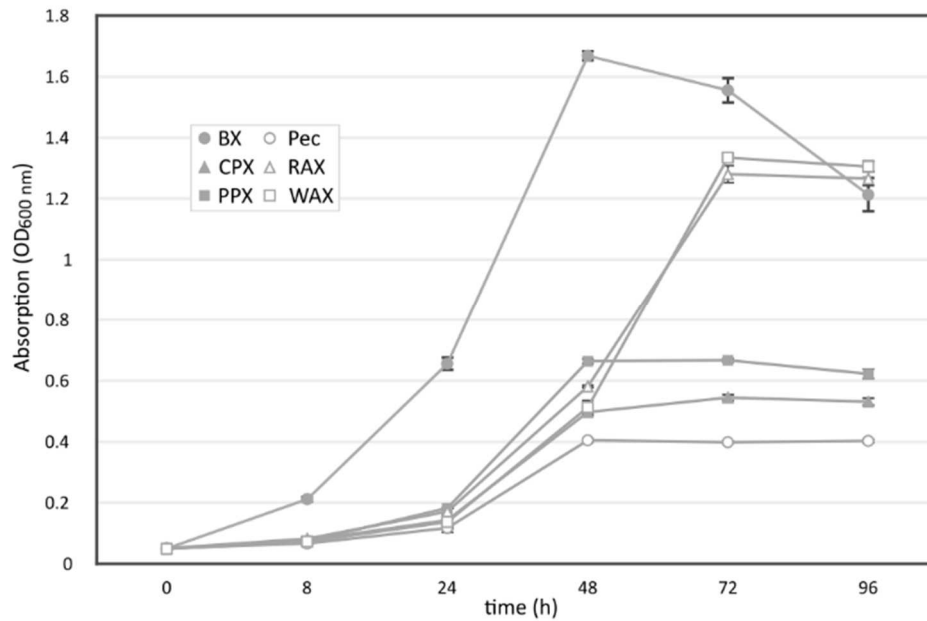


Fig. S1. Growth pattern of *Flavimarina* sp. Hel_I_48 on different substrates. Cultures were grown to stationary phase in MPM Medium with 0.1% of a specific carbon source to determine optimal sampling points for proteomics. Beechwood xylan: BX, *C. proliferata* xylan: CPX, *P. palmata* xylan: PPX, Apple pectin: Pec, rye arabinoxylan: RAX, wheat arabinoxylan: WAX

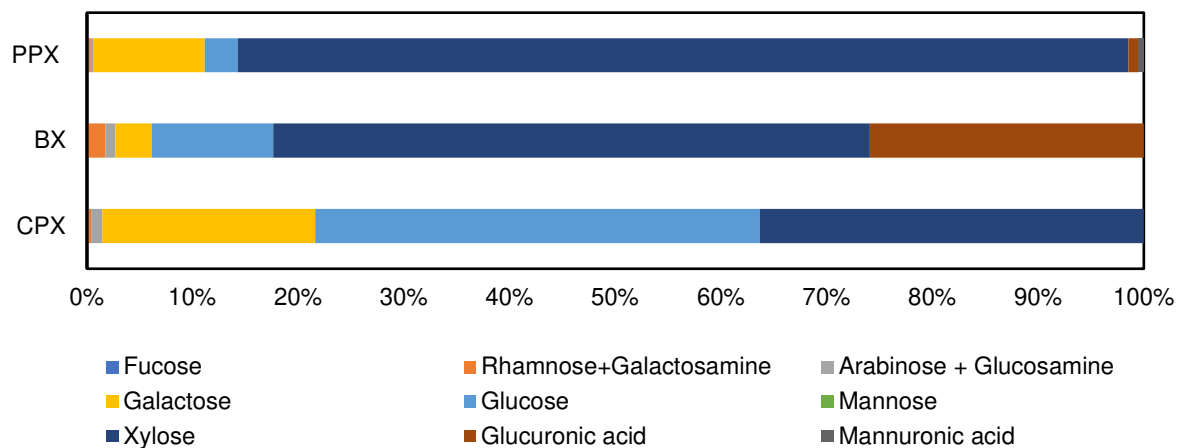


Fig. S2. Monosaccharide composition analysis of the self-extracted polysaccharides from *P. palmata* (PPX), beechwood (BX) and *C. proliferata* (CPX).

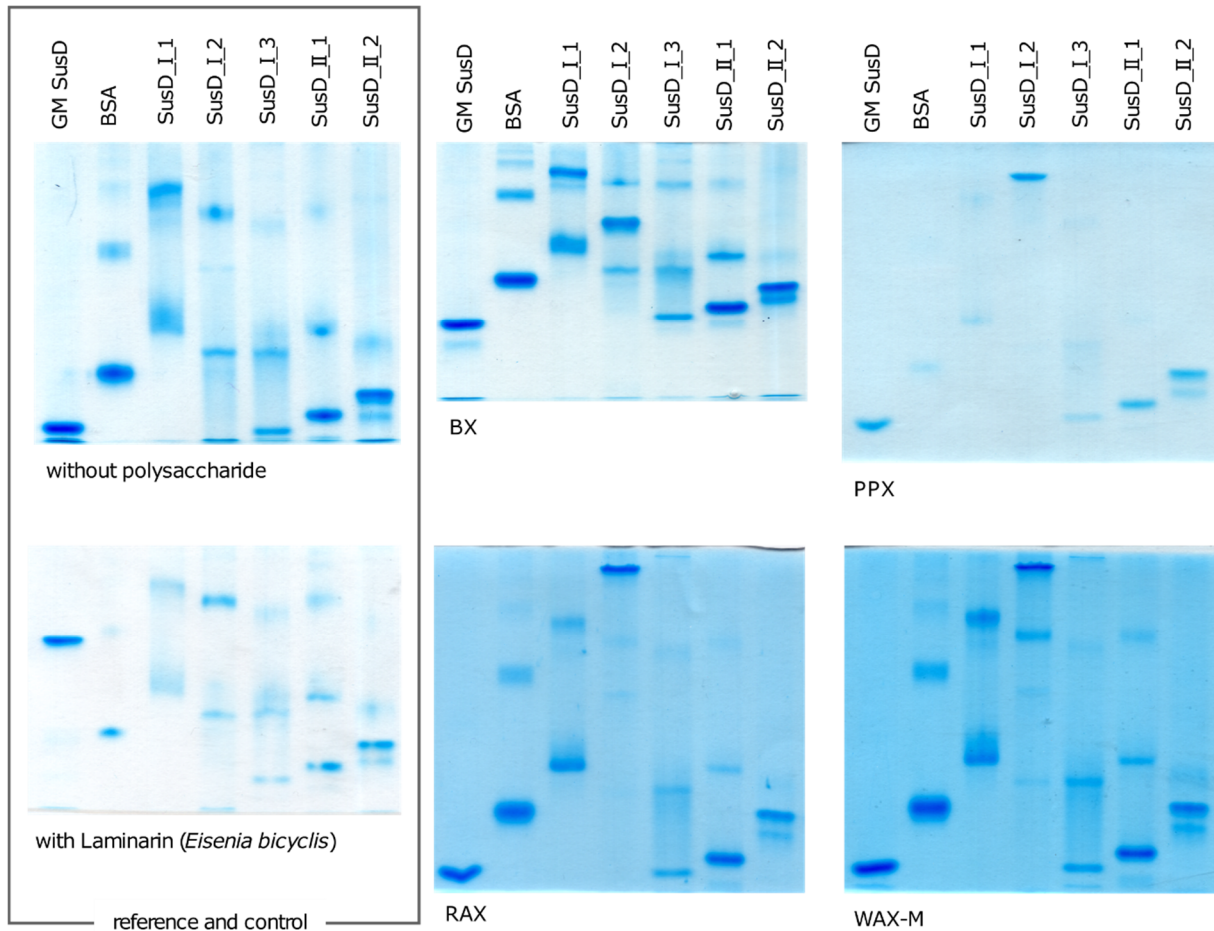


Fig. S3. Tris-acetate based affinity gel electrophoresis of SusD-like proteins from PUL I and II on different xylan polysaccharides and laminarin from *Eisenia bicyclis*. 0.5% of polysaccharides (BX, PPX, RAX, WAX-M, laminarin from *E. bicyclis*) were added to native PAGE gels before polymerization. A gel without polysaccharide served as reference, the gel with 0.5% laminarin from *E. bicyclis* was made to confirm functionality of this method. BSA acts as a reference marker and GM SusD [1] acts as a positive control showing a shift on the laminarin containing gel. 5 μ g of protein were loaded and showed a shift upon binding to the polysaccharide and a separation of their multimers.

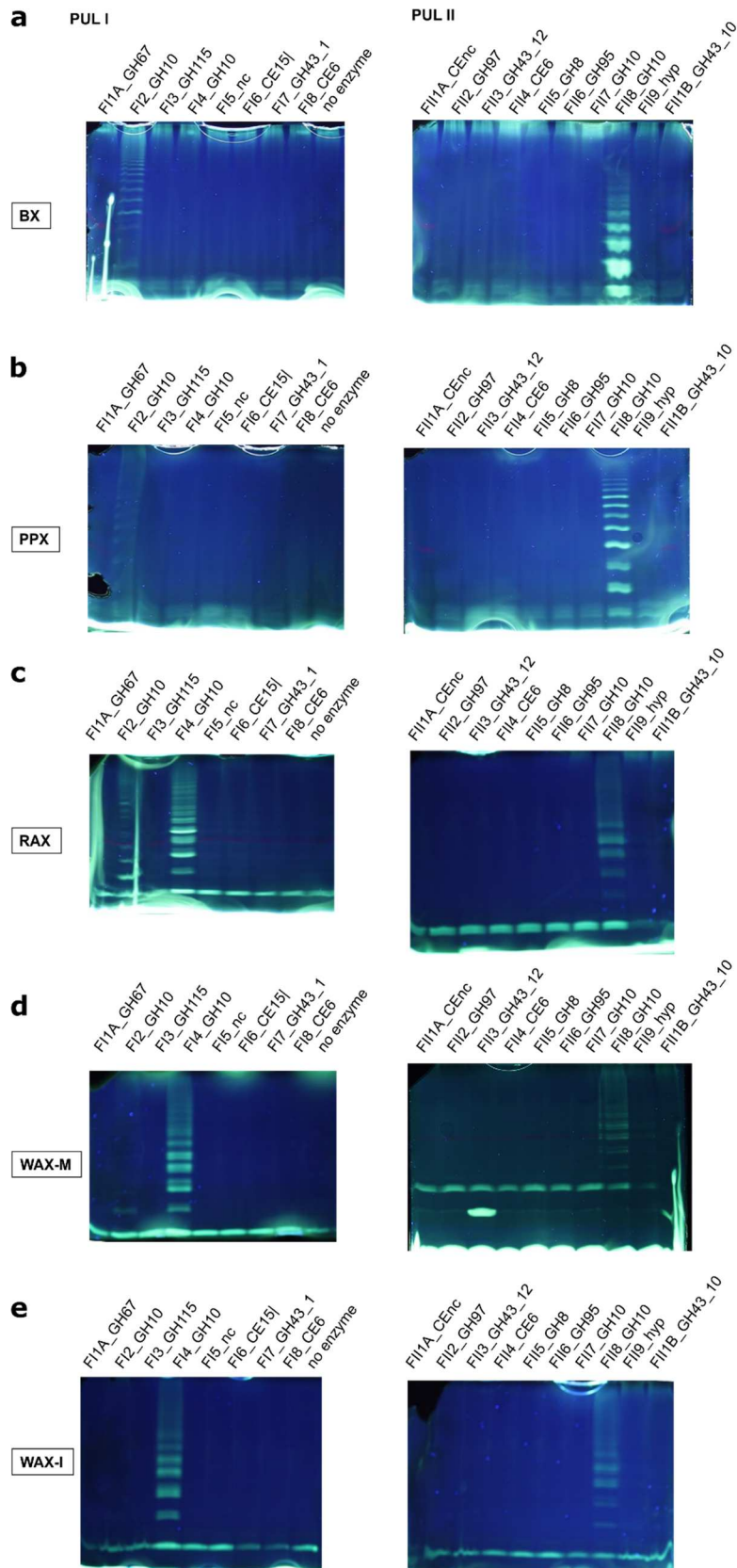


Fig. S4. ANTS-FACE from the initial degradation of different polymeric xylan substrates. Biocatalysis of the recombinantly expressed flavobacterial CAZymes (20 µg/mL) with 10 mg/mL carbohydrate beechwood (BX) (a), *P. palmata* xylan (PPX) (b), rye arabinoxylan (RAX) (c), wheat arabinoxylan medium viscosity (WAX-M) (d) and insoluble fraction (WAX-I) (e). Small degradation products of the polymeric substrate can be fluorescently labelled with the ANTS dye and separated in an electric field. A ladder-like pattern indicates xylanase activity, while single bands indicate exolytic glycoside hydrolases activity. The ANTS-FACE gel of the *C. proflera* (CPX) is not shown

as no degradation products were visible. CPX was excluded from further analysis. The biocatalysis of the F14_GH10 with BX and PPX also results in a ladder pattern which is seen in Fig 3. and a positive red end assay Fig. S5).

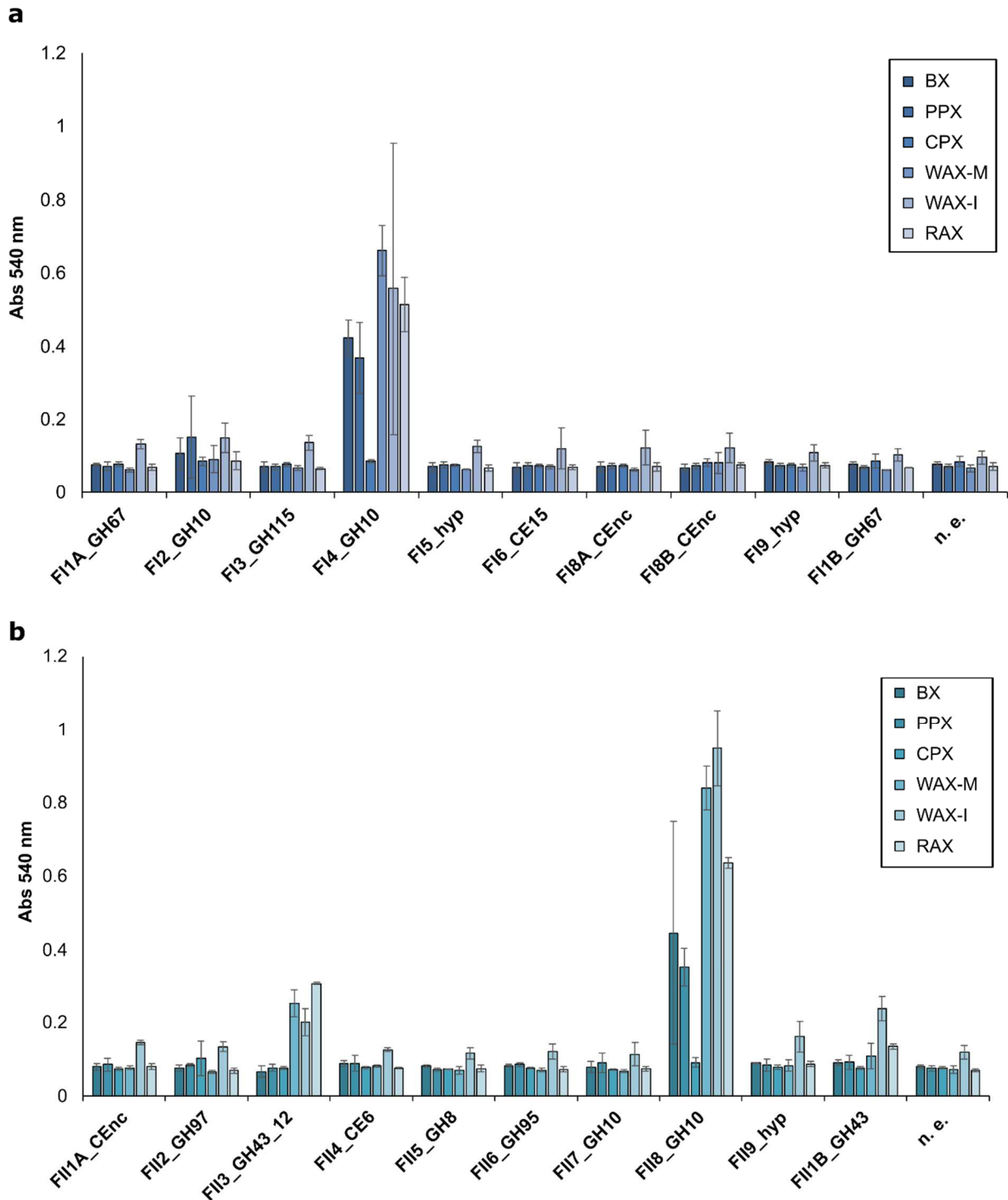


Fig. S5. DNS-reducing end assay screening of the *Flavimarina* sp. PUL proteins. (a) PUL I (b) PUL II enzymes (15 $\mu\text{g/mL}$) were incubated overnight with a 1% xylan solution. The reducing ends were measured with the DNS-assay [12]. The increase of the absorption at 540 nm indicates the increase of reducing ends and thus the carbohydrate degradation.

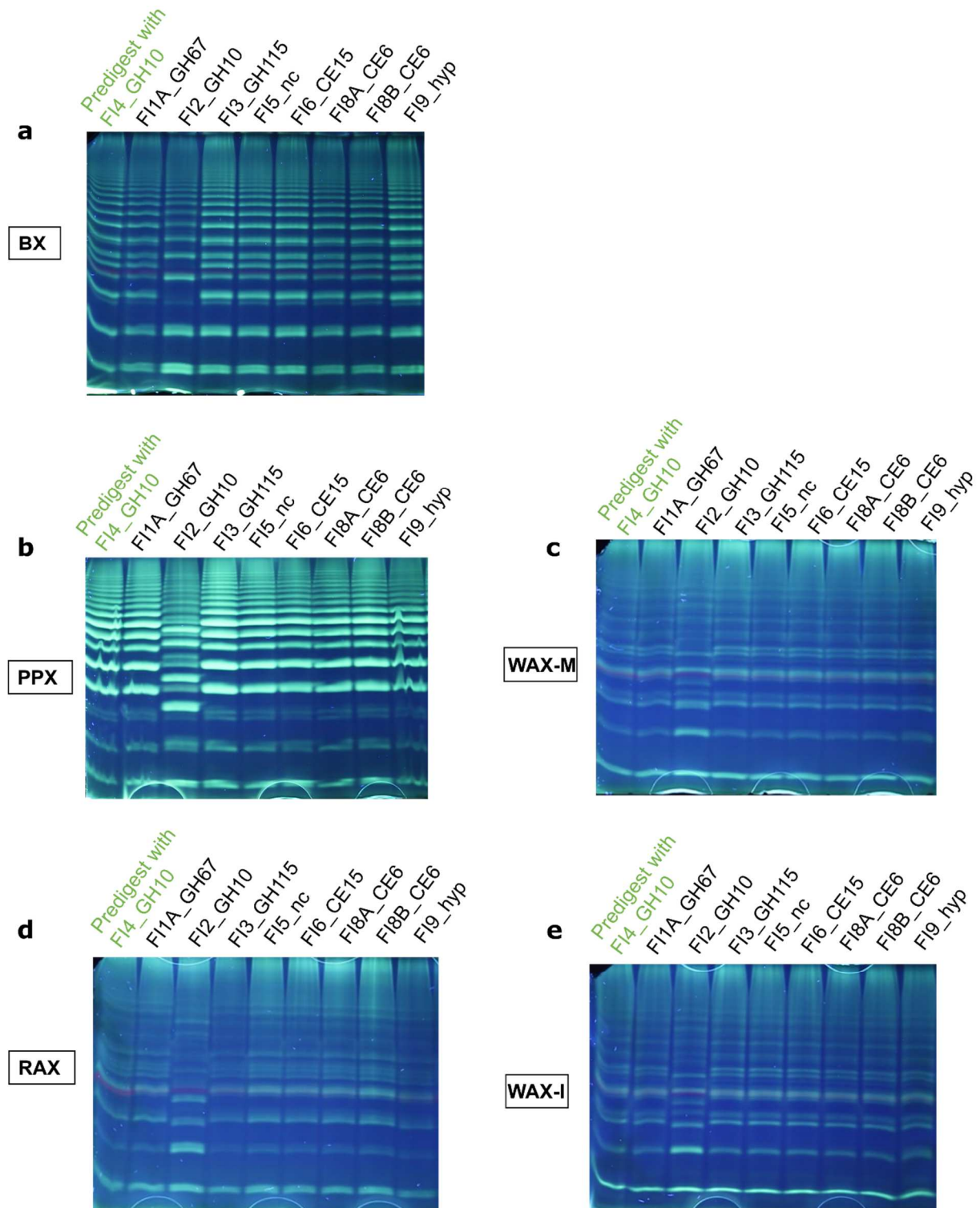


Fig. S6. ANTS-FACE analysis of samples from the biocatalytic reactions utilizing different xylan substrates pre-digested with the PUL I encoded F14_GH10 with further PUL I CAZymes. The carbohydrate beechwood (BX) (a), *P. palmata* xylan (PPX) (b), rye arabinoxylan (RAX) (c), wheat arabinoxylan medium viscosity (WAX-M) (d) and insoluble fraction (WAX-I) (e) were predigested with the F14_GH10 enzyme and combined with further enzymes of the PUL I after heat inactivation of the F14_GH10 enzyme. A shift in the gel pattern indicates further enzyme activity, with the enzyme F2_GH10 enzyme a shift is visible indicating *exo*-xylanase activity.

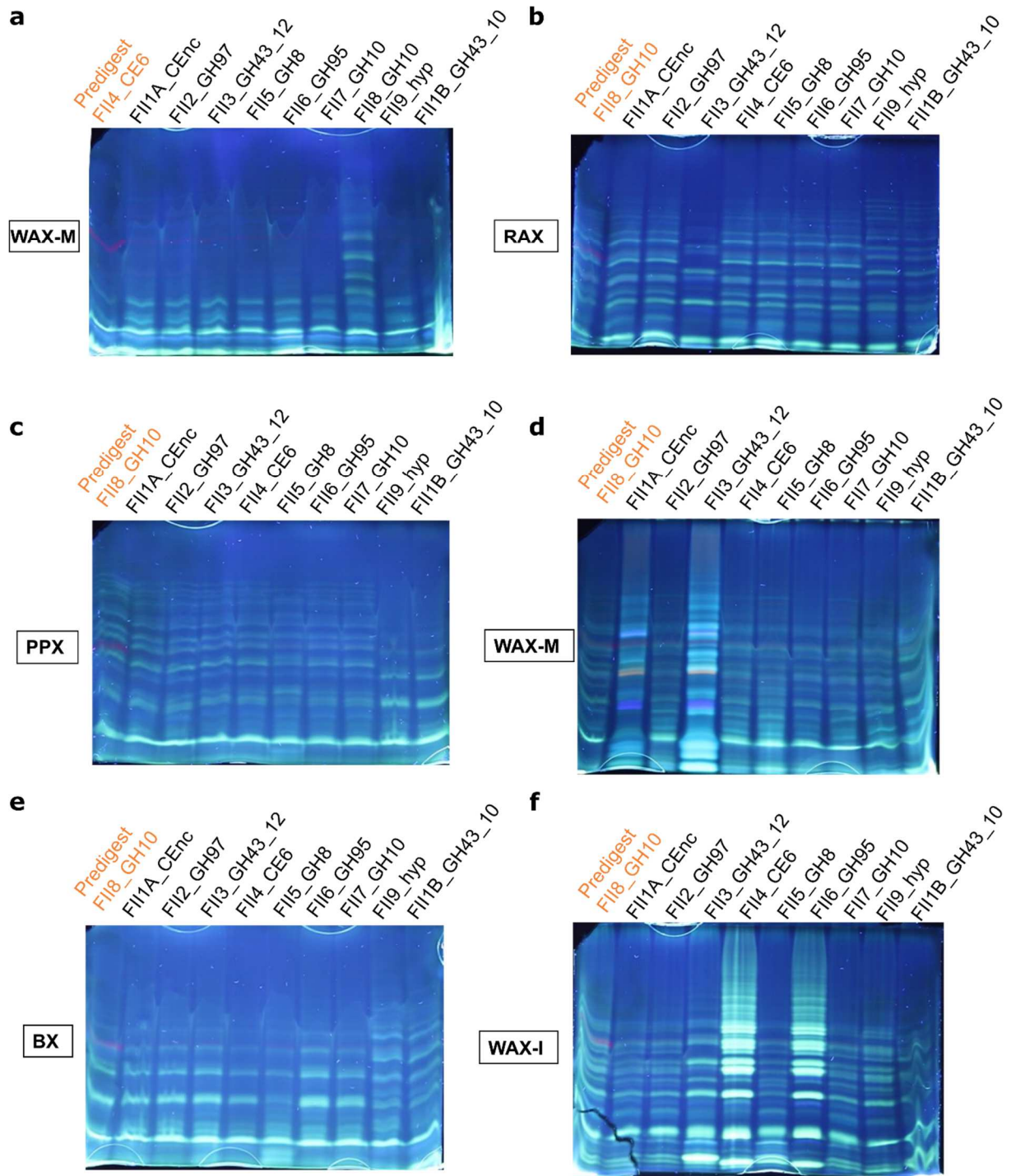


Fig. S7. ANTS-FACE analysis of the xylan pre-digested with the PUL II enzyme FII8_GH10. The carbohydrate wheat arabinoxylan medium viscosity (WAX-M) was pre-digested with FII4_CE6 and combined with further PULII enzymes after heat inactivation (a). The change of band pattern indicated the further hydrolysis of the *endo*-xylanase FII8_GH10. The carbohydrate rye arabinoxylan (RAX) (b), *P. palmata* xylan (PPX) (c) wheat arabinoxylan medium viscosity (WAX-M) (d) beechwood (BX) (e) and insoluble fraction (WAX-I) (f) were pre-digested with the FII8_GH10 enzyme and combined with further enzymes of the PUL II after heat inactivation of the FII4_GH10 enzyme. A shift in the gel pattern indicates further enzyme activity, with the enzyme FII2_GH10 enzyme a shift is visible indicating *exo*-xylanase activity

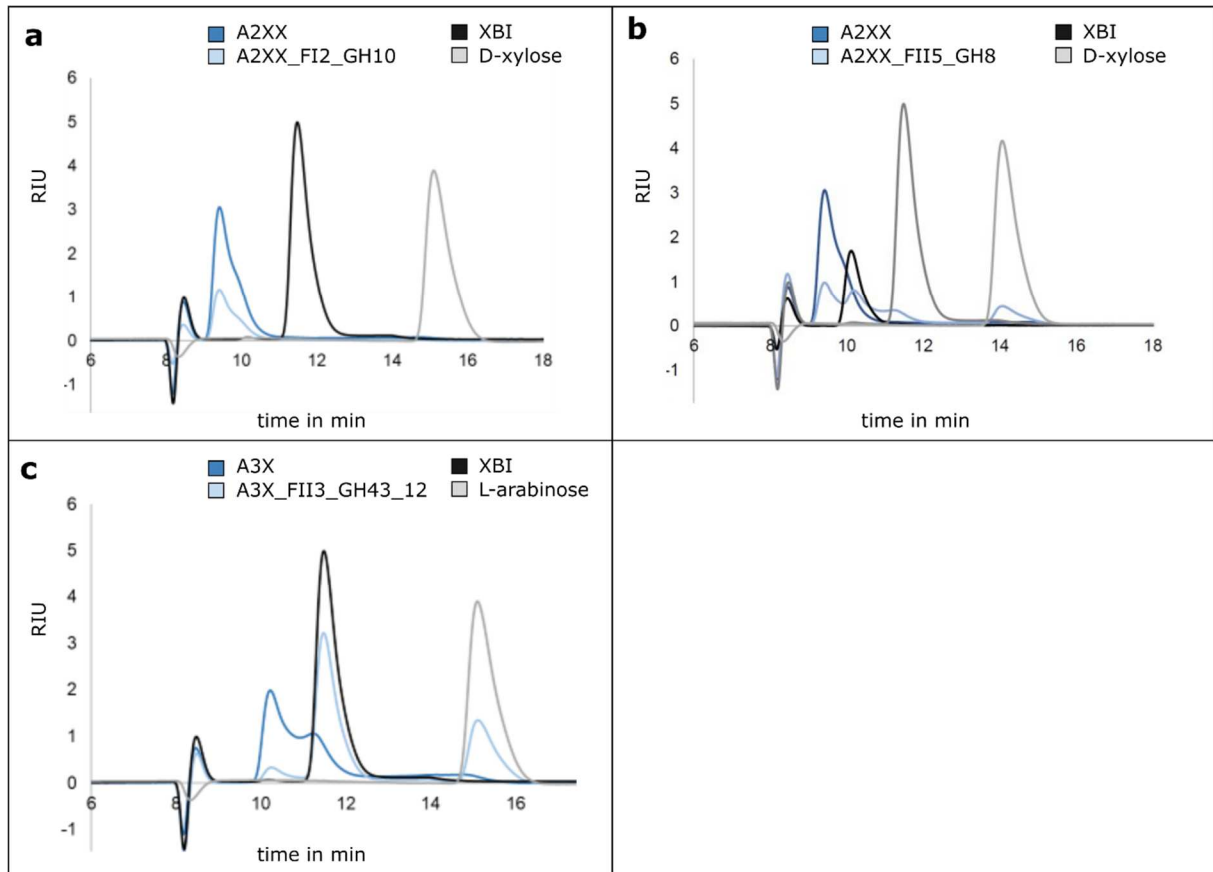


Fig S8. HPLC results of biocatalysis reactions with standard oligosaccharides. The biocatalysis reactions with the 1 mM standard substance purchased from Megazymes (Wicklow, Irland) was analysed after 20 μ M enzyme was added and incubation was performed overnight at room temperature. The sugars were measured via HPLC-RI (Hitachi Chrommaster 5310 column oven, Hitachi Chrommaster 5450 RI detector; SugarSep-H 10 μ m 300 x 8 mm) while the formed products were compared with the single standard solutions. Arabinofuranosidase activity of the FII3_GH43_12 was verified via biocatalysis with (A2XX) 2³- α -L-arabinofuranosyl-xylotriiose (a) and (A3X) 3²- α -L-arabinofuranosyl-xylobiose (c) releasing L-arabinose. The arabinofuranosidase activity and xylosidase activity of the FII5_GH8 was verified with the biocatalysis towards xylotriiose (XTR) (Fig. 3) and A2XX (b). The *exo*-xylanase activity of the FI2_GH10 was verified *via* the biocatalysis with XTR releasing xylobiose (XBI) and D-xylose and the missing release of D-xylose from A2XX (a) show the *exo*-xylanase activity of this enzyme from the non-reducing end.

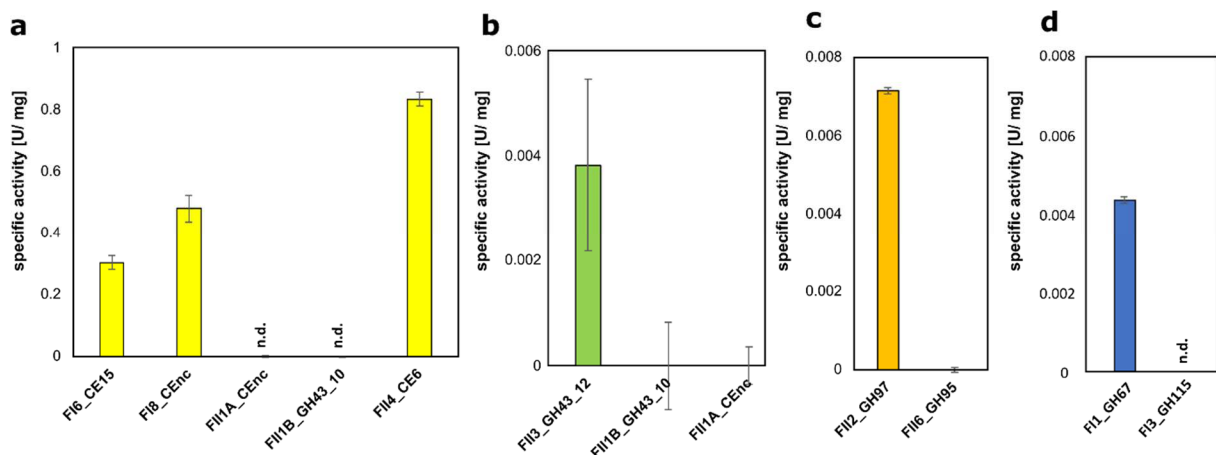


Fig. S9. Screening enzymatic activity with chromogenic substrates (pNP). The autohydrolysis was deducted from each value, an increase of absorption corresponds to the hydrolytic activity of the enzyme. (a) Activity of the carbohydrate esterases towards the artificial substrate pNP-acetate. (b) Carbohydrate esterase substrate activity of putative arabinosidase towards the pNP- α -L-arabinose substrate. (c) Activity of putative galactosidase towards pNP- α -D-galactoside and (d) activity towards pNP- α -D-glucuronide.

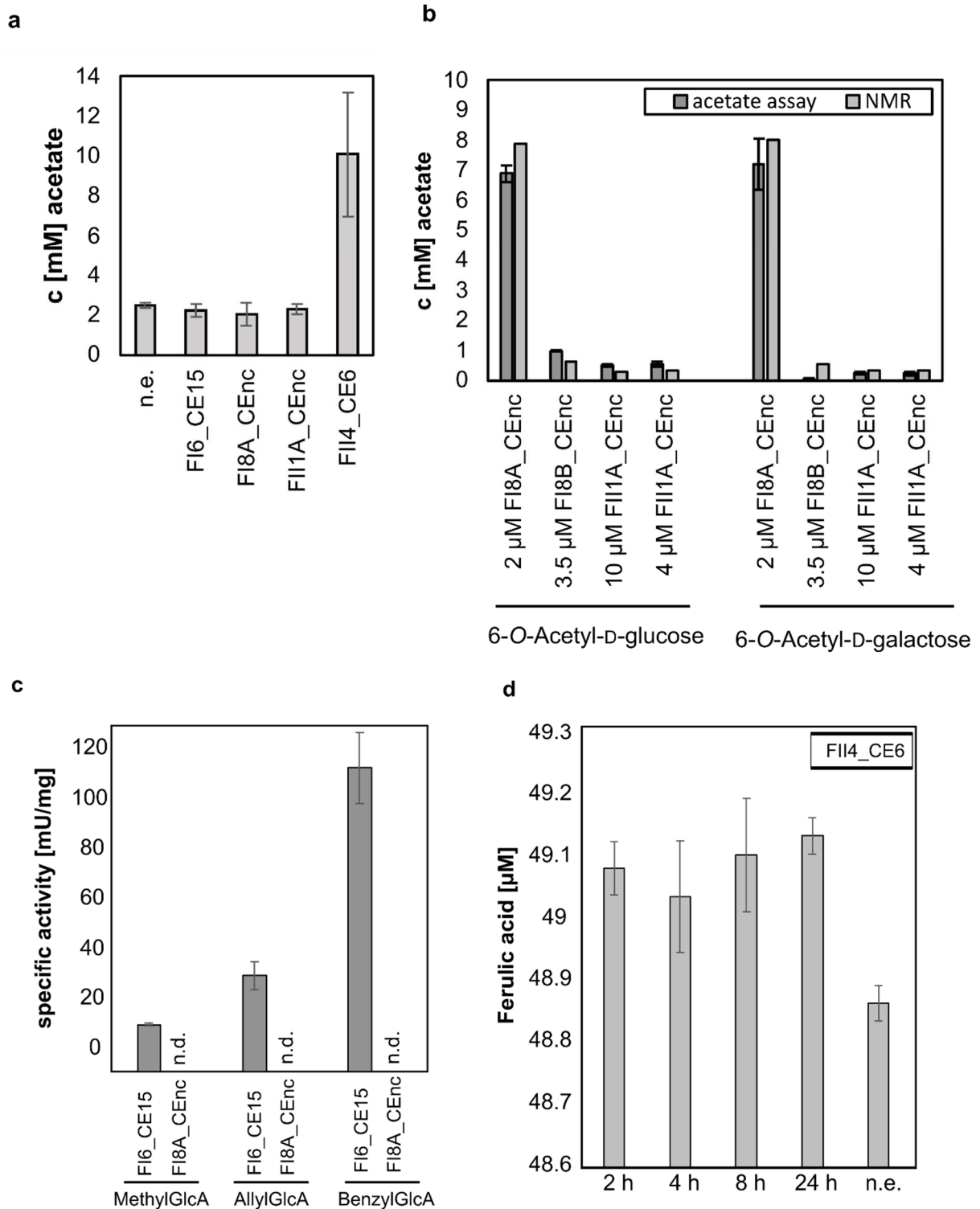


Fig. S10. Analysis of the carbohydrate esterase activities. (a) Release of acetate from polymeric partially acetylated birchwood xylan (Megazymes) measured with the R-Biopharm acetate assay. (b) Release of acetate from monomeric 6-O-acetylglucose and 6-O-acetylgalactosamine measured with the Acetate-Kit (Megazymes) with the additional proof via NMR measurement quantification. (c) Measurement of D-glucuronic acid hydrolysis with the esterases FI6_CE15 and FI8_CE6 with methyl-D-glucuronic acid (MethylGlcA), allyl-D-glucuronic acid (AllylGlcA) and benzyl-D-glucuronic acid (BenzylGlcA) with the K-URONIC Assay Kit (Megazymes). (d) Release of ferulic acid by FI14_CE6 from insoluble wheat arabinoxylan insoluble fraction (WAX-I) measured and quantified via U-HPLC DAD 320 nm.

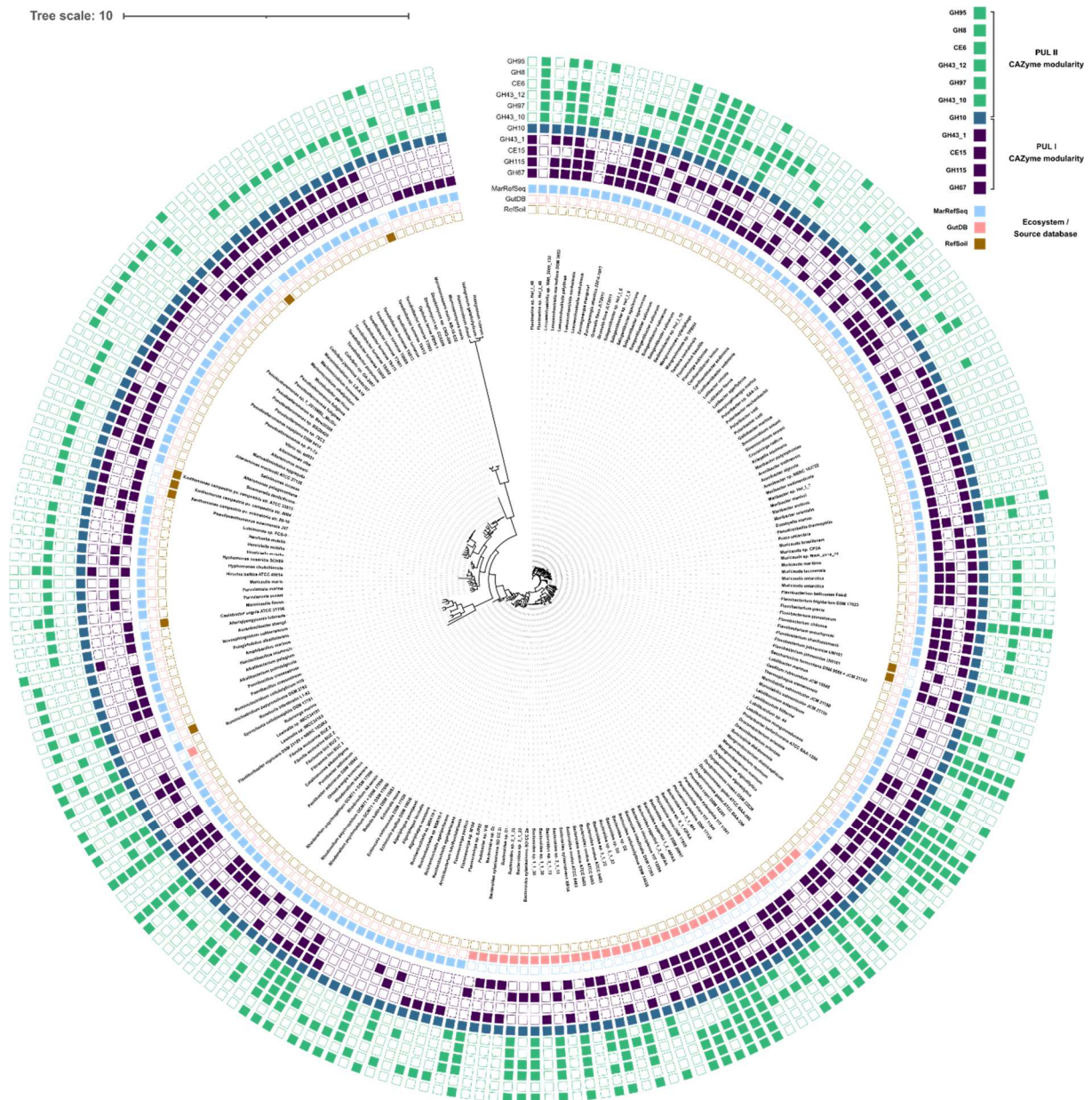


Fig. S11. Taxonomy and xylan PUL modularity of 226 bacteria with PULs similar to *Flavimarina* sp. The phylogenetic tree is based on *rpoB* sequences. The inner rings represent the habitat from which the individual bacteria were initially isolated. Expanded from Fig. 4, modularity of all identified clusters sharing at least a GH10 (blue) and two additional enzymes with the *Flavimarina* sp. PULs (PUL I purple, PUL II green) is depicted on the outer rings.

References

1. Mystkowska AA, Robb C, Vidal-Melgosa S, Vanni C, Fernandez-Guerra A, Höhne M, et al. Molecular recognition of the β -glucans laminarin and pustulan by a SusD-like glycan-binding protein of a marine *Bacteroidetes*. *FEBS J* 2018; **285**: 4465–4481.
2. Zhang Y, Werling U, Edelman W. SLiCE: a novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Res* 2012; **40**: e55.

3. Butt TR, Edavettal SC, Hall JP, Mattern MR. SUMO fusion technology for difficult-to-express proteins. *Protein Expression and Purification* 2005; **43**: 1–9.
4. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed). *The Proteomics Protocols Handbook*. 2005. Humana Press, Totowa, NJ, pp 571–607.
5. Engel A, Händel N. A novel protocol for determining the concentration and composition of sugars in particulate and in high molecular weight dissolved organic matter (HMW-DOM) in seawater. *Marine Chemistry* 2011; **127**: 180–191.
6. Bowers GN Jr, McComb RB, Upreti A. 4-nitrophenyl phosphate--characterization of high-purity materials for measuring alkaline phosphatase activity in human serum. *Clinical Chemistry* 1981; **27**: 135–143.
7. Cubillos-Rojas M, Amair-Pinedo F, Tato I, Bartrons R, Ventura F, Rosa JL. Tris-acetate polyacrylamide gradient gels for the simultaneous electrophoretic analysis of proteins of very high and low molecular mass. In: Kurien BT, Scofield RH (eds). *Electrophoretic Separation of Proteins*. 2019. Springer New York, New York, NY, pp 269–277.
8. Sunner H, Charavgi M-D, Olsson L, Topakas E, Christakopoulos P. Glucuronoyl esterase screening and characterization assays utilizing commercially available benzyl glucuronic acid ester. *Molecules* 2015; **20**: 17807–17817.
9. Duff RB. Esterification of the primary alcoholic group of carbohydrates with acetic acid: a general reaction. 5.
10. Troitzsch A, Loi VV, Methling K, Zühlke D, Lalk M, Riedel K, et al. Carbon source-dependent reprogramming of anaerobic metabolism in *Staphylococcus aureus*. *J Bacteriol* 2021; **203**: e00639-20.
11. Wider G, Dreier L. Measuring protein concentrations by NMR spectroscopy. *J Am Chem Soc* 2006; **128**: 2571–2576.
12. Bernfeld P. Amylases, α and β . *Methods in enzymology*. 1955. Academic Press, pp 149–158.

Article V

Connecting Algal Polysaccharide Degradation to Formaldehyde Detoxification

Stefan Brott,^[a] François Thomas,^[b] Maike Behrens,^[a] Karen Methling,^[c] Daniel Bartosik,^[d] Theresa Dutschei,^[a] Michael Lalk,^[c] Gurvan Michel,^[b] Thomas Schweder,^[d] and Uwe T. Bornscheuer^{*[a]}

Formaldehyde is a toxic metabolite that is formed in large quantities during bacterial utilization of the methoxy sugar 6-O-methyl-D-galactose, an abundant monosaccharide in the red algal polysaccharide porphyran. Marine bacteria capable of metabolizing porphyran must therefore possess suitable detoxification systems for formaldehyde. We demonstrate here that detoxification of formaldehyde in the marine Flavobacterium

Zobellia galactanivorans proceeds via the ribulose monophosphate pathway. Simultaneously, we show that the genes encoding the key enzymes of this pathway are important for maintaining high formaldehyde resistance. Additionally, these genes are upregulated in the presence of porphyran, allowing us to connect porphyran degradation to the detoxification of formaldehyde.

Marine algae are considered to be one of the most important primary producers in the marine ecosystem and one of the largest sources of marine carbohydrates.^[1,2] Serving as energy storage and structural cell wall components, carbohydrates constitute up to 70% of algae dry mass.^[3] Compared to their terrestrial counterparts, marine polysaccharides differ in the backbone structure and side-group modifications.^[4] One bacterial phylum considered to be specialist in the degradation of high molecular weight organic matter such as marine carbohydrates are the Bacteroidetes.^[5–7] Marine Bacteroidetes harbor gene clusters which are referred to as polysaccharide utilization loci (PULs) encoding carbohydrate-active enzymes (CAZymes) as well as specific proteins for the binding and uptake of sugar units.^[5,8] Their tremendous repertoire of CAZymes allows them to depolymerize complex marine carbohydrates and utilize the

imported monosaccharides as a carbon and energy source.^[4,8] Observations that Bacteroidetes are among the first responders after micro- and macroalgal blooms are related to their abilities of rapid growth on colonizable surfaces such as macroalgae as well as their CAZymes production.^[9,10]

A model bacterium for the bioconversion of algal biomass is the marine Flavobacterium *Zobellia galactanivorans* Dsij^T, which was originally isolated from the red alga *Delesseria sanguinea* near the coast of Roscoff (Brittany, France).^[11] In-depth analysis of its complete genome and growth studies revealed that this microorganism possesses 50 PULs, is able to grow on numerous marine polysaccharides and utilizes them as a carbon source.^[12,13] Extensive biochemical studies have elucidated essential CAZymes from *Z. galactanivorans* and their roles in the complex degradation pathways for alginate and laminarin from brown algae^[14–19] as well as for carrageenan, agar and porphyran from red algae.^[20–23] Porphyran is the common name of the agar from red algae of the genus *Porphyra* and is their main cell wall polysaccharide.^[24] The porphyran backbone consists mainly of the alternating monosaccharide units 4-linked- α -L-galactose-6-sulfate (L6S) and 3-linked- β -D-galactose (Gal) or 3,6-anhydro- α -L-galactose (LA) (Figure S1, Supporting Information).^[25,26] In addition, O-methylation of D-galactose is a frequent modification that results in the presence of up to 28% of the methoxy sugar 6-O-methyl-D-galactose (G6Me) within the porphyran chain.^[24,25,27]

Considering the stability of methyl ethers, it is reasonable to assume that G6Me must first be demethylated before it can enter the cellular metabolism. We recently have demonstrated that the oxidative demethylation of G6Me is catalyzed by a cytochrome P450 monooxygenase with the appropriate redox partners ferredoxin and ferredoxin reductase.^[28] The crystal structure of the cytochrome P450 monooxygenase from *Z. galactanivorans* informed on the binding of G6Me as well as other mechanistic insights.^[29] The products of this demethylation are D-galactose and formaldehyde in equimolar amounts.^[28] However, formaldehyde formation leads to a

[a] S. Brott, M. Behrens, T. Dutschei, Prof. U. T. Bornscheuer
Department of Biotechnology & Enzyme Catalysis
Institute of Biochemistry, University of Greifswald
Greifswald 17487 (Germany)
E-mail: uwe.bornscheuer@uni-greifswald.de
Homepage: <http://biotech.uni-greifswald.de>

[b] Prof. F. Thomas, Prof. G. Michel
Laboratory of Integrative Biology of Marine Models (LBI2 M)
Station Biologique de Roscoff (SBR)
Sorbonne Université, CNRS
29688 Roscoff, Bretagne (France)

[c] Dr. K. Methling, Prof. M. Lalk
Department of Cellular Biochemistry & Metabolomics
Institute of Biochemistry, University of Greifswald
Greifswald 17487 (Germany)

[d] D. Bartosik, Prof. T. Schweder
Department of Pharmaceutical Biotechnology
Institute of Pharmacy, University of Greifswald
Greifswald 17487 (Germany)

Supporting information for this article is available on the WWW under <https://doi.org/10.1002/cbic.202200269>

© 2022 The Authors. ChemBioChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

problem for the organism, since formaldehyde is a toxic metabolite in cells due to its high reactivity as an electrophile.^[30] The polarized carbonyl group of formaldehyde can be attacked by nucleophiles such as free amine or thiol groups of amino acids^[31] and proteins^[32] or nucleic acids,^[33] resulting in protein and DNA damages and covalent cross-links.^[34]

Marine bacteria capable of degrading porphyran and utilizing G6Me should therefore possess suitable detoxification pathways for formaldehyde. Focusing on the discovery of possible pathways of formaldehyde detoxification, we first searched through the genomes of the Flavobacteria *Z. galactanivorans* Dsj1^T and *Formosa agariphila* KMM 3901^T,^[35] which possess the cytochrome P450 monooxygenases and thus catalyze the oxidative demethylation of G6Me,^[28] for genes encoding enzymes from well-known detoxification pathways. Both organisms harbor annotated genes for enzymes found in the serine and tetrahydrofolate pathways (Table S1, Supporting Information). However, unlike *F. agariphila*, *Z. galactanivorans* additionally carries the genes for the putative key enzymes of the ribulose monophosphate (RuMP) pathway. This putative RuMP pathway was proposed to be an advantageous adaptive trait for *Z. galactanivorans* to cope with the release of formaldehyde when degrading red algal cell walls.^[12] Altogether, *F. agariphila* and *Z. galactanivorans* should provide different responses to the accumulation of formaldehyde. The RuMP pathway is the most efficient pathway of formaldehyde assimilation in terms of ATP consumption and biomass yield.^[36–38] It can be divided into three parts: fixation, cleavage, and regeneration.^[39] While the cleavage and regeneration part can take place via different routes and are catalyzed by common enzymes of the central carbon cycle, the fixation of formaldehyde takes place via two unique key enzymes: a 3-hexulose-6-phosphate synthase (HPS) and a 6-phospho-3-hexuloisomerase (PHI).^[39,40] HPS, a member of the class 2 aldolases,^[41] catalyzes the Mg²⁺-dependent aldol reaction between formaldehyde and D-ribulose-5-phosphate (R5P) to give the intermediate D-arabino-3-hexulose-6-phosphate (AH6P), which is then isomerized by PHI to D-fructose-6-phosphate (F6P) (Figure 2, top).^[40] F6P is consumed in the cleavage part to generate triose phosphates such as glyceraldehyde-3-phosphate or dihydroxyacetone phosphate which then can be metabolized in the glycolysis or the Entner-Doudoroff pathway.^[39,42] Furthermore, R5P is regenerated from F6P by reactions occurring in the pentose-phosphate cycle.^[40] It has been demonstrated that the RuMP pathway can play an important role in the degradation of methoxylated lignin monomers by non-methylotrophic bacteria.^[44] However, most knowledge about the RuMP pathway originates from methylotrophic bacteria that grow on reduced C1 components such as methane or methanol, which they oxidize to formaldehyde.^[40] In addition to these components, numerous methylated sugars are present in the marine ecosystem^[43] and are thus a potential source of formaldehyde. However, the RuMP pathway has not yet been investigated in the context of marine carbohydrate degradation, we therefore aimed to investigate whether this pathway plays a role in the degradation of porphyran by Bacteroidetes.

We reasoned that the presence of the RuMP pathway in *Z. galactanivorans* should lead to an increased resistance to formaldehyde compared to *F. agariphila*. To test this hypothesis, we cultivated each organism in the presence of increasing formaldehyde concentrations. For *F. agariphila* a significant decrease in growth rate was observed at formaldehyde concentrations greater than 100 μM, while no growth was seen at 500 μM (Figure 1a). In contrast, the presence of formaldehyde concentrations up to 500 μM revealed just minor effects on the growth of *Z. galactanivorans*. However, no growth was detected in the presence of a formaldehyde concentration of 1,000 μM (Figure S2, Supporting Information). In order to prove that the increased resistance towards formaldehyde is caused by the enzymes HPS and PHI, *hxIA* and *hxIB* gene knockout strains of *Z. galactanivorans* were created and the influence of formaldehyde on their growth was investigated again. In the presence of 500 μM formaldehyde, the Δ *hxIA-hxIB* strain was unable to grow, whereas the wild-type (WT) and a control knock-out strain lacking the P450 monooxygenase encoding gene (Δ *mgd*) were able to grow normally (Figure 1b). Knockout of the *hxIA* and *hxIB* genes thus resulted in a formaldehyde-sensitive strain, which displayed normal growth behavior in the absence of formaldehyde. Both findings supported our assumption that these genes were responsible for the detoxification of formaldehyde.

After demonstrating their role for formaldehyde resistance, we were interested to know whether the genes encoding HPS and PHI were also upregulated in the presence of porphyran, considering that this is the origin of formed formaldehyde due to the oxidative demethylation of G6Me. In order to evaluate gene regulation, *Z. galactanivorans* was grown with the marine carbohydrates laminarin, agar or porphyran as a sole carbon source. The β-glucan laminarin was selected as a control considering that it is the most abundant polysaccharide in the marine ecosystem^[45] and agar was chosen as control because it may contain G6Me.^[46] The genes encoding the P450 monooxygenase (*mgd*), HPS (*hxIA*), and PHI (*hxIB*) were upregulated in the presence of porphyran compared to laminarin and agar (Figure 1c). No upregulation in the presence of agar was observed, a possible explanation for this is the absence of G6Me in agar. Upregulation of *mgd* in the presence of porphyran indicates that there is a possible source of formaldehyde, while at the same time, the upregulation of *hxIA* and *hxIB* suggests that formaldehyde detoxification via the RuMP pathway can occur.

Following this demonstration that the genes of HPS and PHI were upregulated in the presence of porphyran, we wanted to verify whether they encode enzymes that catalyze the key reactions of the RuMP pathway. Therefore, we expressed the enzymes in *Escherichia coli* and purified them (Figure S3, Supporting Information). In order to determine the activity of the enzymes, the R5P-dependent disappearance of formaldehyde was determined using the Nash reagent^[47] and the formation of F6P was monitored by an enzyme-coupled assay.^[48] In the presence of D-ribulose-5-phosphate, a decrease in the formaldehyde concentration (Figure 2) and the formation of F6P (Figure S4, Supporting Information) was observed for the

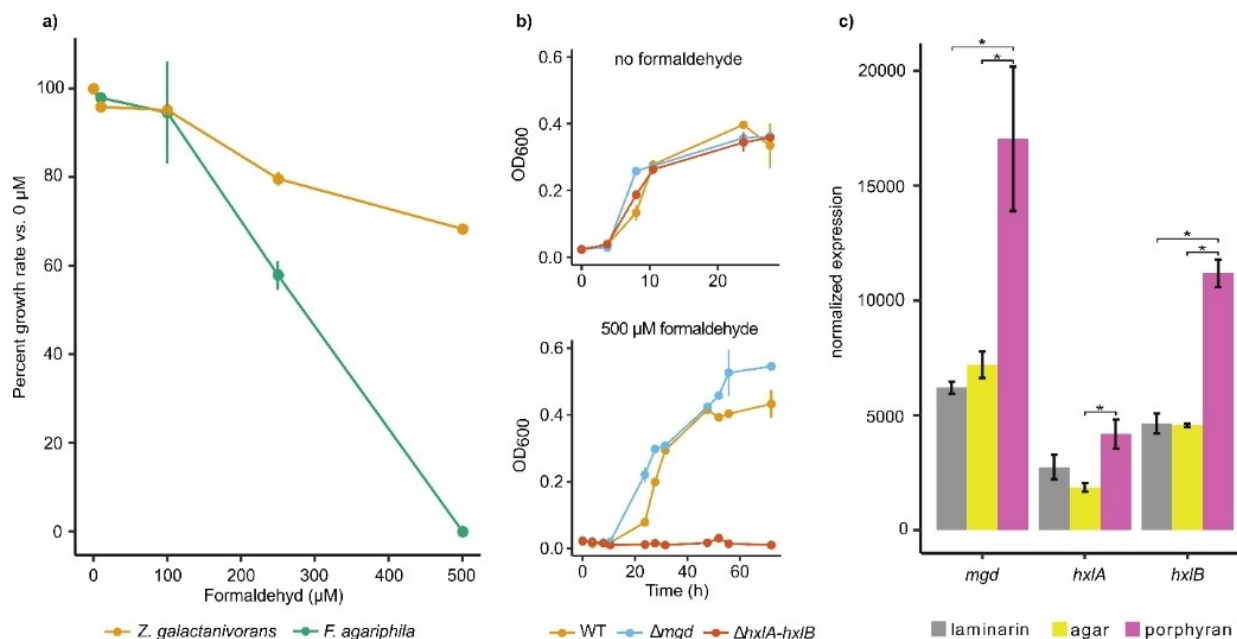


Figure 1. The genes encoding for the key enzymes of the RuMP pathway are crucial for formaldehyde resistance of *Z. galactanivorans* and are upregulated in the presence of porphyran. a) Effect of increasing concentrations of formaldehyde on the growth of *F. agariphila* and *Z. galactanivorans*. For each bacterial strain, the growth rate obtained in the absence of formaldehyde was taken as 100%. b) Growth curve of WT, Δmgd (cytochrome P450 monooxygenase) and $\Delta\text{hxlA-hxlB}$ (HPS and PHI) mutant strains of *Z. galactanivorans* in ZoBell 2216 medium containing no or 500 μM formaldehyde. c) Expression of genes encoding cytochrome P450 monooxygenase (*mgd*), 3-hexulose-6-phosphate synthase (*hxlA*) and 6-phospho-3-hexulose isomerase (*hxlB*) in *Z. galactanivorans* grown with laminarin, agar or porphyran as sole carbon source. The effect of substrate on gene expression was tested by one-way ANOVA on log-transformed data, followed by a post-hoc Tukey test (*, $P < 0.05$). Expression data from the publicly available GEO dataset GSE99940. For a)–c) Values are mean \pm s.e.m (n = 3).

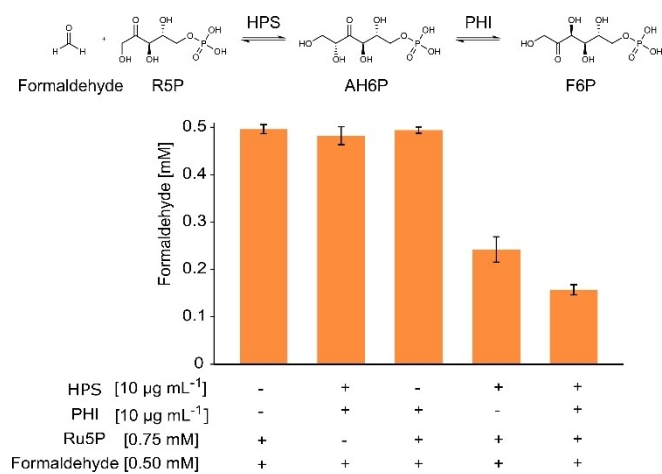


Figure 2. HPS and PHI catalyze the incorporation of formaldehyde to produce fructose-6-phosphate. A protein concentration of 10 $\mu\text{g mL}^{-1}$ for HPS and PHI were used in the biocatalysis. For substrates, 0.75 mM D-ribulose-5-phosphate disodium salt and 0.5 mM formaldehyde were used. The reactions were performed in a 50 mM sodium phosphate buffer pH 7.5 supplemented with 5 mM MgCl_2 for 5 min, at an incubation temperature of 30 $^\circ\text{C}$ and an agitation of 1,000 rpm. The formaldehyde concentration was then determined using the Nash reagent. Mean values are shown, error bars present \pm s.d. (n = 3).

reaction mixture that contained both enzymes. After 5 min incubation at 30 $^\circ\text{C}$, approximately 0.34 mM of the initial formaldehyde concentration of 0.5 mM was removed from the solution, which corresponds to a conversion of 68.5%. Mean-

while, in the control reactions without D-ribulose-5-phosphate and in the absence of either HPS or both enzymes no incorporation of formaldehyde and thus no formation of F6P was observed. For the reaction mixture containing HPS but not PHI, a decrease in formaldehyde could also be detected, which is reasonable considering that HPS catalyzes the reaction of R5P to AH6P independently of PHI. Moreover, the enzymes were able to catalyze the reverse reaction, since very low formation of formaldehyde was observed when F6P was used at a substrate concentration of 20 mM (Figure S5, Supporting Information). In conclusion, *Z. galactanivorans* harbors the active key enzymes of the RuMP pathway.

Since we could prove that *Z. galactanivorans* utilizes the RuMP pathway for formaldehyde detoxification, we were interested in the distribution of this pathway in marine ecosystems. We therefore queried approximately 5,500 marine bacterial genomes from the MarDB and MarRef databases for the key enzymes of the RuMP pathway and identified 197 genomes (equivalent to $\sim 3.58\%$) encoding HPS- and PHI- gene pairs (Figure 3). Among the 197 genomes, only 16 contain similar cytochrome P450 monooxygenase, ferredoxin reductase, and ferredoxin encoding clusters like *Z. galactanivorans* (Figure 3). The key enzymes of the RuMP pathway as well as the enzymes of the cytochrome P450 cluster were highly similar to those of *Z. galactanivorans*, which is exemplarily shown for five selected reference genomes, including *Cellulophaga*, *Maribacter*, and *Zobellia* in Figure 4.

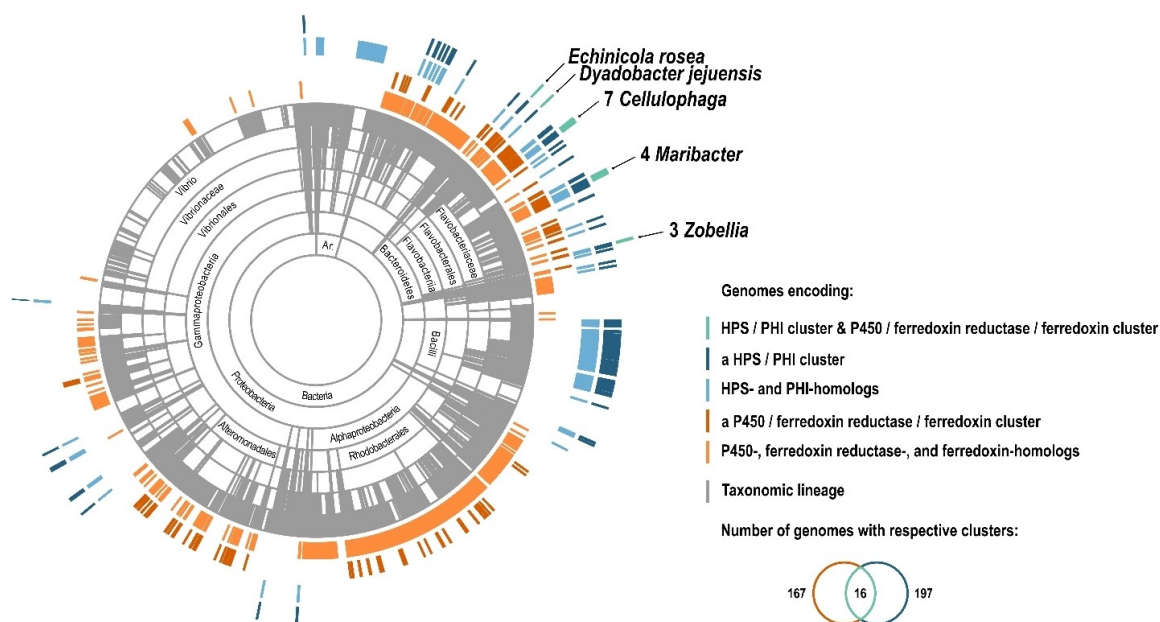


Figure 3. Taxonomic distribution of the RuMP pathway in marine prokaryotes. The colored outer rings indicate the occurrence of the HPS/PHI pairs (dark blue) and the P450 cluster (dark orange). Genomes that encode homologous sequences are shown independently (lighter colors). The intersection of genomes encoding both clusters is shown in green.

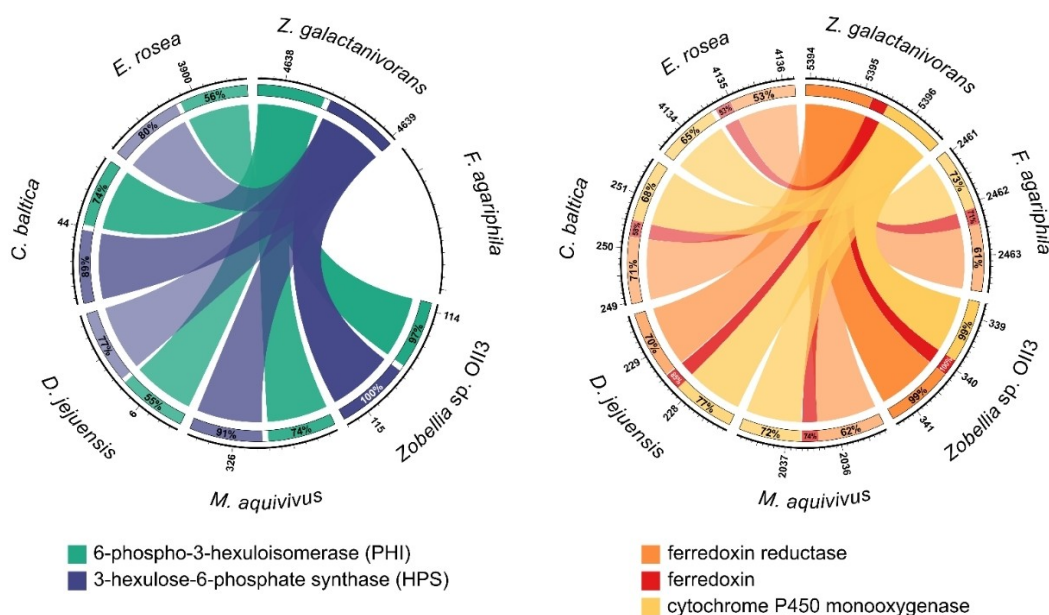


Figure 4. Key enzymes of the RuMP pathway and the enzymes of the P450 cluster from *Zobellia galactanivorans* are highly similar to those in five selected reference genomes of other marine taxa. The similarity is indicated by the opacity of each link as well as the given percentage within each coding sequence (CDS). The outer scale shows the genomic region of the CDS in kbp.

In addition, these 16 bacterial genomes also featured CAZymes belonging to the GH86 and GH117 families which can catalyze the degradation of agar and porphyran. This supports the hypothesis that the RuMP pathway may be responsible for the detoxification of formaldehyde, which is produced during the degradation of marine carbohydrates and thus may provide growth advantages for these marine bacteria over others.

Besides the marine strains with genomically clustered RuMP-based detoxification genes, we found 104 additional marine isolates where putative HPS and PHI homologs are distributed over the genomes. Interestingly the best hits are found for some *Zobellia*, a few *Maribacter*, and *Cellulophaga* as well as *Arenibacter* strains, which are bacterial genera commonly isolated at the surface of macroalgae.^[49,50] This suggests that

marine RuMP-based detoxification is mainly specific to bacteria living on multicellular algae, reminiscent of methylotrophic bacteria of the phyllosphere.^[51]

In summary, we demonstrated in this work that *Z. galactanivorans* exhibited higher resistance to formaldehyde than *F. agariphila* and that this was based on the presence of the RuMP pathway. Consequently, the knockout of the genes, encoding the key enzymes of this pathway, led to a formaldehyde-sensitive strain. We could also demonstrate that in the presence of porphyran the genes encoding the cytochrome P450 monooxygenase and the RuMP pathway were upregulated. This revealed that there is a potential source of formaldehyde through the oxidative demethylation of G6Me and simultaneously a possibility for its detoxification via the RuMP pathway. By verifying the enzyme activity of expressed and purified HPS and PHI, we could demonstrate that the genes encoding the enzymes are indeed responsible for the fixation of formaldehyde. As a result, we were able to provide evidence for a connection between porphyran degradation and formaldehyde detoxification. Genomic analyses in marine genome databases revealed that this pathway is the exception rather than the rule in marine microbes. It may thus provide growth advantages for some marine bacteria over others in the competition for marine polysaccharides.

Acknowledgements

The authors thank Nolwen Le Duff for technical assistance with microbiology experiments. F.T. acknowledges support from the French government via the National Research Agency program ALGAVOR (ANR-18-CE02-0001). D.B. was supported by a scholarship from the Institute of Marine Biotechnology e.V. We thank the German Research Foundation (DFG) for funding through the Research Unit FOR2406 "Proteogenomics of Marine Polysaccharide Utilization" (POMPU) (grant# BO 1862/17-2 to U.T.B. and SCHW 595/10-2 to T.S.). Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: Bacteroidetes · carbohydrates · CAZymes
formaldehyde detoxification · RuMP pathway

[1] D. Krause-Jensen, C. M. Duarte, *Nat. Geosci.* **2016**, *9*, 737–742.

[2] C. Arnosti, M. Wietz, T. Brinkhoff, J. H. Hehemann, D. Probandt, L. Zeugner, R. Amann, *Annu. Rev. Mar. Sci.* **2021**, *13*, 81–108.

- [3] A. Usman, S. Khalid, A. Usman, Z. Hussain, Y. Wang, in *Algae Based Polymers, Blends and Compositions: Chemistry, Biotechnology and Material Sciences* (Eds.: K. M. Zia, M. Zuber, M. Ali), Elsevier, Amsterdam, **2017**, pp. 115–153.
- [4] M. Bäumgen, T. Dutschei, U. T. Bornscheuer, *ChemBioChem* **2021**, *22*, 2247–2256.
- [5] F. Thomas, J. H. Hehemann, E. Rebuffet, M. Czjzek, G. Michel, *Front. Microbiol.* **2011**, *2*, 93.
- [6] J. M. Grondin, K. Tamura, G. Déjean, D. W. Abbott, H. Brumer, *J. Bacteriol.* **2017**, *199*, 15.
- [7] R. Munoz, R. Rosselló-Móra, R. Amann, *Syst. Appl. Microbiol.* **2016**, *39*, 281–296.
- [8] P. Lapébie, V. Lombard, E. Drula, N. Terrapon, B. Henrissat, *Nat. Commun.* **2019**, *10*, 2043.
- [9] H. Teeling, B. M. Fuchs, D. Becher, C. Klockow, A. Gardebrecht, C. M. Bennke, M. Kassabgy, S. Huang, A. J. Mann, J. Waldmann, M. Weber, A. Klindworth, A. Otto, J. Lange, J. Bernhardt, C. Reinsch, M. Hecker, J. Peplies, F. D. Bockelmann, U. Callies, G. Gerdtts, A. Wichels, K. H. Wiltshire, F. O. Glöckner, T. Schweder, R. Amann, *Science* **2012**, *336*, 608–611.
- [10] M. Brunet, F. de Bettignies, N. Le Duff, G. Tanguy, D. Davoult, C. Leblanc, A. Gobet, F. Thomas, *Environ. Microbiol.* **2021**, *23*, 1638–1655.
- [11] T. Barbeyron, S. L'Haridon, E. Corre, B. Kloareg, P. Potin, *Int. J. Syst. Evol. Microbiol.* **2001**, *51*, 985–997.
- [12] T. Barbeyron, F. Thomas, V. Barbe, H. Teeling, C. Schenowitz, C. Dossat, A. Goesmann, C. Leblanc, F. O. Glöckner, M. Czjzek, R. Amann, G. Michel, *Environ. Microbiol.* **2016**, *18*, 4610–4627.
- [13] F. Thomas, P. Bordron, D. Eveillard, G. Michel, *Front. Microbiol.* **2017**, *8*, 1808.
- [14] F. Thomas, T. Barbeyron, T. Tonon, S. Génicot, M. Czjzek, G. Michel, *Environ. Microbiol.* **2012**, *14*, 2379–2394.
- [15] F. Thomas, L. C. E. Lundqvist, M. Jam, A. Jeudy, T. Barbeyron, C. Sandström, G. Michel, M. Czjzek, *J. Biol. Chem.* **2013**, *288*, 23021–23037.
- [16] M. Dudek, A. Dieudonné, D. Jouanneau, T. Rochat, G. Michel, B. Sarels, F. Thomas, *Nucleic Acids Res.* **2020**, *48*, 7786–7800.
- [17] D. Jouanneau, L. J. Klau, R. Larocque, A. Jaffrenou, G. Duval, N. Le Duff, T. Roret, A. Jeudy, F. L. Aachmann, M. Czjzek, F. Thomas, *Glycobiology* **2021**, *31*, 1364–1377.
- [18] A. Labourel, M. Jam, A. Jeudy, J. Hehemann, M. Czjzek, G. Michel, *J. Biol. Chem.* **2014**, *289*, 2027–2042.
- [19] A. Labourel, M. Jam, L. Legentil, B. Sylla, J. H. Hehemann, V. Ferrières, M. Czjzek, G. Michel, *Acta Crystallogr. Sect. D* **2015**, *71*, 173–184.
- [20] E. Ficko-Blean, D. Duffieux, É. Rebuffet, R. Larocque, A. Groisillier, G. Michel, M. Czjzek, *Acta Crystallogr.* **2015**, *71*, 209–223.
- [21] E. Ficko-Blean, A. Préchoux, F. Thomas, T. Rochat, R. Larocque, Y. Zhu, M. Stam, S. Génicot, M. Jam, A. Calteau, B. Viart, D. Ropartz, D. Pérez-Pascual, G. Correc, M. Matard-Mann, K. A. Stubbs, H. Rogniaux, A. Jeudy, T. Barbeyron, C. Médigue, M. Czjzek, D. Vallenet, M. J. McBride, E. Ducaud, G. Michel, *Nat. Commun.* **2017**, *8*, 1685.
- [22] J. H. Hehemann, G. Correc, F. Thomas, T. Bernard, T. Barbeyron, M. Jam, W. Helbert, G. Michel, M. Czjzek, *J. Biol. Chem.* **2012**, *287*, 30571–30584.
- [23] E. Rebuffet, A. Groisillier, A. Thompson, A. Jeudy, T. Barbeyron, M. Czjzek, G. Michel, *Environ. Microbiol.* **2011**, *13*, 1253–1270.
- [24] D. A. Rees, E. Conway, *Biochem. J.* **1962**, *84*, 411–416.
- [25] N. S. Anderson, D. A. Rees, *J. Chem. Soc.* **1965**, 5880–5887.
- [26] G. Correc, J. H. Hehemann, M. Czjzek, W. Helbert, *Carbohydr. Polym.* **2011**, *83*, 277–283.
- [27] D. Ropartz, A. Giuliani, M. Fanuel, C. Hervé, M. Czjzek, H. Rogniaux, *Anal. Chim. Acta* **2016**, *933*, 1–9.
- [28] L. Reisky, H. C. Büchschütz, J. Engel, T. Song, T. Schweder, J. H. Hehemann, U. T. Bornscheuer, *Nat. Chem. Biol.* **2018**, *14*, 342–344.
- [29] C. S. Robb, L. Reisky, U. T. Bornscheuer, J. H. Hehemann, *Biochem. J.* **2018**, *475*, 3875–3886.
- [30] N. H. Chen, K. Y. Djoko, F. J. Veyrier, A. G. McEwan, *Front. Microbiol.* **2016**, *7*, 257.
- [31] J. J. A. G. Kamps, R. J. Hopkinson, C. J. Schofield, T. D. W. Claridge, *Commun. Chem.* **2019**, *2*, 126.
- [32] T. Tayri-Wilk, M. Slavin, J. Zamel, A. Blass, S. Cohen, A. Motzik, X. Sun, D. E. Shalev, O. Ram, N. Kalisman, *Nat. Commun.* **2020**, *11*, 3128.
- [33] S. Shishodia, D. Zhang, A. H. El-Sagheer, T. Brown, T. D. W. Claridge, C. J. Schofield, R. J. Hopkinson, *Org. Biomol. Chem.* **2018**, *16*, 4021–4032.
- [34] K. Lu, W. Ye, L. Zhou, L. B. Collins, X. Chen, A. Gold, L. M. Ball, J. A. Swenberg, *J. Am. Chem. Soc.* **2010**, *132*, 3388–3399.

- [35] A. J. Mann, R. L. Hahnke, S. Huang, J. Werner, P. Xing, T. Barbeyron, B. Huettel, K. Stüber, R. Reinhardt, J. Harder, F. O. Glöckner, R. I. Amann, H. Teeling, *Appl. Environ. Microbiol.* **2013**, *79*, 6813–6822.
- [36] T. Strom, T. Ferenci, J. R. Quayle, *Biochem. J.* **1974**, *144*, 465–476.
- [37] I. Goldberg, J. S. Rock, A. Ben-Bassat, R. I. Mateles, *Biotechnol. Bioeng.* **1976**, *18*, 1657–1668.
- [38] W. B. Whitaker, N. R. Sandoval, R. K. Bennett, A. G. Fast, E. T. Papoutsakis, *Curr. Opin. Biotechnol.* **2015**, *33*, 165–175.
- [39] J. E. N. Müller, F. Meyer, B. Litsanov, P. Kiefer, E. Potthoff, S. Heux, W. J. Quax, V. F. Wendisch, T. Brautaset, J. C. Portais, J. A. Vorholt, *Metab. Eng.* **2015**, *28*, 190–201.
- [40] N. Kato, H. Yurimoto, R. K. Thauer, *Biosci. Biotechnol. Biochem.* **2006**, *70*, 10–21.
- [41] S. Desmons, R. Fauré, S. Bontemps, *ACS Catal.* **2019**, *9*, 9575–9588.
- [42] H. He, C. Edlich-Muth, S. N. Lindner, A. Bar-Even, *ACS Synth. Biol.* **2018**, *7*, 1601–1611.
- [43] C. Panagiotopoulos, D. J. Repeta, L. Mathieu, J. F. Rontani, R. Sempéré, *Mar. Chem.* **2013**, *154*, 34–45.
- [44] R. Mitsui, Y. Kusano, H. Yurimoto, Y. Sakai, N. Kato, M. Tanaka, *Appl. Environ. Microbiol.* **2003**, *69*, 6128–6132.
- [45] S. Becker, J. Tebben, S. Coffinet, K. Wiltshire, M. H. Iversen, T. Harder, K. U. Hinrichs, J. H. Hehemann, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 6599–6607.
- [46] A. Chiovitti, A. Bacic, D. J. Craik, G. T. Kraft, M. L. Liao, *Carbohydr. Res.* **2004**, *339*, 1459–1466.
- [47] T. Nash, *Biochem. J.* **1953**, *55*, 416–421.
- [48] H.-J. Hohorst, *Methods Enzym. Anal.* **1965**, 134–138.
- [49] M. Martin, D. Portetelle, G. Michel, M. Vandenbol, *Appl. Microbiol. Biotechnol.* **2014**, *98*, 2917–2935.
- [50] M. Martin, T. Barbeyron, R. Martin, D. Portetelle, G. Michel, M. Vandenbol, *Front. Microbiol.* **2015**, *6*, 1487.
- [51] J. A. Vorholt, *Nat. Rev. Microbiol.* **2012**, *10*, 828–840.

Manuscript received: May 10, 2022

Accepted manuscript online: May 13, 2022

Version of record online: May 30, 2022

ChemBioChem

Supporting Information

Connecting Algal Polysaccharide Degradation to Formaldehyde Detoxification

Stefan Brott, François Thomas, Maike Behrens, Karen Methling, Daniel Bartosik, Theresa Dutschei, Michael Lalk, Gurvan Michel, Thomas Schweder, and Uwe T. Bornscheuer*

Supporting Information

Table of Contents

Experimental Procedures.....	2
Figure S1: Structure of porphyrin.....	5
Figure S2: Effect of increasing concentrations of formaldehyde on the growth of <i>F. agariphila</i> and <i>Z. galactanivorans</i>	5
Figure S3: SDS-PAGE of purified proteins.....	6
Figure S4: Coupled enzyme assay for the detection of F6P formation.....	6
Figure S5: Formaldehyde formation in the reverse reaction.....	7
Table S1: Annotated genes for <i>F. agariphila</i> and <i>Z. galactanivorans</i> from known formaldehyde detoxification pathways.....	8
Table S2: Primers used in this study.....	8
Table S3: Bacterial strains and plasmids used in this study.....	9
References.....	9
Author Contribution.....	10

Experimental Procedures

Reagents

D-Ribulose-5-phosphate and D-fructose-6-phosphate were purchased as disodium salts from Sigma Aldrich. The 37% formaldehyde solution was also purchased from Sigma Aldrich. All other chemicals were purchased at the highest purity from Sigma-Aldrich, Carl Roth, Alfa Aesar or Acros.

Gene deletions in *Z. galactanivorans*

Deletion mutants of genes encoding a cytochrome P450 monooxygenase (ZGAL_4677) and both, a putative 3-hexulose-6-phosphate synthase (ZGAL_3942, HxlA) and a 6-phospho-3-hexulose isomerase (ZGAL_3941, HxlB) were constructed using a *sacB* system described previously.^[1] All primers and strains are listed in **Tables S2** and **S3**. To delete *zgal_4677*, a 2,049 bp fragment including the first 36 bp of *zgal_4677* and 2,013 bp of upstream sequence was amplified using primers OFT0046 and OFT0048 on genomic DNA from *Z. galactanivorans* Dsij^T. The fragment was digested with BamHI and XbaI and ligated into pYT313 that had been digested with the same enzymes, to generate pFT14. A 2,222 bp fragment including the final 57 bp of *zgal_4677* and 2,165 bp of downstream sequence was amplified using primers OFT0047 and OFT0049. The fragment was cloned into XbaI and PstI sites of pFT2 to generate the *zgal_4677* deletion construct pFT15. To delete *zgal_3941* and *zgal_3942*, a 2,013 bp fragment including the first 27 bp of *zgal_3941* and 1,986 bp of upstream sequence was amplified using primers OFT0052 and OFT0054. The fragment was digested with XbaI and Sall and ligated into pYT313 that had been digested with the same enzymes, to generate pFT16. A 1,491 bp fragment including the final 27 bp of *zgal_3942* and 1,464 bp of downstream sequence was amplified using primers OFT0053 and OFT0055. The fragment was cloned into Sall and PstI sites of pFT16 to generate the *zgal_3941-3942* deletion construct pFT17. Plasmids pFT15 and pFT17 were introduced individually into the wild-type *Z. galactanivorans* Dsij^T by conjugation from *E. coli* S17-1. Conjugants with plasmids integrated in the genome were isolated on Cytophaga-agar containing 50 µg mL⁻¹ erythromycin. Single erythromycin-resistant colonies were grown overnight at 30 °C in Cytophaga medium without antibiotics. Cells that lost the plasmid through a second recombination were selected on Cytophaga-agar containing 5% sucrose. Isolated colonies were checked for erythromycin sensitivity. Deletions were confirmed by PCR and sequencing on isolated colonies using primer pairs OFT0050- OFT0051 to identify the *zgal_4677* deletion mutant (mZG_0084), and primers OFT0056-OFT0057 to identify the *zgal_3941-3942* deletion mutant (mZG_0082).

Bacterial growth

F. agariphila KMM 3901^T and *Z. galactanivorans* Dsij^T strains were routinely grown from glycerol stocks in Zobell 2216E medium at 25 °C. Their resistance to formaldehyde was tested by inoculating them (initial OD₆₀₀ 0.05) in 50 mL flasks containing 5 mL Zobell 2216E medium with increasing initial formaldehyde concentration (10 µM – 1 mM). To test the effect of gene deletions, *Z. galactanivorans* WT, mZG0082 (Δ *zgal_3941-3942*) and mZG0084 (Δ *zgal_4677*) strains were grown in 5 mL Zobell 2216E medium with or without 500 µM formaldehyde. All tests were performed in triplicates and growth was followed by monitoring OD₆₀₀ on 180 µL of culture using a microplate spectrophotometer (Spark Tecan, Männedorf, Switzerland).

Gene expression analysis

Expression data were retrieved from a previously published study^[2] with a publicly available GEO dataset GSE99940. Briefly, *Z. galactanivorans* Dsij^T was grown in marine minimum medium with 2 g L⁻¹ laminarin, agar or porphyrin as a sole carbon source. After 48 h, RNA was retrieved from cells for cDNA synthesis and analyzed on a custom microarray. The effect of substrate on gene expression was tested by one-way ANOVA on normalized and log-transformed data, followed by a post-hoc Tukey test.

Cloning and expression in *E. coli*

Synthetic genes, codon optimized for expression in *E. coli*, encoding HPS and PHI from *Z. galactanivorans*, were synthesized and cloned into the pET-51b vector by BioCat GmbH (Heidelberg, Germany). The constructs encoded the recombinant proteins as fusions to a cleavable N-terminal Strep-tag for affinity purification. Chemically competent *E. coli* BL21 (DE3) cells were transformed with the plasmids harboring HPS or PHI and were spread on lysogeny broth (LB) agar plates containing 100 µg mL⁻¹ ampicillin. The agar plates were incubated overnight at 37 °C. One colony was picked and used to inoculate 5 mL LB medium which contained 100 µg mL⁻¹ ampicillin and was then incubated at 37 °C and 180 rpm overnight. For overexpression the cultivation was performed with 200 mL LB medium containing 100 µg mL⁻¹ ampicillin in a 1 L flask. The LB medium was inoculated with the overnight culture so that a starting optical density (OD₆₀₀) of 0.05 was obtained. Cells were then incubated at 37 °C and 180 rpm until an OD₆₀₀ of 1 was reached. Expression of target enzymes was then induced by the addition of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). For the expression of HPS, 1 mM MgCl₂ was supplemented simultaneously and the cultivation was then continued at 25 °C and 180 rpm overnight. For PHI, cultivation was subsequently continued at 20 °C and 180 rpm overnight after the addition of IPTG. Cells were harvested by centrifugation at 10,000 x g and 4 °C for 30 min, washed with 50 mM sodium phosphate buffer pH 7.5, and subsequently stored at -20 °C until cell disruption.

Purification

For cell disruption, the cell pellet was resuspended in 10 mL of lysis buffer (100 mM TRIS-HCl buffer (pH 8.0) containing 500 mM NaCl, 5 mM MgCl₂, 0.1% Triton-X-100, and 1 mM phenylmethylsulfonyl fluoride). Cell disruption on ice was performed using a Sonoplus HD 2070 ultrasonic homogenizer (Bandelin electronic GmbH & Co. KG, Berlin, Germany) with the program: 2 × 3 min, 50% power, 50% cycle time. Cell debris was subsequently removed by centrifugation at 10,000 x g and 4 °C for 30 min. Purification utilizing gravity flow columns was performed using 10 mL of the Strep-Tactin® Sepharose® 50% suspension (IBA Lifesciences GmbH, Göttingen, Germany) as column material. After equilibration of the column with the wash buffer (100 mM TRIS-HCl buffer (pH 8.0), which contained 500 mM NaCl, 5 mM MgCl₂ and 0.1% Triton-X-100), the clarified lysate was applied. Unbound proteins were then removed from the column by excessive washing with the wash buffer. Elution of the target enzymes was then performed with the elution buffer 100 mM TRIS-HCl buffer (pH 8.0), which contained 2.5 mM D-desthiobiotin in addition to 500 mM NaCl and 5 mM MgCl₂. Elution fractions were pooled and concentrated using a Vivaspin 6 centrifugal concentrator with a 10 kDa molecular weight cut-off (Sartorius AG, Göttingen, Germany). PD-10 desalting columns (Cytiva Europe GmbH, Freiburg, Germany) were then used to desalt the sample and exchange the elution buffer to a 50 mM sodium phosphate buffer pH 7.5 supplemented with 5 mM MgCl₂.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed to verify the purity of the target enzymes. 20 µL protein sample was mixed with 5 µL of a 5-fold stock of SDS sample buffer (100 mM TRIS-HCl buffer (pH 6.8) containing 4% w/v SDS, 20% v/v glycerol, 2% (v/v) β-mercaptoethanol, 25 mM EDTA and 0.04% w/v bromophenol blue) and denatured by incubation at 95 °C for 10 min. For the SDS-PAGE a 12.5% acrylamide gel (separating gel) and a 4.0% loading gel were used. Electrophoresis was carried out at 200 V. Proteins were stained with Coomassie Blue (PhastGel® Blue R, Sigma Aldrich, Taufkirchen, Germany). As reference the Pierce™ Unstained protein molecular weight marker (Thermo Scientific, Waltham, MA, USA) was used.

Determination of protein concentration

Protein concentrations were determined using the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA).

Enzyme assays

Activity of HPS was assayed by the D-ribulose-5-phosphate-dependent disappearance of formaldehyde and by the formation of F6P. A protein concentration of 10 µg mL⁻¹ for HPS and PHI were used in the biocatalysis. As substrates, 0.75 mM D-ribulose-5-phosphate disodium salt and 0.5 mM formaldehyde were used. The reaction volume was 0.2 mL and the reactions were performed in a 50 mM sodium phosphate buffer (pH 7.5) supplemented with 5 mM MgCl₂ for 5 min at an incubation temperature of 30 °C and an agitation of 1000 rpm. The formaldehyde concentration was then determined using the Nash reagent.^[3]

In the reverse reaction a protein concentration of 50 µg mL⁻¹ for each enzyme and 20 mM D-fructose-6-phosphate disodium salt as substrate were used. The reaction volume was 0.2 mL and the reactions were performed in a 50 mM sodium phosphate buffer pH 7.5 supplemented with 5 mM MgCl₂ for 10 minutes at an incubation temperature of 30 °C and an agitation of 1000 rpm. Formaldehyde formation was then detected using the Nash reagent.^[3]

Additionally, F6P formation was detected by coupling HPS and PHI with the phosphoglucose isomerase (PGI) from yeast (Roche Holding AG, Basel, Switzerland) and the glucose-6-phosphate dehydrogenase (G6pDHG) from baker's yeast Type XV (Sigma Aldrich, St. Louis, MO, USA).^[4] A protein concentration of 10 µg mL⁻¹ was used for HPS and PHI, and 5 U mL⁻¹ were used for PGI and G6pDHG. For substrates, 0.75 mM D-ribulose-5-phosphate disodium salt, 0.5 mM formaldehyde and 0.5 mM NADP⁺ were used. The reaction was carried out in a 50 mM TRIS-HCl buffer (pH 7.5) containing 5 mM MgCl₂ at 30 °C. The absorbance at 340 nm was measured every 2 min using the Infinite® M200 pro microplate reader (Tecan Group Ltd., Männedorf, Switzerland).

Formaldehyde quantification

For the determination of the formaldehyde concentration, 0.2 mL freshly prepared Nash reagent^[3] was added immediately to the reaction mixture, this mix was then immediately incubated at 50 °C at an agitation of 1,000 rpm for 10 min. After centrifugation at 17,000 x *g* for 2 min to remove precipitated proteins, the mixture was transferred to a microtiter plate and absorbance was measured at 420 nm. The formaldehyde concentration was then determined by a formaldehyde standard curve, which was prepared in parallel to the reaction mixture.

Computational analysis

RefSeq assemblies of genomes deposited in MarRef (v1.5) and MarDB (v1.5)^[5] were downloaded from NCBI^[6] to create a target database. Translated coding sequences were compared to TIGRFAM profiles TIGR03127.1 (6-phospho-3-hexuloisomerase) and TIGR03128.1 (3-hexulose-6-phosphate synthase) as well as to PFAM models PF07992.17 together with PF14759.9 (to aim for ferredoxin reductase homologs, ZOBGAL_RS21970), PF00111.30 (for putative ferredoxins, ZOBGAL_RS21975) and PF00067.25 (for putative cytochrome P450 monooxygenases, ZOBGAL_RS21980) using the hmmscan function of HMMER v3.3.2^[7] with model-specific noise cutoff threshold (--cut_nc). Results were then compared to *Z. galactanivorans* Dsij^T sequences ZOBGAL_RS18540 (HPS), ZOBGAL_RS18545 (PHI), ZOBGAL_RS21970, ZOBGAL_RS21975, and ZOBGAL_RS21980 using Protein-Protein BLAST v2.11.0+^[8] with default settings. Circos was used to visualize similarity on protein level^[9].

Supporting Figures

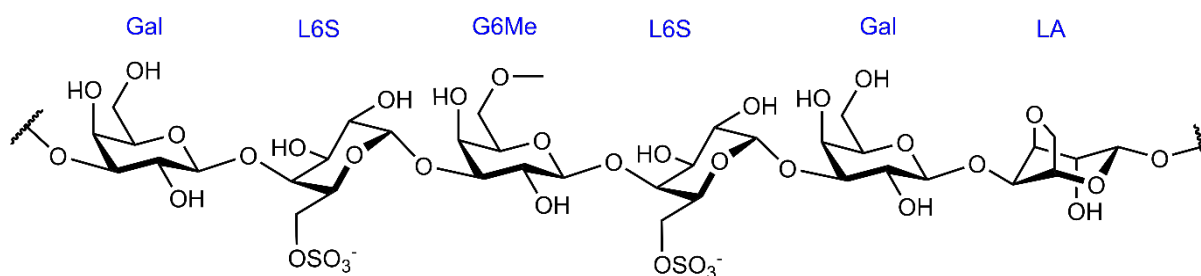


Figure S1: Structure of porphyran. The porphyran backbone consists of chains composed mainly of the alternating monosaccharide units 4-linked- α -L-galactose-6-sulfate (L6S) and 3-linked- β -D-galactose (Gal) or 3,6-anhydro- α -L-galactose (LA). In addition, O-methylation of D-galactose results in the presence of the methoxy sugar 6-O-methyl-D-galactose (G6Me).

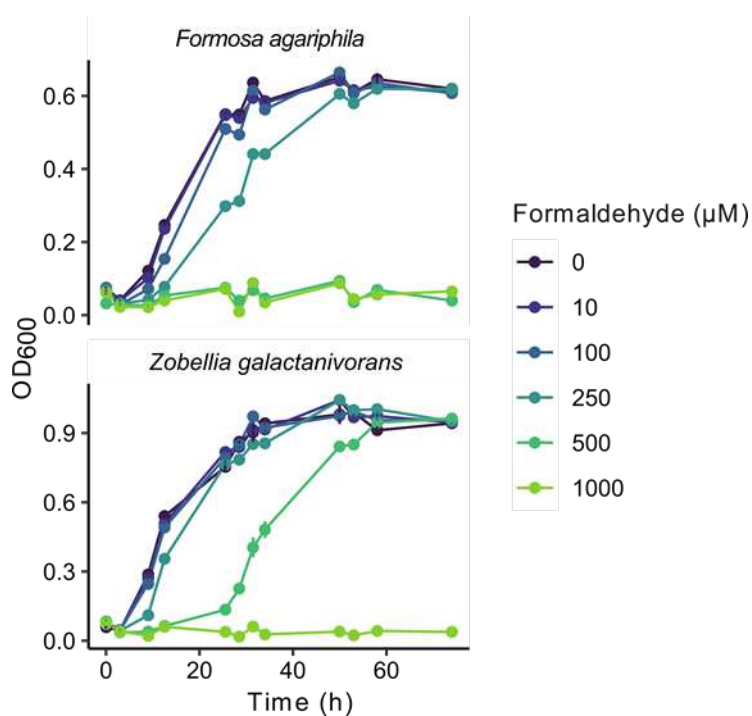


Figure S2: Effect of increasing concentrations of formaldehyde on the growth of *F. agariphila* and *Z. galactanivorans*. Growth was performed with Zobell 2216E medium with increasing initial formaldehyde concentration at 25 °C. Values are mean \pm s.e.m. (n = 3).

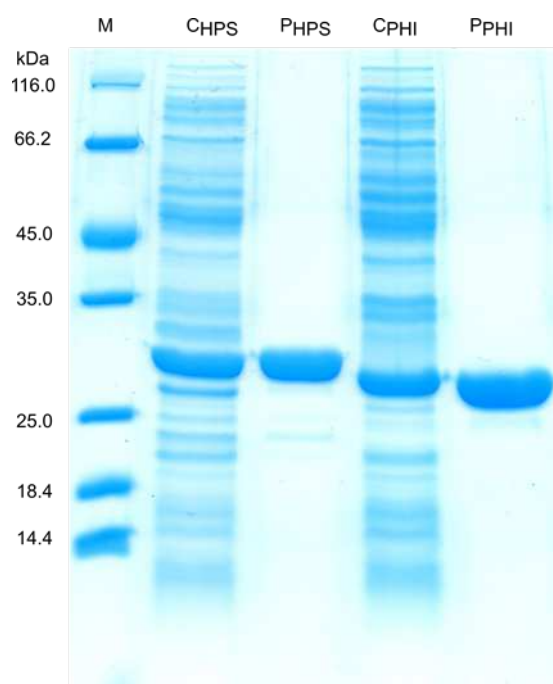


Figure S3: SDS-PAGE of purified proteins. The purified proteins (P) and the crude cell extract (C) were separated on a 12.5% gel and stained with Coomassie blue. 7.5 μg of the proteins were loaded onto the gel. As reference (M) the Pierce™ Unstained protein molecular weight marker (Thermo Scientific, Waltham, MA, USA) was used. The experiment was repeated independently with similar results.

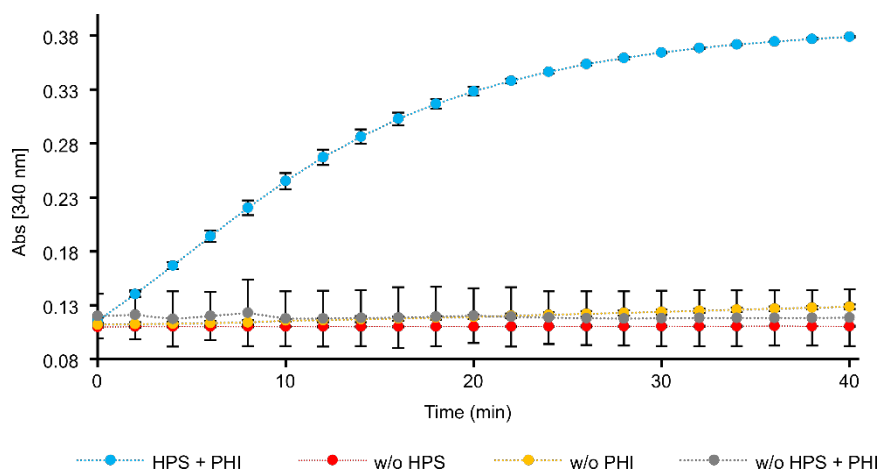


Figure S4: Coupled enzyme assay for the detection of F6P formation. Formed F6P is converted by PGI to glucose-6-phosphate, which is then oxidized by G6pDHG under NADP^+ consumption to D-glucono-1,5-lactone-6-phosphate. A protein concentration of $10 \mu\text{g mL}^{-1}$ was used for HPS and PHI, and 5 U mL^{-1} was used for PGI and G6pDHG. For substrates, 0.75 mM D-ribulose-5-phosphate, 0.5 mM formaldehyde and 0.5 mM NADP^+ were used. The reaction was carried out in a 50 mM TRIS-HCl buffer (pH 7.5) containing 5 mM MgCl_2 at $30 \text{ }^\circ\text{C}$. The absorbance at 340 nm was measured every 2 min using the Infinite® M200 pro microplate reader. Mean values are shown, error bars present \pm s.d. ($n = 3$).

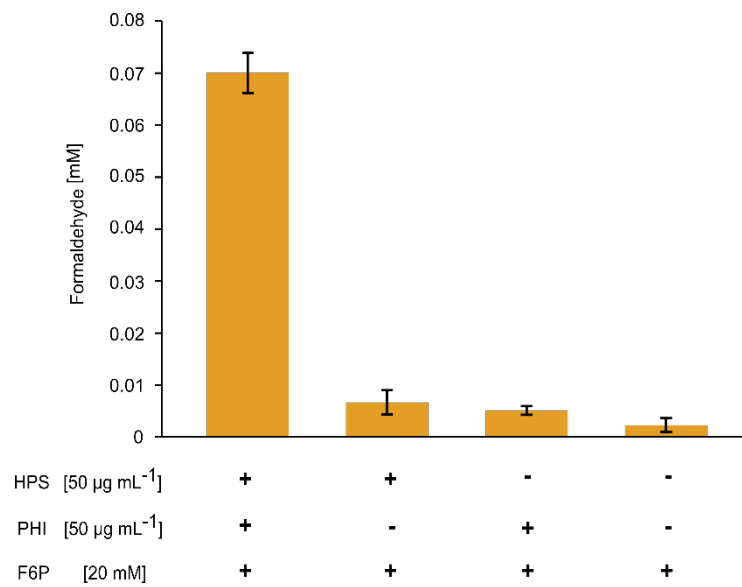


Figure S5: Formaldehyde formation in the reverse reaction. A protein concentration of 50 $\mu\text{g mL}^{-1}$ for each enzyme and 20 mM D-fructose-6-phosphate as substrate were used. The reaction volume was 0.2 mL and the reactions were performed in a 50 mM sodium phosphate buffer pH 7.5 supplemented with 5 mM MgCl_2 for 10 min at an incubation temperature of 30 °C and an agitation of 1,000 rpm. Mean values are shown, error bars present \pm s.d. (n = 3).

Supporting Tables

Table S1: Annotated genes found in the NCBI database for *F. agariphila* KMM3901^T and *Z. galactanivorans* Dsij^T, which encode enzymes involved in known formaldehyde detoxification pathways. The accession number is given for each gene.

Annotated enzyme activity	<i>F. agariphila</i> KMM3901 ^T	<i>Z. galactanivorans</i> Dsij ^T
3-Hexulose-6-phosphate synthase	-	WP_013995270.1
6-Phospho-3-hexuloisomerase	-	WP_013995269.1
Serine hydroxymethyltransferase	WP_038533459.1	WP_013996064.1
Bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase	WP_038527446.1	WP_013993916.1
Formyltetrahydrofolate deformylase	WP_038527618.1	WP_215931961.1

Table S2: Primers used in this study.

Primers	Sequence and Description
OFT0046	5' TTTTTTGGATCC ^{TCCTTATAGTCGGGTATATCAAGG} 3'; forward primer used in construction of pFT14; BamHI site underlined
OFT0047	5' GAGTTGACCACAGAACCATAACC 3'; reverse primer used in construction of pFT15; PstI site downstream in amplified fragment
OFT0048	5' TTTTTTCTAGATTTTCAAACGGGTCTGGAAGC 3'; reverse primer used in construction of pFT14; XbaI site underlined
OFT0049	5' TTTTTTCTAGACAGCGTAAAGTAGGTTTTCATAAC 3'; forward primer used in construction of pFT15; XbaI site underlined
OFT0050	5' GGCTCTAATATGGGTTGCATCCG 3'; forward primer to confirm deletion of <i>zgal_4677</i>
OFT0051	5' ATATCGGTCTCTATCTCACTGGC 3'; reverse primer to confirm deletion of <i>zgal_4677</i>
OFT0052	5' TTTTTTCTAGAAGTTGGCATTGAAAGCTGTAGG 3'; forward primer used in construction of pFT16; XbaI site underlined
OFT0053	5' TTTTTTCTGCAGCAGAAGTAAAAATCCAATGACTTTTAGC 3'; reverse primer used in construction of pFT17; PstI site underlined
OFT0054	5' TTTTTTGTCGACCTTGCTTTCATCAAGTATGTTCTCC 3'; reverse primer used in construction of pFT16; Sall site underlined
OFT0055	5' TTTTTTGTCGACGAGCTAAAGGAATTATTGGAAGCC 3'; forward primer used in construction of pFT17; Sall site underlined
OFT0056	5' CGGACGAGGGGTTAAATAGCC 3'; forward primer to confirm deletion of <i>zgal_3941-3942</i>
OFT0057	5' TTCGTCTTTTGAATTATGAGGAGGC 3'; reverse primer to confirm deletion of <i>zgal_3941-3942</i>

Table S3: Bacterial strains and plasmids used in this study.

	Description ^a	Ref.
<i>E. coli</i> strains		
NEB5α	Strain used for general cloning Genotype: <i>fhuA2 (argF-lacZ)U169 phoA glnV44 80 (lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17</i>	New England Biolabs (Ipswich, MA, USA)
S17-1 λ pir	Strain used for conjugation with <i>Z. galactanivorans</i> Genotype: <i>λpir hsdR pro thi</i> ; chromosomal integrated RP4-2 Tc::Mu Km::Tn7	[10]
Marine strains		
<i>F. agariphila</i> KMM 3901 ^T	wild type <i>F. agariphila</i> strain	[11]
<i>Z. galactanivorans</i> Dsij ^T (DSM 12802)	wild type <i>Z. galactanivorans</i> strain	[12]
mZG0082	Δ <i>zgal_3941-3942</i> in <i>Z. galactanivorans</i> Dsij ^T	This study
mZG0084	Δ <i>zgal_4677</i> in <i>Z. galactanivorans</i> Dsij ^T	This study
Plasmids		
pYT313	Suicide vector carrying <i>sacB</i> under <i>F. johnsoniae ompA</i> promoter; Ap ^r (Em ^r)	[1]
pFT14	2,049 bp region upstream of <i>Z. galactanivorans zgal_4677</i> amplified with primers OFT0046 and OFT0048 and cloned into BamHI and XbaI sites of pYT313; Ap ^r (Em ^r)	This study
pFT15	Construct used to delete <i>Z. galactanivorans zgal_4677</i> ; 2,222 bp region downstream of <i>zgal_4677</i> amplified with primers OFT0047 and OFT0049 and cloned into PstI and XbaI sites of pFT14; Ap ^r (Em ^r)	This study
pFT16	2,013 bp region upstream of <i>Z. galactanivorans zgal_3941</i> amplified with primers OFT0052 and OFT0054 and cloned into XbaI and Sall sites of pYT313; Ap ^r (Em ^r)	This study
pFT17	Construct used to delete <i>Z. galactanivorans zgal_3941-3942</i> ; 1,491 bp region downstream of <i>zgal_3942</i> amplified with primers OFT0053 and OFT0055 and cloned into PstI and Sall sites of pFT16; Ap ^r (Em ^r)	This study

^aAntibiotic resistance phenotypes: ampicillin, Ap^r; erythromycin, Em^r. Antibiotic resistance phenotypes are those expressed in *E. coli*. The antibiotic resistance phenotypes given in parentheses are those expressed in *Z. galactanivorans* but not in *E. coli*.

References

- [1] Y. Zhu, F. Thomas, R. Larocque, N. Li, D. Duffieux, L. Cladière, F. Souchaud, G. Michel, M. J. McBride, *Environ. Microbiol.* **2017**, *19*, 2164–2181.
- [2] F. Thomas, P. Bordron, D. Eveillard, G. Michel, *Front. Microbiol.* **2017**, *8*, 1808.
- [3] T. Nash, *Biochem. J.* **1953**, *55*, 416–421.
- [4] H.-J. Hohorst, *Methods Enzym. Anal.* **1965**, 134–138.
- [5] T. Klemetsen, I. A. Raknes, J. Fu, A. Agafonov, S. V. Balasundaram, G. Tartari, E. Robertsen, N. P. Willassen, *Nucleic Acids Res.* **2018**, *46*, D693–D699.
- [6] E. W. Sayers, J. Beck, J. R. Brister, E. E. Bolton, K. Canese, D. C. Comeau, K. Funk, A. Ketter, S. Kim, A. Kimchi, P. A. Kitts, A. Kuznetsov, S. Lathrop, Z. Lu, K. McGarvey, T. L. Madden, T. D. Murphy, N. O'Leary, L. Phan, V. A. Schneider, F. Thibaud-Nissen, B. W. Trawick, K. D. Pruitt, J. Ostell, *Nucleic Acids Res.* **2020**, *48*, D9–D16.
- [7] R. D. Finn, J. Clements, S. R. Eddy, *Nucleic Acids Res.* **2011**, *39*, 29–37.
- [8] B. B. Salgaonkar, M. Kabilan, J. M. Braganca, *World J. Microbiol. Biotechnol.* **2011**, *27*, 403–410.
- [9] M. Krzywinski, J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, M. A. Marra, *Genome Res.* **2009**, *19*, 1639–1645.
- [10] V. de Lorenzo, K. N. Timmis, *Methods Enzymol.* **1994**, *235*, 386–405.
- [11] A. J. Mann, R. L. Hahnke, S. Huang, J. Werner, P. Xing, T. Barbeyron, B. Huettel, K. Stüber, R. Reinhardt, J. Harder, F. O. Glöckner, R. I. Amann, H. Teeling, *Appl. Environ. Microbiol.* **2013**, *79*, 6813–6822.
- [12] T. Barbeyron, S. L'Haridon, E. Corre, B. Kloareg, P. Potin, *Int. J. Syst. Evol. Microbiol.* **2001**, *51*, 985–997.

Author Contributions

M.G., T.S. and U.T.B. initiated the study and directed the project. F.T. conducted the growth studies and created the knock-out strains. S.B. and M.B. expressed and purified the enzymes and performed the biocatalysis. D.B. performed the computational analysis. K.M. and M.L. performed metabolite analysis, S.B. prepared the main manuscript, which was revised by F.T., T.D, M.G, T.S. and U.T.B. and approved by all authors.

Article VI

1 **A unique alcohol dehydrogenase involved in algal sugar utilization by marine**
2 **bacteria**

3

4 Stefan Brott¹, Ki Hyun Nam², François Thomas³, Theresa Dutschei¹, Lukas Reisky¹, Maike
5 Behrens¹, Hanna C. Grimm¹, Gurvan Michel³, Thomas Schweder⁴, and Uwe T. Bornscheuer^{1*}

6

7 ¹Department of Biotechnology & Enzyme Catalysis, Institute of Biochemistry, University of
8 Greifswald, Greifswald 17487, Germany

9 ²Department of Life Science, Pohang University of Science and Technology, Pohang 37673,
10 South Korea

11 ³Laboratory of Integrative Biology of Marine Models (LBI2M), Station Biologique de Roscoff
12 (SBR), Sorbonne Université, CNRS 29688 Roscoff, Bretagne, France

13 ⁴Department of Pharmaceutical Biotechnology, Institute of Pharmacy, University of Greifswald,
14 Greifswald 17487, Germany

15 *Corresponding author: E-mail: uwe.bornscheuer@uni-greifswald.de

16

17 ORCID-IDs:

18 Stefan Brott: 0000-0003-2346-7706

19 Ki Hyun Nam: 0000-0003-3268-354X

20 François Thomas: 0000-0003-1896-0774

21 Theresa Dutschei: 0000-0003-3053-9536

22 Lukas Reisky: 0000-0001-8957-4083

23 Maike Behrens: 0000-0003-3757-588X

24 Hanna C. Grimm: 0000-0001-6352-0212

25 Gurvan Michel: 0000-0002-3009-6205

26 Thomas Schweder: 0000-0002-7213-3596

27 Uwe T. Bornscheuer: 0000-0003-0685-2696

28

29 **Abstract**

30 Marine algae produce complex polysaccharides, which can be degraded by marine
31 heterotrophic bacteria utilizing carbohydrate-active enzymes. The red algal polysaccharide
32 porphyran contains the methoxy sugar 6-O-methyl-D-galactose (G6Me). In the degradation of
33 porphyran, an oxidative demethylation of this monosaccharide towards D-galactose and
34 formaldehyde occurs, which is catalyzed by a cytochrome P450 monooxygenase and its redox
35 partners. In direct proximity to the genes encoding for the key enzymes of this oxidative
36 demethylation, a gene encoding for a zinc-dependent alcohol dehydrogenase (ADH) was
37 identified, which seems to be conserved in porphyran utilizing marine *Flavobacteriia*.
38 Considering the fact that dehydrogenases could play an auxiliary role in carbohydrate
39 degradation we aimed to elucidate the physiological role of this marine ADH. Although, our
40 results reveal that the ADH is not involved in formaldehyde detoxification, a knockout of the
41 ADH gene causes a dramatic growth defect of *Zobellia galactanivorans* with G6Me as
42 substrate. This indicates that the ADH is required for G6Me utilization. Complete biochemical
43 characterizations of the ADHs from *F. agariphila* KMM 3901^T (FoADH) and *Z. galactanivorans*
44 Dsij^T (ZoADH) were performed and the substrate screening revealed that these enzymes
45 preferentially convert aromatic aldehydes. Additionally, we elucidated the crystal structures of
46 FoADH and ZoADH in complex with NAD⁺ and show that the strict substrate specificity of this
47 new auxiliary enzyme is based on a narrow active site.

48 **Key points**

- 49 ▪ Knockout of the ADH encoding gene revealed its role in 6-O-methyl-D-galactose utilization,
50 suggesting a new auxiliary activity in marine carbohydrate degradation
- 51 ▪ Complete enzyme characterization indicated no function in a subsequent reaction of the
52 oxidative demethylation such as formaldehyde detoxification
- 53 ▪ These marine ADHs preferentially convert aromatic compounds and their strict substrate
54 specificity is based on a narrow active site

55

56 **Keywords**

57 Alcohol dehydrogenase · Porphyran · CAZyme · Bacteroidetes · *Zobellia galactanivorans*
58 · Auxiliary activity

59 Introduction

60 Marine algae represent one of the most crucial primary producers within the marine carbon
61 cycle and contribute to approximately 50% of the total global primary production (Field 1998).
62 For instance, macroalgae sequester approximately 173 Tg of carbon dioxide per year (Krause-
63 Jensen and Duarte 2016) and accumulate the excess carbon in form of carbohydrates, which
64 they utilize as cell wall constituents or energy storage (Arnosti et al. 2021). Degradation of
65 these marine polysaccharides can be extremely complicated due their complexity and the
66 occurrence of side chain modifications like sulfatations, methylations or acetylations (Bäumgen
67 et al. 2021a). It was shown that complex enzymatic cascades are required for the breakdown
68 of a single algal polysaccharide (Reisky et al. 2019; Sichert et al. 2020). Members of the
69 bacterial phylum *Bacteroidetes* are considered specialists in the pivotal degradation of marine
70 polysaccharides (Thomas et al. 2011a) and are observed as first responders after micro- and
71 macroalgal blooms (Teeling et al. 2012; Brunet et al. 2021). They contain specific gene clusters
72 referred to as polysaccharide utilization loci (PULs) (Grondin et al. 2017), which encode for
73 carbohydrate-active enzymes (CAZymes) that catalyze the breakdown of the carbohydrates
74 (Lapébie et al. 2019) as well as proteins essential for the binding and uptake of smaller sugar
75 molecules (Bauer et al. 2006; Martens et al. 2009). Characterizing individual CAZymes helps
76 elucidating complete degradation pathways of marine carbohydrates and provides a deeper
77 understanding of the global carbon cycle. Which has been successfully performed for instance
78 for ulvan from green algae (Reisky et al. 2019; Bäumgen et al. 2021b), fucoidan from brown
79 algae (Sichert et al. 2020) and carrageenan from red algae (Ficko-Blean et al. 2017).

80 Recently, we have demonstrated that in the degradation process of the red algal galactan
81 porphyran (Fig. 1a) by marine bacteria, an oxidative demethylation of the methoxy sugar
82 6-O-methyl-D-galactose (G6Me) occurs (Reisky et al. 2018). This reaction which is catalyzed
83 by a cytochrome P450 monooxygenase (CYP) and its respective redox partners consisting of
84 ferredoxin reductase and ferredoxin leads to the formation of equimolar amounts of
85 D-galactose and formaldehyde (Fig. 1b) (Reisky et al. 2018). It was hypothesized that this
86 reaction is crucial in terms of G6Me utilization, as it removes the highly stable methyl ether,
87 consequently generating an easily metabolizable compound (Reisky et al. 2018). The crystal
88 structure of the CYP from *Zobellia galactanivorans* Dsij^T provided additional information on the
89 binding of G6Me as well as other mechanistic insights (Robb et al. 2018). In addition to the key
90 enzymes for the oxidative demethylation of G6Me, glycoside hydrolases (GH2 and GH16), an
91 esterase and a putative zinc-dependent alcohol dehydrogenase (ADH) were also observed in
92 the genomic context of the marine *Flavobacteriia Formosa agariphila* KMM 3901^T (Fig. 1c)
93 (Reisky et al. 2018). A similar genomic context was also found in *Zobellia galactanivorans* Dsij^T
94 (Fig. 1d).

95 Considering the fact that dehydrogenases play only a minor auxiliary role in the carbohydrate
96 degradation and are poorly represented in the Carbohydrate-Active enZymes (CAZy)
97 database, with some exceptions in the AA3, AA6 and AA12 families (Takeda et al. 2015;
98 Kracher and Ludwig 2016; Sützl et al. 2018), it remains unclear which biological function this
99 ADH provides for the organism. ADHs belong to the enzyme class of oxidoreductases and
100 catalyze the reversible oxidation of an alcohol to the corresponding aldehyde or ketone
101 employing the nicotinamide adenine dinucleotide (NAD⁺) or nicotinamide adenine dinucleotide
102 phosphate (NADP⁺) cofactor. Depending on the size of the substrate-binding domain, it is
103 possible for ADHs to possess a broad substrate scope; while some exhibit only activities for
104 small aliphatic alcohols, others can convert sterically challenging cyclic components (Persson
105 et al. 2008; Sirota et al. 2021). A major family of ADHs includes the group of zinc-dependent
106 ADHs, which exhibit a typical Rossmann fold (Rao and Rossmann 1973) and contain a
107 catalytic zinc in the active site as well as an additional non-catalytic zinc supporting the stability
108 of an external loop structure (Hambidge et al. 2000). Various biological functions are observed
109 within this family (Sirota et al. 2021) including polyol dehydrogenases catalyzing the
110 conversion between sugar and sugar alcohol (Lu et al. 2019), cinnamyl alcohol
111 dehydrogenases (Larroy et al. 2002; Pick et al. 2013) and glutathione-dependent
112 formaldehyde dehydrogenases (Gutheil et al. 1992; Sanghani et al. 2000; Achkor et al. 2003)
113 which play an important part in the detoxification of formaldehyde (Vorholt 2002). Additionally,
114 ADHs provide numerous advantageous properties for organic synthesis including high
115 enantioselectivity and applicability under mild reaction conditions (Koesoema et al. 2020).
116 Consequently, they are now employed in numerous biotechnological applications such as the
117 preparation of chiral alcohols (Zhang et al. 2015), rare sugars (Lu et al. 2019), fine chemicals
118 as well as the synthesis of building blocks for various essential pharmaceuticals (Hall and
119 Bommarius 2011; Zheng et al. 2017). Discovering and characterizing additional ADHs with
120 unique biochemical properties, is thus also desirable for potential industrial applications.

121 In this study, we aimed to elucidate the putative function of the ADH, which is consistently
122 located in close proximity to genes that are essential for the oxidative demethylation of G6Me
123 of polysaccharide utilizing marine *Flavobacteriia*. We provide a detailed biochemical
124 characterization as well as the crystal structures for the ADHs from *Formosa agariphila*
125 KMM 3901^T (FoADH) and *Zobellia galactanivorans* Dsij^T (ZoADH). We propose the putative
126 biological functions of these ADHs and demonstrate their importance for the utilization of G6Me
127 via growth studies with a *Z. galactanivorans* knockout strain.

128 **Materials and Methods**

129 **Materials, strains and plasmids**

130 All chemicals and reagents used, unless otherwise specified, were purchased from Sigma-
131 Aldrich (St. Louis, MO, USA), Thermo Fisher Scientific (Waltham, MA, USA), Th. Geyer (Berlin,
132 Germany), ABCR GmbH (Karlsruhe, Germany), Honeywell Fluka™ (Morristown, NJ, USA),
133 Carl Roth GmbH (Karlsruhe, Germany), chemPUR GmbH (Karlsruhe, Germany), TCI
134 Deutschland GmbH (Eschborn, Germany) and Cayman Chemical Company (Ann Arbor, MI,
135 USA). Porphyrin and G6Me were obtained from Biosynth Carbosynth (Staad, Switzerland).
136 Primers were obtained from Invitrogen (Waltham, MA, USA). Phage-resistant *Escherichia coli*
137 (*E. coli*) BL21 (genotype: *fhuA2* [*lon*] *ompT gal* (λ *DE3*) [*dcm*] Δ *hsdS* [λ *DE3* = λ *sBamHI*
138 Δ *EcoRI-B* int::(*lacI::PlacUV5::T7 gene1*) *i21* Δ *nin5*) was obtained from New England Biolabs
139 (Ipswich, MA, USA). The conjugative strain *E. coli* S17-1 λ *pir* (genotype λ *pir hsdR pro thi*;
140 chromosomal integrated RP4-2 Tc::Mu Km::Tn7) (de Lorenzo and Timmis 1994) was grown
141 from in-house glycerol stocks. A construct for the expression of the FoADH (GenBank
142 accession number: OP548117) from *F. agariphila* KMM 3901^T was prepared using the
143 FastCloning strategy (Li et al. 2011) with genomic DNA as template for the amplification of the
144 insert. *F. agariphila* KMM 3901^T (collection number DSM-15362) was obtained from the DSMZ
145 (Braunschweig, Germany). The pET28a vector was amplified with the 5-GCG GCC GCA CTC
146 GAG CA-3' and 5-CAT ATG GCT GCC GCG C-3' oligonucleotides while the insert was
147 amplified with the 5'-CAC AGC AGC GGC CTG GTG CCG CGC GGC AGC CAT ATG TCC
148 ATA ATT TCA AAA TGC GCT ATT G-3' and 5'-CAG TGG TGG TGG TGG TGG TGC TCG
149 AGT GCG GCC GCT TAA AAA ATA ATT ACA CCC TTT GCA TTC-3' oligonucleotides. A
150 synthetic gene, codon optimized for expression in *E. coli*, encoding the ZoADH (GenBank
151 accession number: OP548118) from *Z. galactanivorans* Dsij^T, was synthesized and cloned into
152 a pET28a vector by BioCat GmbH (Heidelberg, Germany). The constructs encoded the
153 recombinant protein as fusion to a N-terminal Strep-tag for affinity purification.

154 Computational analysis for FoADH and ZoADH

155 Sequences of FoADH (Uniprot ID: T2KM87) and ZoADH (Uniprot ID: G0L712) were blasted
156 against the MarDB and MarRef database using the Marine Metagenomic Portal (Klemetsen et
157 al. 2018; Priyam et al. 2019) with the -e value of 1e⁻⁵ and maximal target sequences of 1000.
158 The automated fasta hit table of both blasts were fused and used for the generation of a
159 sequence similarity network (Zallot et al. 2019). An alignment score of 150 was chosen for the
160 refinement and generation of a genome neighborhood analysis of ten genes down and
161 upstream of the ADHs genes (Zallot et al. 2019). Resulting diagrams were visualized via
162 Cytoscape (Paul Shannon et al. 2003) and genome neighborhood diagrams were generated

163 from the online server. Only shared sequences of the MarDB/MarRef database with the
164 UniProtKB databases could be incorporated in the genome neighborhood analysis.

165 ADH knockout in *Z. galactanivorans* and growth studies

166 The deletion mutant of the ADH gene *zgal_4674* in *Z. galactanivorans* Dsij^T (collection number
167 DSM-12802) was constructed using a *sacB* system (Zhu et al. 2017) as previously described
168 for the deletion variant of the CYP gene (Brott et al. 2022). Briefly, to delete *zgal_4674*, a
169 2,448 bp fragment including the last 43 bp of *zgal_4674* and 2,405 bp of downstream sequence
170 was amplified using primers OFT0041 and OFT0043 on genomic DNA from
171 *Z. galactanivorans* Dsij^T. The fragment was digested with BamHI and XbaI and ligated into
172 pYT313 that had been digested with the same enzymes, to generate pFT12. A 2,077 bp
173 fragment including the first 29 bp of *zgal_4674* and 2,048 bp of upstream sequence was
174 amplified using primers OFT0040 and OFT0042. The fragment was cloned into XbaI and Sall
175 sites of pFT12 to generate the *zgal_4674* deletion construct pFT13. Conjugative transfer of
176 pFT13 from *E. coli* S17-1 into the wild-type *Z. galactanivorans* Dsij^T and second recombination
177 steps were carried out as described previously (Zhu et al. 2017). Deletions were confirmed by
178 PCR and sequencing on isolated colonies using primer pairs OFT0044- OFT0045 to identify
179 the *zgal_4674* deletion mutant (mZG_0080). Primers employed are displayed in Table S1 in
180 the Supplementary Information (SI). For growth studies, precultures of three
181 *Z. galactanivorans* strains (wild type, knockout ADH and knockout CYP) were prepared in
182 Zobell 2216E medium (Zobell 1941). The 3-day precultures were then rinsed twice with sterile
183 saline solution. Marine minimal medium (Thomas et al. 2011a) amended with D-galactose or
184 G6Me (4 g L⁻¹) was then inoculated so that an initial optical density (OD₆₀₀) of 0.05 was
185 achieved. Appropriate cultures were incubated for 3 days at room temperature.

186 Enzyme production and purification

187 Chemically competent *E. coli* BL21 (DE3) cells were transformed with the plasmids harboring
188 FoADH or ZoADH and were spread on lysogeny broth (LB) agar plates containing 50 µg mL⁻¹
189 kanamycin. The agar plates were incubated overnight at 37 °C. One colony was picked and
190 used to inoculate 5 mL LB medium which contained 50 µg mL⁻¹ kanamycin and was then
191 incubated at 37 °C and 180 rpm overnight. For overexpression the cultivation was performed
192 in terrific broth (TB) medium containing 50 µg mL⁻¹ kanamycin. The TB medium was inoculated
193 with the overnight culture so that a starting OD₆₀₀ of 0.05 was obtained. Cells were then
194 incubated at 37 °C and 180 rpm until an OD₆₀₀ of 0.8 was reached. Expression of target
195 enzymes was induced by the addition of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG).
196 The cultivation was performed at 25 °C and 180 rpm overnight. Cells were harvested by
197 centrifugation at 10,000 x *g* and 4 °C for 1 h, washed with 50 mM sodium phosphate buffer

198 (NaPi) pH 7.5, and subsequently stored at -20 °C until cell disruption. The purification
199 procedures of FoADH and ZoADH for crystallization and enzyme assays are identical. Cells
200 were resuspended in 50 mM Tris-HCl buffer pH 8.0 containing 200 mM NaCl. Following cell
201 lysis by ultra-sonication (2 × 3 min, 50% power, 50% cycle), cell debris was removed by
202 centrifugation at 10,000 x g, at 4 °C for 20 min. The clarified supernatant was loaded on a
203 gravity flow column containing Strep-Tactin XT Sepharose® 50% suspension (IBA-
204 Lifesciences GmbH, Göttingen, Germany) as column material. The column was washed with
205 100 mM Tris-HCl buffer pH 8.0 containing 150 mM NaCl in order to remove unbound and
206 undesirable proteins. The target enzymes were then eluted with same buffer containing
207 additionally 50 mM biotin. Elution fractions were pooled and concentrated using a Vivaspin 6
208 centrifugal concentrator with a 10 kDa molecular weight cut-off (Sartorius AG, Göttingen,
209 Germany). Size exclusion chromatography was subsequently performed via the Äkta™ pure
210 chromatography system (Cytiva Europe GmbH, Germany). The concentrated enzyme solution
211 was applied to a HiPrep™ 16/60 Sephacryl® S-200 HR column (Cytiva Europe GmbH, Freiburg,
212 Germany) that was previously equilibrated with 10 mM Tris-HCl buffer pH 8.0 containing
213 200 mM NaCl. Elution fractions were collected and the purity was verified by sodium dodecyl
214 sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Pure fractions were combined and
215 concentrated like mentioned above. The enzyme solution was stored at 4 °C for crystallization.
216 For application in enzyme assays, a PD-10 desalting column (Cytiva Europe GmbH, Freiburg,
217 Germany) was employed to desalt the protein sample and exchange the buffer.

218 SDS-PAGE and determination of protein content

219 SDS-PAGE was performed to verify the purity of the target enzymes. 20 µL protein sample
220 was mixed with 5 µL of a 5-fold stock of SDS sample buffer (100 mM Tris-HCl buffer pH 6.8
221 containing 4% (w/v) SDS, 20% (v/v) glycerol, 2% (v/v) β-mercaptoethanol, 25 mM
222 ethylenediaminetetraacetic acid (EDTA) and 0.04% (w/v) bromophenol blue) and denatured at
223 99 °C for 15 min. For the SDS-PAGE a 12.5% acrylamide gel (separating gel) and a 4.0%
224 loading gel were used. Electrophoresis was carried out at 200 V. Proteins were stained with
225 Coomassie Blue (PhastGel® Blue R). As reference the Pierce™ Unstained protein molecular
226 weight marker (Thermo Fisher Scientific, Waltham, MA, USA) was used. Protein
227 concentrations were determined using the Pierce™ BCA Protein Assay Kit (Thermo Fisher
228 Scientific, Waltham, MA, USA) with bovine serum albumin as protein standard.

229 Crystallization

230 Purified FoADH (25 mg mL⁻¹) and ZoADH (25 mg mL⁻¹) were incubated with 20 mM NAD⁺
231 overnight. Initial crystallization screen was performed using sitting drop vapor-diffusion method
232 at 22 °C. The droplets contained 0.2 µL of protein and 0.2 µL of reservoir solution. Microcrystals

233 of FoADH were obtained from reservoir solution containing 0.1 M Tris-HCl pH 7.5, 0.2 M KCl
234 and 22% (w/v) polyethylene glycol 3350. Microcrystals of ZoADH were obtained from reservoir
235 solution containing 0.1 M Tris-HCl pH 7.5, 0.2 M KCl and 20% (w/v) polyethylene glycol 3350.
236 Further crystal optimization was performed by scale-up of the droplets containing 2 μ L of
237 protein and 2 μ L of reservoir solution, using the hanging drop vapor-diffusion method at 22 $^{\circ}$ C.
238 Suitable FoADH and ZoADH crystals for X-ray diffraction were obtained from 0.1 M Tris-HCl,
239 pH 7.5, 0.2 M KCl and 20-22% (w/v) polyethylene glycol 3350 within one day.

240 Data collection

241 X-ray diffraction data were collected at beamline 11C at Pohang Light Source II (PLS-II,
242 Pohang, South Korea) with a Pilatus 6M detector (Dectris, Swiss). The FoADH crystals were
243 equilibrated in a cryoprotectant buffer containing reservoir buffer plus 20% (v/v) ethylene glycol.
244 ZoADH crystals were equilibrated in a cryoprotectant buffer containing reservoir buffer plus
245 20% (v/v) glycerol. The crystal was mounted on the goniometer and cooled under a nitrogen
246 gas stream at 100 K. The diffraction data were indexed, integrated, and scaled using HKL2000
247 program (Otwinowski and Minor 1997). A data collection statistic is given in Table S2.

248 Structure determination

249 The electron density maps of FoADH and ZoADH were obtained via the molecular replacement
250 method using the MOLREP program (Vagin and Teplyakov 2010). The crystal structure of an
251 ADH from *Artemisia annua* (PDB code: 6LJH, unpublished) was used as search model for both
252 FoADH and ZoADH. Model building and refinement were performed with the COOT program
253 (Emsley and Cowtan 2004) and phenix.refinement in PHENIX (Liebschner et al. 2019),
254 respectively. The geometry of final models was evaluated with MolProbity (Williams et al. 2018).
255 Structural figures were generated with PyMOL (www.pymol.org). Structure-based sequence
256 alignments were generated using Clustal-Omega (Sievers et al. 2011) and ESPript (Gouet et
257 al. 1999). Tetrameric interfaces of ADHs were analyzed by PDBePISA (Krissinel and Henrick
258 2007). The interaction between ADHs and ligands were analyzed using PLIP (Salentin et al.
259 2015). The structure factor and coordinates are deposited in the Protein Data Bank under PDB
260 codes 8H2A (FoADH-NAD) and 8H2B (ZoADH-NAD).

261 Enzyme activity determination and substrate screening

262 For determining the enzyme activity of the ADHs, the absorbance maximum of NADH at
263 340 nm was utilized. The absorbance at 340 nm was measured every minute over a 10 min
264 period using a microplate spectrophotometer (BioTek Synergy H1, Agilent Technologies,
265 Santa Clara, CA, USA), and the slope over time was used to determine activities or relative
266 activities. One unit of activity is defined as oxidation or formation of 1 μ mol of NADH per minute.

267 For calculation of activity, the molar absorption coefficient of NADH was determined via a
268 standard curve that covered the range of 0 to 0.5 mM. For the initial substrate screening,
269 several alcohols/aldehydes/ketones were employed at a final concentration of 10 mM. For
270 increased substrate solubility, these reactions contained 3.5% (v/v) dimethyl sulfoxide (DMSO).
271 The total volume for all reactions was 0.2 mL. The oxidation and reduction were both
272 conducted at an incubation temperature of 70 °C. Reduction of aldehydes was performed in
273 the presence of a 50 mM succinate buffer pH 6.5, while oxidation reactions were assayed in
274 the presence of a 50 mM NaPi buffer pH 8.5. The final enzyme concentrations used to provide
275 a linear absorbance increase or decrease ranged from 20-100 $\mu\text{g mL}^{-1}$ for the oxidation
276 reactions and from 0.25-2.5 $\mu\text{g mL}^{-1}$ for the reduction reactions. The reaction was initialized by
277 the addition of 0.5 mM NAD^+ or NADH. For the measurement with sugar substrates, a reduced
278 reaction temperature of 40 °C and an increased measuring time of 30 min was chosen. Various
279 sugars were used at a final substrate concentration of 30 mM. A concentration of 0.2% (w/v)
280 was used for porphyran. Oxidation and reduction reactions were performed in the identical
281 buffers as used for substrate screening, the final enzyme concentration was 0.1 mg mL^{-1} . The
282 reaction was initialized by the addition of 0.5 mM NAD^+ or NADH. In order to test for thiol-
283 dependent formaldehyde detoxification, different thiols were evaluated as potential cofactors.
284 For this reaction, the thiol cofactor and formaldehyde were used in a 1:1 ratio at a final
285 concentration of 0.5 mM. The measurement was performed in the 50 mM NaPi buffer pH 8.5
286 at 70 °C with a final enzyme concentration of 0.2 mg mL^{-1} . The reaction was started by the
287 addition of 0.5 mM NAD^+ . The ADH catalyzed disproportionation of formaldehyde into
288 methanol and formate was monitored by a pH change utilizing the phenol red assay (Martínez-
289 Martínez et al. 2018). This measurement was performed in a microtiter plate and the reaction
290 volume was 0.2 mL. 5 mM formaldehyde was used as substrate, 0.5 mM NAD^+ as cosubstrate
291 and 0.1 mg mL^{-1} as final enzyme concentration. The pH indicator phenol red was used at a
292 final concentration of 91 μM . The reaction was performed in a 5 mM HEPES buffer pH 8.5 at
293 40 °C. Absorbance at 560 nm was measured every minute for 20 min.

294 Influence of pH and buffer components

295 To determine the pH optimum of the enzymes, the oxidation and reduction reactions were both
296 investigated in the presence of varying pH values. All buffers had a concentration of 50 mM. A
297 citrate buffer was used in the pH range of 5 to 6, a NaPi buffer in the range of 6 to 8.5, a CHES
298 buffer in the range of 8.5 to 10 and a CAPS buffer in the range of 10 to 12.5. The assay
299 conditions for the oxidation reaction were as follows: the reaction volume was 200 μL , 10 mM
300 benzyl alcohol and 0.5 mM NAD^+ was used as substrate. The reaction was started by the
301 addition of 0.1 mg mL^{-1} ADH. For the reduction reaction, instead of benzyl alcohol and NAD^+ ,
302 10 mM benzaldehyde and 0.5 mM NADH were used. Since benzaldehyde was less soluble in

303 the buffer than benzyl alcohol, both reactions contained 3.5% (v/v) DMSO, in order to achieve
304 better comparability. The reaction was carried out at 25 °C in the respective buffers. To
305 examine the influence of buffer components on enzyme activity, different buffers with a
306 concentration of 50 mM were used. The buffers had a pH of 6.5 for the reduction reaction,
307 whereas it was 8.5 for the oxidation reaction. The reaction was carried out under the same
308 conditions as those for the pH optimum. Relative activities were determined as described
309 above.

310 Influence of temperature and thermostability

311 The temperature optimum was determined by conducting the oxidation reaction at different
312 temperatures in the range between 20 and 90 °C. For this, the reaction mixture without enzyme
313 was preheated to the desired temperature in a reaction tube by using a heating block
314 (Eppendorf ThermoMixer®C, Eppendorf SE, Hamburg, Germany) for at least 45 min. The
315 reaction mixture had a volume of 200 µL. 30 mM benzyl alcohol and 0.5 mM NAD⁺ were
316 employed as substrates, and the reaction was carried out at different temperatures ranging
317 from 20 to 90 °C in a 50 mM NaPi buffer pH 7.5. The reaction was initiated by the addition of
318 enzyme with a final concentration of 0.1 mg mL⁻¹. For the thermostability determination, the
319 purified ADH (1 mg mL⁻¹) was incubated in 50 mM NaPi buffer pH 7.5 for 1 or 4 h in a gradient
320 thermal cycler (FlexCycler², Analytik Jena, Jena, Germany) at various temperatures ranging
321 from 20 to 80 °C. Residual activity was then determined as described above and compared
322 with a control that was incubated on ice. The assay conditions were as follows: the reaction
323 volume was 200 µL, the final enzyme concentration was 0.1 mg mL⁻¹, 10 mM benzyl alcohol
324 was used as substrate, the reaction was performed at 40 °C in 50 mM NaPi buffer pH 7.5. The
325 reaction was initiated by the addition of 0.5 mM NAD⁺.

326 Influence of sodium chloride

327 Determination of NaCl influence on enzyme activity was performed by carrying out the
328 oxidation reaction in the presence of different NaCl concentrations varying from 0 to 800 mM.
329 The relative activities were determined as described above and were compared with a control
330 where no additional NaCl was present. Assay conditions were as follows: the reaction volume
331 was 200 µL, 10 mM benzyl alcohol was used as substrate, the final enzyme concentration was
332 0.1 mg mL⁻¹ and the NaCl concentration was between 0 and 800 mM. The reaction was carried
333 out at 25 °C in a 50 mM NaPi buffer pH 8.5 or in a 50 mM Tricine buffer pH 8.5 and started by
334 the addition of 0.5 mM NAD⁺.

335 Influence of metal ions and other small molecules

336 For the determination of the influence of various metal ions on enzyme activity, the ADHs with
337 a concentration of 1 mg mL⁻¹ were incubated with either 1 or 10 mM metal ion at RT for 1 h
338 before activity measurement. A sample without additional metal ion served as a control. For
339 the activity measurement, the standard assay was used with the following conditions: the
340 reaction mixture had a total volume of 200 μL, 10 mM benzyl alcohol was used as substrate,
341 a final enzyme concentration of 0.1 mg mL⁻¹ was employed and the reaction was performed in
342 50 mM HEPES buffer pH 8.5 at 25 °C. The reaction was initiated by the addition of 0.5 mM
343 NAD⁺. In order to determine the effect of EDTA, dithiothreitol (DTT) and 2-mercaptoethanol
344 (2-ME) on enzyme activity, the ADHs were incubated at a protein concentration of 1 mg mL⁻¹
345 with these components at concentrations of 1, 10 or 25 mM for 1h at RT before activity
346 determination. The untreated enzyme served as a control. The activity measurement was
347 performed as described for the influence of metal ions.

348 Influence of solvents and formaldehyde

349 To evaluate the influence of selected water-miscible solvents on the activity of both ADHs, the
350 oxidation reaction was conducted in the presence of 5, 10, and 20% (v/v) solvent and
351 compared with a control containing no additional solvent. The relative activity was determined
352 as described above. The total reaction volume was 0.2 mL and 0.1 mg mL⁻¹ of enzyme was
353 used as final enzyme concentration. The reactions were performed in 50 mM NaPi buffer at
354 25 °C. 10 mM benzyl alcohol was employed as substrate and the reactions were started by
355 adding 0.5 mM NAD⁺. The enzymes were incubated at a concentration of 1 mg mL⁻¹ with
356 different concentrations of formaldehyde varying from 0 to 50 mM for 1 h at RT prior to activity
357 measurement to evaluate the effect of formaldehyde on enzyme activity. Relative activity was
358 determined as described above. For the activity measurement, the same conditions were used
359 as for the influence of solvent.

360 **Results**

361 Distribution and gene neighborhood analysis

362 In order to obtain an overview regarding the distribution and function of these ADHs in marine
363 bacteria, we queried the MarDB and MarRef databases for ADHs with similar sequences to
364 FoADH and ZoADH and constructed a sequence similarity network based on an alignment
365 score of 150 and a sequence identity of 63.14%. This analysis revealed six main clusters,
366 which we define here as clusters containing at least 34 sequences, with FoADH and ZoADH
367 included in main cluster 2 (Fig. S1). This main cluster primarily contained sequences that were
368 annotated as zinc-dependent ADHs, histidine kinases, ADH GroES-like domains and some
369 glutathione dependent formaldehyde dehydrogenases/ADHs. However, glutathione-
370 dependent and mycothiol-dependent formaldehyde dehydrogenases were identified

371 predominantly in clusters 1 and 4, respectively. Based on main cluster 2, we performed a
372 genome neighborhood analysis to obtain a general sense of which genes are located in close
373 proximity to the ADH gene. Similar genomic arrangements consisting of CYP, redox partners,
374 an esterase and the ADH can be identified in several marine bacteria that are capable of
375 degrading marine polysaccharides (Fig. S2), including members of the genera *Polaribacter*,
376 *Maribacter* and *Arenibacter*. Minor differences in gene arrangement can be observed among
377 some organisms such as *F. agariphila* or *Algibacter lectus*, where genes encoding for
378 CAZymes (GH2 and GH16) are located between the ADH and the esterase gene. Additionally,
379 some genes encoding for sulfatases and SusC/SusD homologs, which are responsible for the
380 binding and transport of sugar molecules (Martens et al. 2009), are located up- and
381 downstream of the ADH gene. Considering that the ADH gene consistently appears in the
382 proximity of the genes, which encode for CAZymes and key enzymes for the oxidative
383 demethylation of G6Me, it is conceivable that the ADH possesses a specific function in
384 carbohydrate utilization or a subsequent reaction.

385 Knockout of the ADH encoding gene in *Z. galactanivorans* and growth studies

386 As an attempt to elucidate the biological relevance of the ADHs for the organisms, a knock-out
387 of the gene, which encodes for the ADH in *Z. galactanivorans* was performed followed by
388 growth experiments. The controls employed for these growth studies were the wild type (WT)
389 and an additional knock-out strain of *Z. galactanivorans* in which the CYP gene was deleted.
390 When G6Me was employed as the sole carbon source, impaired growth was observed for the
391 ADH and CYP knock-out strains, while the WT exhibited normal growth (Fig. 2). In contrast,
392 regular growth was observable for all three strains in a control, which contained D-galactose
393 as sole carbon source. Consequently, the ADH possessed an impact on the G6Me utilization
394 of *Z. galactanivorans*.

395 Functional overexpression and purification of the ADH

396 Since we could demonstrate a biological significance of the ADH for the utilization of G6Me by
397 the gene knockout in *Z. galactanivorans*, our next aim was to identify the enzyme function. We
398 therefore cloned the gene encoding for the ADH from *F. agariphila* into a pET28a vector. For
399 the ADH from *Z. galactanivorans* a synthetic gene was ordered in the pET28a vector. Both
400 enzymes were successfully overexpressed and purified (Fig. S3), which established the basis
401 to elucidate putative biological functions of this ADHs by performing biochemical and structural
402 biological characterizations.

403 Substrate spectrum of the ADHs

404 In order to obtain a preliminary understanding over the substrate spectrum of these ADHs,
405 their ability for the alcohol oxidation as well as the reduction of various aldehydes and ketones
406 were examined. Both enzymes converted predominantly aromatic substrates (Tables 1 and 2).
407 The highest specific activity of 64.1 U mg⁻¹ for FoADH and 54.9 U mg⁻¹ for ZoADH was
408 observed for the reduction of pyridine-3-carbaldehyde. In addition to compounds containing a
409 benzene ring, substrates harboring a furan or thiophene ring, such as furfural and thiophene-
410 3-carbaldehyde, were also preferentially converted. Positions of additional substituents at the
411 benzene ring influenced the activity. A difference in the specific activities for the constitutional
412 isomers of terephthalaldehyde and tolualdehyde were observed for both enzymes. In particular,
413 substrates that possessed an additional substituent in *ortho*-position were converted
414 significantly less efficiently. In addition, the length of the aldehyde substituent at the benzene
415 ring also affected the activity. For instance, hydrocinnamaldehyde was converted by both
416 enzymes, whereas for phenylacetaldehyde no activity was observable. In contrast to
417 benzaldehyde, the structurally similar acetophenone could not be oxidized. Thus, both ADHs
418 were unable to convert ketones to secondary alcohols. In comparison to the reduction reaction,
419 significant reduced specific activities were noticed for the oxidation reactions (Table 2). The
420 highest specific activity of 490 mU mg⁻¹ for FoADH and 290 mU mg⁻¹ for ZoADH was observed
421 for 2,5-bis(hydroxymethyl)furan. Both ADHs lacked any activity for smaller aliphatic alcohols
422 such as methanol and ethanol. Since the ADHs exhibited predominantly activities for
423 substrates containing a ring structure, several sugars were also considered as possible
424 substrates. However, no activity was observed for the oxidation or reduction of galactose,
425 G6Me and additional monosaccharides and disaccharides (Table S3). Additionally, the marine
426 carbohydrate porphyran was also evaluated as a potential substrate, however, no activity was
427 detected either. Both ADHs utilize NAD⁺ as cofactor, in the presence of NADP⁺ no activity for
428 the oxidation reaction was observed.

429 Testing for formaldehyde detoxification activity

430 Since no activity was observed for neither galactose nor G6Me, we hypothesized that the
431 ADHs may participate in formaldehyde detoxification, considering that formaldehyde is formed
432 as a by-product in the oxidative demethylation reaction. Members of the zinc-dependent ADHs
433 may catalyze the glutathione-dependent formaldehyde detoxification, therefore various thiols
434 were considered as potential cofactors. Thiol-dependent detoxification of formaldehyde
435 proceeds via a spontaneous reaction between the sulfhydryl group of the thiol cofactor and the
436 carbon atom of formaldehyde, resulting in the formation of an alcohol (Fig. 3a) (Chen et al.
437 2016). Subsequently, this alcohol can be oxidized by the ADH to a thioester, which is then
438 converted by an esterase to formate and the starting thiol cofactor (Gonzalez et al. 2006).
439 Based on the results of our genome neighborhood analysis, where we have also demonstrated

440 that a gene encoding for an esterase is located in the vicinity of the ADH gene, it is quite
441 possible that a thiol-dependent detoxification of formaldehyde can proceed via both enzymes.
442 In addition to glutathione, mainly mycothiol (Misset-Smits et al. 1997; Newton and Fahey 2002)
443 and bacillithiol (Newton et al. 2009; Chandrangu et al. 2018) are well known cofactors in
444 formaldehyde detoxification (Fig. 3b). However, no activity was detected for these thiols.
445 Furthermore, common thiols abundant in nature such as cysteine, coenzyme A and
446 L-ergothioneine (Hand and Honek 2005) were also investigated as cofactors. Nevertheless, no
447 activity was observed for these substrates in combination with formaldehyde either.
448 Considering that the ADHs mainly exhibited activity for aromatic substrates, aromatic thiols
449 such as 2-mercaptoimidazole or 4-mercaptofenol were considered as possible substrates as
450 well. However, even with these compounds, no oxidation reaction was detected. Furthermore,
451 neither enzyme exhibited activity for the oxidation or reduction of formaldehyde in the presence
452 of only NAD^+ or NADH as cofactors. In addition, a disproportionation reaction of formaldehyde
453 into methanol and formate catalyzed by the ADH was also checked. However, no activity could
454 be detected. Consequently, the ADHs possessed no activities for the substrate nor for the
455 products of the oxidative demethylation of G6Me. To provide additional insights into these
456 ADHs, we performed further biochemical characterizations of both enzymes.

457 Influence of pH and buffer components on enzyme activity

458 In order to determine the optimal pH for the enzymatic reaction, several buffers were
459 investigated in the pH range from 5.5 to 12.5. A similar pH optimum was observed for both
460 enzymes (Fig. 4). The reduction reaction was most efficiently catalyzed at pH 6.5, while
461 oxidation was found to be most efficient at pH 8.5 (Fig.4 a-b and d-e). At pH 5 and at 12.5, no
462 activity was detected for either enzymes; precipitation was noticed at pH 5 while employing
463 higher protein concentrations. Since a considerable difference in activity was observed
464 between NaPi and CHES buffer at pH 8.5, other buffers were also evaluated at pH 6.5 (Fig. 4 c)
465 and 8.5 (Fig. 4f) to investigate the influence of buffer components on the activity. For the
466 oxidation reaction at pH 8.5, it was shown that by employing a Tris-HCl buffer, an
467 approximately 60 to 80% increased activity was obtained compared to the activity in the NaPi
468 buffer. In contrast, a significant activity decrease of 95% was observed for both enzymes in
469 the presence of a borate-NaOH buffer. For the reduction reaction at pH 6.5, a slight increase
470 in activity of ~8 to 16% could be detected using citrate and succinate buffer compared to the
471 NaPi buffer, with the highest activity found for the succinate buffer.

472 Influence of temperature and enzyme thermostability

473 In addition to the pH value, the temperature influence is essential for enzymatic activity. At the
474 same time, elevated temperatures promote substrate solubility and thus the application of

475 higher concentrations which also may shift the reaction equilibrium towards product formation
476 (Unsworth et al. 2007). Therefore, the impact of temperature in the range between 20 and
477 90 °C was investigated for both enzymes. The ADHs possessed a similar temperature profile,
478 where activity increased with rising temperature, reaching an optimum between 65 to 75 °C
479 (Fig. 5a). However, at higher temperatures the activity decreased rapidly, whereas at room
480 temperature only a relative activity of about 18% for FoADH and 10% for ZoADH was observed.
481 The measurement for the temperature optimum was performed for 10 min to ensure that any
482 influence of thermostability would not affect the results. Thermostability of enzymes is an
483 important parameter for biocatalysis, since many industrial processes operate at higher
484 temperatures for longer time periods, leading to increased product yields. The thermostability
485 of the ADHs was therefore evaluated next by incubating the enzymes for 1 or 4 h at various
486 temperatures ranging from 20 to 80 °C followed by determination of residual activity. After 1 h
487 incubation at 59 °C as well as lower temperatures, no decrease in activity was detected for
488 FoADH compared to a control incubated on ice (Fig. 5b). Residual activity only diminished at
489 higher incubation temperatures and a residual activity of roughly 20% was still observed for
490 80 °C. In contrast, after 4 h incubation, almost no residual activity was observed at this
491 temperature. Nevertheless, even after this extended incubation period, a high remaining
492 activity of approximately $\leq 85\%$ was detected for the temperature range of 20 to 59 °C. ZoADH
493 exhibited a similar behavior in thermostability as FoADH, however an initial activity decrease
494 of 20% was observed for the 1 h incubation already at 57 °C (Fig. 5c). A severe activity loss of
495 almost 95 to 100% was observed for ZoADH when incubated for 4 h at temperatures >73 °C.

496 Influence of sodium chloride

497 Enzymes originating from marine organisms may possess habitat-related characteristics such
498 as an increased salt tolerance (Trincone 2011). Considering that both enzymes originate from
499 marine bacteria, the influence of NaCl on the enzyme activity was tested. For this purpose, the
500 relative activities for the oxidation reaction were determined in the presence of different NaCl
501 concentrations ranging from 0 to 800 mM in the NaPi and Tricine buffer, respectively. Both
502 ADHs displayed a similar behavior in the presence of rising NaCl concentrations (Fig. S4). An
503 increase in relative activity of approximately 10% was observed in the range from 0 to 150 mM
504 NaCl for FoADH using the Tricine buffer. In contrast, only a minor increase in activity was
505 observed for the NaCl concentration of 100 mM in the NaPi buffer. A difference between the
506 NaCl influence depending on the selected buffer was also noticed for ZoADH, with a higher
507 effect in the Tricine buffer. For ZoADH, an increase in relative activity of 20% was also detected
508 in the range of 0 to 200 mM NaCl. At NaCl concentrations ≥ 400 mM, a diminished relative
509 activity was observed for both enzymes.

510 Influence of metal ions and other small molecules

511 Both enzymes are annotated as zinc-dependent ADHs, which contain a catalytic zinc ion in
512 the active site. An influence of various metal ions on the enzyme activity is thus possible and
513 was therefore investigated next. For this purpose, the enzymes were incubated with different
514 metal ions at concentrations of 1 or 10 mM for 1 h prior to activity measurement and the relative
515 activities were determined. A high dependence on metal ions was observed for both ADHs,
516 with nearly all ions assayed exhibiting a beneficial effect on enzyme activity (Table 3 and
517 Fig. S5). Particularly higher concentrations of Ni^{2+} , Co^{2+} and Mn^{2+} led to a 10 to 14-fold increase
518 in relative activity for both enzymes compared to the control which contained no additional
519 metal ion. In contrast, complete inhibition for both enzymes was only observed for Cu^{2+} , Zn^{2+}
520 as well as 10 mM Fe^{3+} . Additionally, we analyzed whether the chelating agent EDTA, which is
521 capable of complexing bivalent metal ions, affects the enzymatic activity. After 1 h incubation
522 in the presence of 25 mM EDTA, a reduction in the relative activity for both enzymes was found,
523 while complete inhibition was not observable (Table 3 and Fig. S6). The influence of DTT and
524 2-mercaptoethanol (2-ME) on activity was also investigated, since these compounds can affect
525 enzyme stability. DTT had a lesser impact on both enzymes than 2-ME. A major decline in
526 relative activity of over 70% was observed for both enzymes after 1 h incubation with 10 mM
527 2-ME (Table 3 and Fig. S6). When compared to ZoADH, the effect of the reducing agents was
528 more pronounced for the activity of FoADH.

529 Influence of solvents and formaldehyde

530 The influence of water-miscible solvents on the enzyme activity of both ADHs was also
531 investigated. Increasing the amount of solvent in the reaction led to a decrease in the relative
532 activity for all tested solvents (Fig. S7). Compared to the other solvents, methanol and DMSO
533 had the weakest negative effects on the enzyme activity, leading to a relative activity of still
534 50% in the presence of 10% (v/v) solvent. In addition, the presence of formaldehyde on the
535 enzyme activity was examined, since formaldehyde is released during the oxidative
536 demethylation of G6Me and the ADHs are most likely involved in this reaction. Therefore, the
537 ADHs were incubated with a variety of formaldehyde concentrations in the range between
538 0 and 50 mM for 1 h at RT and the relative activities was determined. In the presence of 0 to
539 1 mM formaldehyde, no reduction in activity was observed. An initial decrease in relative
540 activity of approximately 10-20% could be perceived in the presence of 2.5 mM formaldehyde
541 (Fig. S8). At higher formaldehyde concentrations, a more severe activity decrease was found,
542 while no activity was observed for both enzymes in the presence of 50 mM formaldehyde.

543 Overall structures of FoADH and ZoADH

544 In order to gain a deeper understanding of the molecular function, we performed X-ray
545 crystallography studies of FoADH and ZoADH. For the determination of the functional states

546 of both ADHs, the essential NAD⁺ cofactor was added to purified FoADH and ZoADH proteins
547 before crystallization. The crystal structures of FoADH and ZoADH in complex with NAD⁺ were
548 determined at a resolution of 2.5 and 2.1 Å, respectively (Table S2). FoADH and ZoADH
549 crystals belong to space group monoclinic P2₁ and orthorhombic P2₁2₁2₁, respectively and
550 contain four and eight molecules in asymmetric unit, respectively (Fig. S9). The electron
551 density map of FoADH and ZoADH clearly showed the almost entire polypeptide chain, except
552 for a partially disordered fragment of the loop between the β5 and β6-strands (Gly111-His115
553 in both enzymes), which is involved in substrate binding and specificity. The monomer
554 structures of FoADH and ZoADH comprise the catalytic domain (residues 1–149 and residues
555 283–326 for both enzymes) and the cofactor-binding domain (residues 150–282 for both
556 enzymes) (Fig. 6a), which are separated by a cleft containing a deep pocket, which
557 accommodates the substrate and the NAD⁺ cofactor. The catalytic domain contains two zinc-
558 binding sites, Zn1 and Zn2, which are responsible for catalytic activity and structural stability,
559 respectively. The cofactor binding domain adopts a typical Rossmann fold with the conserved
560 sequence “GXGXXG”. FoADH and ZoADH had a 76.0% similarity in amino acid sequence
561 (Fig. S10), and their monomer structures showed a similarity with a root-mean-square
562 deviation (r.m.s.d.) of 0.350-0.772 Å (Table S4).

563 In FoADH, molecules A/B/C/D and E/F/G/H form a tetrameric formation (Fig. S9). In
564 superimposition of monomeric FoADH molecules, the A, B, C, E and G molecules showed
565 structural similarity (denoted as closed form) with r.m.s.d. of 0.256-0.353 Å, whereas
566 molecules D and H (denoted as open form) showed the relatively high r.m.s.d. value of 0.457-
567 0.626 Å when superimposed with molecule A, B, C, E and G (Fig. 6b and Table S5). On the
568 other hand, molecule F maintains the intermediate conformation between the closed and open
569 conformations. When the cofactor binding domains of molecules A and H of FoADH were
570 superimposed, the catalytic binding of molecule H was shifted by approximately 2.0-3.3 Å in
571 the opposite direction of the substrate-binding cleft compared to molecule A (Fig. 6b).

572 In ZoADH, superimposition of molecules A, B and C exhibited a similar conformation (denoted
573 as closed form) with r.m.s.d. of 0.198-0.226 Å, whereas molecule D (denoted as open form)
574 showed a relatively high r.m.s.d. value of 0.314-0.471 Å when superimposed with molecules
575 A, B and C (Fig. 6b and Table S6). Superposition of the cofactor binding domains of molecules
576 A and D clearly revealed the conformational difference between the catalytic domains. The
577 catalytic domain of molecule D is shifted about 2.2-3.3 Å to the outside of the substrate binding
578 cleft of ZoADH compared to molecule A. Accordingly, in the structure of NAD⁺-bound FoADH,
579 molecules A/B/C and D represent closed and open conformations of the substrate binding site,
580 respectively. Collectively, the crystal structures of NAD⁺-bound ZoADH and FoADH contain

581 the open and closed conformations between catalytic and cofactor-binding domains (see
582 below).

583 The crystal structures of FoADH and ZoADH showed the tetrameric formation via the
584 arrangement of a dimer of dimers (Fig. 6c). In both ADHs, the β 17 and β 18-strands of the
585 cofactor binding domains are stabilized by forming an antiparallel β -sheet with the β 17* and
586 β 18*strands (asterisk indicates the second monomer), respectively (Figs. S11 and S12). For
587 FoADH, the dimeric interface is stabilized by the main chain interactions of Ile297-Ile299*
588 (* denoting the partner molecule) and Ile299-Ile297* between the β 17 strands and Tyr310-
589 Tyr310* between β 18 strands (Fig. S11). In addition, numerous hydrogen and salt bridges
590 were observed in the dimer interface with a buried surface area of 1654 Å² (Table S7). The
591 dimer of dimers was stabilized by hydrogen interaction and the buried interface of dimers of
592 dimers is 1193 Å² (Table S7). For ZoADH, the dimeric interface is stabilized by the main chain
593 interactions of Ile298-Ile300* and Ile300-Ile298* between the β 17 strand and Tyr311-Tyr311*
594 between β 18 strand (Fig. S12). Moreover, numerous hydrogen and salt bridges were observed
595 at the dimer interface with a buried surface area of 1640 Å² (Table S8). The dimer of dimers
596 was stabilized by hydrogen interactions and salt bridges and the buried interface of dimers of
597 dimers is ~1205 Å² (Table S8). All active sites of the tetrameric ADH in the crystal were
598 exposed to solvent (Fig. 6c). Superposition of tetrameric molecules of FoADH and ZoADH in
599 the asymmetric unit shows an r.m.s.d. of 0.327-0.888 Å for whole C α atoms (Fig. 6d).

600 Structural homology search by DALI revealed that both FoADH and ZoADH share structural
601 similarities to the class II alcohol dehydrogenase (ADH4) from human (PDB code: 3COS,
602 Z-score= 45.8 for FoADH and 45.3 for ZoADH, sequence identity= 32% for FoADH [357 α
603 atoms] and 30% for ZoADH [357 α atoms]), an ADH from *E. coli* (PDB code: 5vm2, Z-score=
604 48.1 for FoADH and 38.1 for ZoADH, sequence identity= 28% for FoADH [329 α atoms] and
605 27% for ZoADH [328 α atoms]) as well as an ADH from *Thermotoga maritima* (PDB code: 3IP1,
606 Z-score= 35.8 for FoADH and 36.8 for ZoADH, sequence identity= 25% for FoADH [328 α
607 atoms] and 23% for ZoADH [332 α atoms]). Although these structural homologous ADHs share
608 low amino acid sequence similarities with less than 32% compared to FoADH and ZoADH, the
609 active site residues involved in the Zn²⁺ and NAD⁺ binding are highly conserved (Fig. S10). In
610 addition, the NAD⁺-binding domain exhibits a typical Rossmann fold motif and has the classical
611 conserved sequence "GXGXXG" as in other ADHs and the topologies of those ADHs are highly
612 similar (Fig. S10). The overall topology of those homolog structures was similar with FoADH
613 and ZoADH (Fig. S13). However, superimposition of those ADH structures revealed that there
614 is a large difference in conformation between catalytic and cofactor-binding domains with a
615 r.m.s.d. of 1.373-2.963 Å for FoADH and 1.376-2.191 for ZoADH (Fig. 6e), indicating that they
616 possess large distinct NAD⁺ and substrate-binding clefts. Meanwhile, ADHs from *E. coli* and

617 *T. maritima* also formed the tetrameric formation in crystal structures like FoADH and ZoADH
618 (Fig. S13). These ADHs have the similar tetrameric assembly, however the superimposition of
619 the tetrameric ADHs showed that these tetrameric assemble have low similarity with a r.m.s.d
620 of 17.68~29.94 Å.

621 NAD⁺ and Zn²⁺-binding sites of FoADH and ZoADH

622 While NAD⁺ is the required cofactor for alcohol oxidation, Zn²⁺ interacts with the alcohol
623 molecule in the active site. The electron density maps of a NAD⁺ molecule and two zinc ions
624 are clearly observed in a substrate binding cleft of both FoADH and ZoADH (Fig. S14). The
625 binding configuration of NAD⁺ and the Zn²⁺ ions of ZoADH and FoADH are highly similar
626 (Fig. 7a). The adenine ring of NAD⁺ was located in the hydrophobic pocket formed by
627 hydrophobic interaction (Ile219, Leu245, Thr268, Ile270 and Leu273 for FoADH, Ile220,
628 Leu246, Thr269, Ile271 and Leu274 for ZoADH). The adenine ribose appears to be in a C2'-
629 endo conformation, and the O2' and O3' -hydroxyl group of ribose forms a hydrogen bond with
630 the side chain of aspartate (Asp218 for FoADH and Asp219 for ZoADH). The pyrophosphate
631 moiety of the NAD⁺ interacts with the nitrogen atoms of the main chain of glycine-valine residue
632 (Gly197-Val198 for FoADH and Gly198 and Val199 for ZoADH) that forms the loop between
633 strand β5 and helix α4. The nicotinamide ribose is in a C2'-endo conformation, and hydrogen
634 bonds are formed between the ribose O2' -hydroxyl group and threonine (Thr43 for FoADH
635 and ZoADH). The nicotinamide ring is in the anti-conformation. The carboxamide nitrogen
636 atom of the nicotinamide ring interacted with main-chain of proline (Pro313 for FoADH and
637 Pro314 for ZoADH) and valine (Val290 for FoADH and Val291 for ZoADH). The carboxamide
638 oxygen atom of the nicotinamide ring interacted with main-chain of tyrosine (Tyr315 for FoADH
639 and Tyr316 for ZoADH). Therefore, in both FoADH and ZoADH, the NAD⁺ molecules are
640 stabilized by hydrophobic and hydrogen bonds interactions.

641 In both FoADH and ZoADH, two zinc ions are commonly observed in the active site (Zn1 site)
642 and in a loop between α2 and β7 (Zn2 site) (Fig. 7a and Fig. S14). The zinc ion at Zn1 site is
643 coordinated by conserved cysteine and histidine residues (Cys41, His58, and Cys169 for
644 FoADH and Cys41, His58, and Cys170 for ZoADH) in the catalytic domain. The zinc ion at Zn2
645 site is involved in the protein stability and is tetrahedrally coordinated by conserved cysteine
646 residues (Cys88, Cys91, Cys94, and Cys102 for both enzymes) (Fig. S14). There result
647 indicated that ZoADH and FoADH showed high structural similarity for the NAD⁺ and zinc
648 binding configuration.

649 Different structural conformations were observed between monomeric ADHs in the tetrameric
650 formation of FoADH and ZoADH (Fig. 6b), indicating that they exhibit structurally different
651 substrate binding cleft and active site. In both results of superimposition of the active sites of

652 FoADH and ZoADH, the positions of the NAD⁺ and Zn₂ sites were similar, whereas a
653 significant difference was observed in the positions of the catalytic Zn₁ sites (Fig. 7b). In
654 FoADH and ZoADH, the maximum distances between metals from the Zn₁ site were 2.57 and
655 2.60 Å, respectively, from closed and open conformation of two domains of ADHs (Fig. 7b).

656 Since the substrate binds to the Zn₁ site and a dehydrogenase reaction occurs through the
657 interaction of NAD⁺ with the hydroxyl group, the size of the space between NAD⁺ and Zn₁ is
658 involved in substrate selectivity. The closest/longest distance between the Zn²⁺ and C5 atom
659 of the nicotinamide ring of NAD⁺ in FoADH and ZoADH were 3.21/4.91 Å, and 3.46/5.49 Å,
660 respectively (Fig. 7c). These different distance between Zn²⁺ and NAD⁺ were caused by
661 different from closed and open conformation of FoADH and ZoADH.

662 The electrostatic surfaces of FoADH and ZoADH showed that the substrate binding sites
663 commonly exhibited a hydrophobic surface (Fig. 7c). The space of the substrate binding site
664 of FoADH in closed and open conformation were approximately 3.4 x 4.2 Å and 3.9 x 5.4 Å,
665 respectively (Fig. S15). In the closed and open conformation of FoADH, His42 and Ala270 are
666 apart by 3.60 and 5.60 Å, respectively, showing the both surface structure surrounding the
667 NAD⁺ (Fig. 7c). ZoADH also exhibits open and closed conformations similar to FoADH, but the
668 distance of open conformation is relatively wide. The space of the substrate binding site of
669 ZoADH in closed and open conformation were approximately 3.0 x 3.8 Å and 3.8 x 4.9 Å,
670 respectively (Fig. S15). In the closed conformation of ZoADH, the catalytic domain and the
671 cofactor domain are close to each other, especially His42 and Ala270 by a distance of 3.88 Å,
672 indicating the surface structure surrounding the NAD⁺ (Fig. 7c). On the other hand, in the open
673 conformation of ZoADH, His42 and Ala270 are apart by 6.81 Å, and accordingly, the entire
674 NAD⁺ molecule in the surface structure is exposed to the solvent (Fig. 7c).

675 **Discussion**

676 In the present work, FoADH from *F. agariphila* KMM 3901^T and ZoADH from *Z. galactanivorans*
677 Dsij^T were characterized in detail to draw conclusions about their biological function. Three
678 main conclusions regarding the biological function can be derived by the knockout of the genes
679 encoding for ZoADH and CYP in *Z. galactanivorans* and subsequent growth studies on
680 D-galactose and G6Me. First, we confirmed the hypothesis of Reisky et al. that in the absence
681 of CYP-catalyzed oxidative demethylation, a G6Me utilization as sole carbon source is
682 infeasible for the organism (Reisky et al. 2018). Surprisingly, knockout of the ZoADH gene also
683 caused diminished growth of *Z. galactanivorans* in the presence of G6Me. Second, due this
684 observation, we can conclude a significant role of this ADHs in G6Me utilization in these marine
685 bacteria. From an ecological perspective, this has additional importance for the marine
686 carbohydrate degraders. G6Me can occur up to 28% within the porphyran chain (Rees and

687 Conway 1962). Thus, a reduced utilization of G6Me would represent a substantial potential
688 loss as a carbon source for the organism. Third, since normal growth was observed in the
689 presence of D-galactose as sole carbon source, a function in D-galactose metabolism can be
690 excluded. This was also supported by the observation that both ADHs lacked activity for
691 D-galactose. The ADHs are therefore probably involved in oxidative demethylation or a
692 subsequent reaction. Since no activity was observed for G6Me, the substrate of oxidative
693 demethylation could be excluded. Consequently, we hypothesized that the ADHs are involved
694 in the detoxification of formaldehyde, which is a by-product of the oxidative demethylation
695 reaction. This was also supported by the resistance of both ADHs to formaldehyde exposure.
696 Formaldehyde is a toxic metabolite due to its properties as a highly reactive electrophile. It can
697 react with free amino and thiol groups of proteins and nucleic acids, leading to protein and
698 DNA damages as well as cross-link formations (Chen et al. 2016; Shishodia et al. 2018; Tayri-
699 Wilk et al. 2020). It has been shown that higher concentrations of formaldehyde can negatively
700 affect the growth of *Z. galactanivorans* (Brott et al. 2022). Thus, a reduced growth of the ADH
701 knock-out strain could be explained by the potential accumulation of formaldehyde. There are
702 numerous metabolic pathways in which formaldehyde can be detoxified (Yurimoto et al. 2005;
703 Klein et al. 2022). However, in the thiol-dependent formaldehyde detoxification, a zinc-
704 dependent ADH and an esterase perform the key reactions (Sanghani et al. 2000; Gonzalez
705 et al. 2006). Genome neighborhood analysis revealed that in most marine bacteria that
706 possess the ADH gene, it was located in close proximity to a gene encoding for an esterase in
707 addition to the CYP gene. We therefore investigated whether the ADH catalyzed a thiol-
708 dependent detoxification of formaldehyde. However, with glutathione, mycothiol, and
709 bacillithiol as thiol cofactors, no activity was detected for either ADH. These observations can
710 be further explained with the crystal structures of both ADHs; sterically demanding compounds
711 such as mycothiol or bacillithiol cannot fit into the narrow active site of these enzymes. These
712 observations are also consistent with the results from the sequence similarity network, in which
713 glutathione- and mycothiol-dependent formaldehyde dehydrogenases were predominantly
714 present in different clusters (main clusters 1 and 4) than the ADHs (main cluster 2). Since no
715 activity could be detected with literature-known cofactors, additional thiols were considered;
716 however, no activity could be observed either. Thiol cofactors are still being discovered
717 (Newton and Rawat 2019), perhaps marine organisms also possess an unidentified thiol,
718 which can serve as a cofactor for this reaction. Since no activity was observed for
719 formaldehyde without an additional thiol cofactor, the biological function of a thiol-independent
720 formaldehyde dehydrogenase was excluded. In addition, some ADHs can possess dismutase
721 activities (Trivić et al. 1999). A formaldehyde dismutase catalyzes the disproportionation of
722 formaldehyde to methanol and formic acid in the presence of a covalent-bound NAD⁺
723 (Yonemitsu and Kikuchi 2018). However, this reaction could not be detected. Both organisms

724 harbor other metabolic pathways for the detoxification of formaldehyde (Brott et al. 2022). For
725 instance, in *Z. galactanivorans*, the genes encoding for the key enzymes of the ribulose
726 monophosphate pathway are upregulated in the presence of porphyran (Brott et al. 2022), so
727 an accumulation of formaldehyde is unlikely. Eventually, the ADHs might have a completely
728 different biological function such as the regeneration of NADH (Hilberath et al. 2021; Kokorin
729 et al. 2021). In the oxidative demethylation reaction, NADH is oxidized to NAD⁺, a reduced
730 growth in the ADH knockout strain due to cofactor depletion might be possible. NADH could
731 be regenerated by oxidation of an unknown component or by the thiol-dependent
732 formaldehyde detoxification pathway. However, it is doubtful that the loss of one single enzyme
733 would cause such a tremendous effect on NADH/NAD⁺ homeostasis. Additionally, the ADHs
734 displayed predominantly activity for the reduction of aldehydes under NADH consumption,
735 therefore recycling of a cofactor is improbable.

736 Both ADHs possessed predominantly activity for aromatic substances, resulting in a substrate
737 specificity resembling partially those of cinnamyl alcohol and/or benzyl alcohol
738 dehydrogenases (Larroy et al. 2002; Willson et al. 2022). However, the highest activity was
739 observed for pyridine-3-carbaldehyde and furan derivatives. Furfural is generally produced as
740 a side product by pretreating lignocellulosic biomass for the production of bioethanol. Under
741 acidic conditions and high temperatures, dehydration of pentoses and hexoses proceeds,
742 leading to the formation of furfural or hydroxymethylfurfural. Furfural acts as an inhibitor in
743 subsequent bioethanol-producing fermentations by bacteria by prolonging the lag phase of
744 growth and thereby the fermentation time (Mariscal et al. 2016). Consequently, these marine
745 bacteria possess ADHs that catalyze the potential removal of furfural although the biological
746 function may be different. The ADHs lacked activity for various sugar substrates, which
747 excluded a polyol dehydrogenase activity. Activity for any other monosaccharides,
748 disaccharides or even oligosaccharides formed during porphyran degradation is unlikely as
749 well, considering the substrate specificity of the enzymes based on the narrow active site. The
750 data from biochemical characterizations are discussed in the SI.

751 We have determined the crystal structures of FoADH and ZoADH complexed with NAD⁺ and
752 two zinc ions. These ADHs showed high structural similarity in terms of topology and assembly.
753 On the one hand, these two ADHs showed similarities in topology with other ADHs from human,
754 *E. coli* and *T. maritima*, but showed distinct conformation between the cofactor and catalytic
755 domains of those ADHs. On the other hand, the crystal structures of FoADH and ZoADH
756 showed open and closed conformations, indicating that the conformation between the two
757 domains can change in the state where the substrate is not bound. These distinct
758 conformations of FoADH and ZoADH represent different substrate binding pockets. When they
759 exhibit an open conformation between the two domains of FoADH and ZoADH, they form a

760 broadened substrate-binding pocket. Accordingly, in terms of substrate accessibility, we
761 consider that substrate accessibility will be easier when FoADH and ZoADH have an open
762 conformation.

763 During substrate recognition, when the converting functional group from the substrate
764 approaches the Zn1 site on the substrate binding pocket of FoADH and ZoADH, the rest of the
765 substrate is exposed to the nicotinamide of NAD⁺ or the hydrophobic surface. Considering that
766 the nicotinamide group of NAD⁺ is involved in the oxidoreductase mechanism of the ADH, the
767 substrate would prefer to be located to the hydrophobic surface rather than the nicotinamide
768 group of NAD⁺. Accordingly, FoADH and ZoADH may prefer substrates having a hydrophobic
769 body. Our biochemical studies showed that both enzymes prefer aromatic substrates. We
770 expected that the aromatic ring of the substrate may be located on a hydrophobic surface
771 nearby the substrate binding pocket of FoADH and ZoADH. In this case, the aromatic ring of
772 the substrate could interact with the Phe136 residue in the hydrophobic surfaces of the
773 enzymes. Based on the active site structures of both ADH computational docking of a substrate
774 will be able to provide an insight into the molecular mechanism and substrate specificity.
775 However, from the results of this study, ZoADH and FoADH have various conformations
776 between catalytic and cofactor binding domain in NAD⁺ and two zinc ion binding states,
777 indicating the computational docking results could be different depending on the applied model
778 structure. Also, based on our results, we concluded that the docking results may be different
779 from biochemical experiments if the active sites of ZoADH and FoADH may have different
780 conformations. Therefore, to better understand the substrate specificity, the crystal structures
781 of ZoADH and FoADH in complex with the biological substrate will be needed in the future.

782 In summary, in this study we determined the putative functions of a conserved ADH from
783 marine *Flavobacteriia*. Additionally, we provided the crystal structures of the enzymes of
784 *F. agariphila* and *Z. galactanivorans*. Enzymatic studies revealed the preferential conversion
785 of aromatic aldehydes. We revealed that these enzymes are not involved in formaldehyde
786 detoxification or in subsequent reaction of the oxidative demethylation of G6Me. Based on
787 gene knockouts, we demonstrated the essential role of these ADHs in the utilization of marine
788 algal sugars. Our study indicates a potential auxiliary activity of these ADHs in the utilization
789 of marine sugars by marine *Flavobacteriia*.

790 **Declarations**

791 **Funding:** François Thomas acknowledges support from the French government via the
792 National Research Agency program ALGAVOR (ANR-18-CE02-0001). We thank the German
793 Research Foundation (DFG) for funding through the Research Unit FOR2406

794 “Proteogenomics of Marine Polysaccharide Utilization” (POMPU) (grants# BO 1862/17-1 and
795 BO 1862/17-2 to U.T.B. and SCHW 595/10-2 to T.S.).

796 **Conflicts of interests:** The authors have no relevant financial or non-financial interests to
797 disclose.

798 **Author contributions:** M.G., T.S. and U.T.B. initiated the study and directed the project. F.T.
799 conducted the growth studies and created the knock-out strain. K.H.N performed the
800 crystallization and structural analyses. T.D. performed the computational analysis. L. R. and
801 H.C.G. performed the cloning and initial experiments on enzyme function of FoADH. S.B. and
802 M.B. expressed and purified the enzymes and performed further experiments on enzyme
803 function and characterization. S.B. and K.H.N prepared the main manuscript, which was
804 revised by F.T., T.D, H.C.G., L. R., M.G, T.S. and U.T.B. and was approved by all authors.

805 **Ethical approval:** Not applicable.

806 **Data availability:** The datasets generated during and/or analyzed during the current study are
807 available from the corresponding author on reasonable request.

808

809 **References**

810 Achkor H, Díaz M, Fernández MR, Biosca JA, Parés X, Martínez MC (2003) Enhanced
811 formaldehyde detoxification by overexpression of glutathione-dependent formaldehyde
812 dehydrogenase from *Arabidopsis*. *Plant Physiol.* 132:2248-2255

813 Arnosti C, Wietz M, Brinkhoff T, Hehemann JH, Probandt D, Zeugner L, Amann R (2021) The
814 biogeochemistry of marine polysaccharides: sources, inventories, and bacterial drivers of
815 the carbohydrate cycle. *Ann Rev Mar Sci* 13:81–108.

816 Bauer M, Kube M, Teeling H, Richter M, Lombardot T, Allers E, Würdemann CA, Quast C,
817 Kuhl H, Knaust F, Woebken D, Bischof K, Mussmann M, Choudhuri J V., Meyer F,
818 Reinhardt R, Amann RI, Glöckner FO (2006) Whole genome analysis of the marine
819 Bacteroidetes “*Gramella forsetii*” reveals adaptations to degradation of polymeric organic
820 matter. *Environ. Microbiol.* 8:2201–2213.

821 Bäumgen M, Dutschei T, Bartosik D, Suster C, Reisky L, Gerlach N, Stanetty C, Mihovilovic
822 MD, Schweder T, Hehemann J-H, Bornscheuer UT (2021a) A new carbohydrate-active
823 oligosaccharide dehydratase is involved in the degradation of ulvan. *J. Biol. Chem.*
824 297:101210.

825 Bäumgen M, Dutschei T, Bornscheuer UT (2021b) Marine Polysaccharides: Occurrence,
826 Enzymatic Degradation and Utilization. *ChemBioChem* 22:2247–2256.

827 Brott S, Thomas F, Behrens M, Methling K, Bartosik D, Dutschei T, Lalk M, Michel G, Schweder
828 T, Bornscheuer UT (2022) Connecting Algal Polysaccharide Degradation to
829 Formaldehyde Detoxification. *ChemBioChem* 23:e202200269.

830 Brunet M, de Bettignies F, Le Duff N, Tanguy G, Davoult D, Leblanc C, Gobet A, Thomas F
831 (2021) Accumulation of detached kelp biomass in a subtidal temperate coastal ecosystem
832 induces succession of epiphytic and sediment bacterial communities. *Environ. Microbiol.*

- 833 23:1638–1655.
- 834 Chandrangu P, Loi V Van, Antelmann H, Helmann JD (2018) The role of bacillithiol in Gram-
835 Positive *Firmicutes*. *Antioxid. Redox Signal.* 28:445–462.
- 836 Chen NH, Djoko KY, Veyrier FJ, McEwan AG (2016) Formaldehyde stress responses in
837 bacterial pathogens. *Front Microbiol.* 7:257.
- 838 de Lorenzo V, Timmis KN (1994) Analysis and construction of stable phenotypes in gram-
839 negative bacteria with Tn5- and Tn10-derived minitransposons. *Methods Enzymol.*
840 235:386–405.
- 841 Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr*
842 *D Biol Crystallogr.* 60:2126–2132.
- 843 Ficko-Blean E, Préchoux A, Thomas F, Rochat T, Larocque R, Zhu Y, Stam M, Génicot S, Jam
844 M, Calteau A, Viart B, Ropartz D, Pérez-Pascual D, Correc G, Matard-Mann M, Stubbs
845 KA, Rogniaux H, Jeudy A, Barbeyron T, Médigue C, Czjzek M, Vallenet D, McBride MJ,
846 Duchaud E, Michel G (2017) Carrageenan catabolism is encoded by a complex regulon
847 in marine heterotrophic bacteria. *Nat. Commun.* 8:1685.
- 848 Field CB (1998) Primary production of the biosphere: Integrating terrestrial and oceanic
849 components. *Science* 281:237–240.
- 850 Gonzalez CF, Proudfoot M, Brown G, Korniyenko Y, Mori H, Savchenko A V., Yakunin AF
851 (2006) Molecular basis of formaldehyde detoxification: Characterization of two
852 S-formylglutathione hydrolases from *Escherichia coli*, FrmB and YeiG. *J. Biol. Chem.*
853 281:14514–14522.
- 854 Gouet P, Courcelle E, Stuart DI, Métoz F (1999) ESPript: analysis of multiple sequence
855 alignments in PostScript. *Bioinformatics* 15:305–308.
- 856 Grondin JM, Tamura K, Déjean G, Abbott DW, Brumer H (2017) Polysaccharide utilization loci:
857 Fueling microbial communities. *J. Bacteriol.* 199:e00860-16.
- 858 Gutheil WG, Holmquist B, Vallee BL (1992) Purification, Characterization, and partial
859 sequence of the Glutathione-Dependent Formaldehyde Dehydrogenase from *Escherichia*
860 *coli*: A Class III Alcohol Dehydrogenase. *Biochemistry* 31:475–481.
- 861 Hall M, Bommaris AS (2011) Enantioenriched compounds via enzyme-catalyzed redox
862 reactions. *Chem. Rev.* 111:4088–4110.
- 863 Hambidge M, Cousins RJ, Costello RB (2000) Zinc and health: Current status and future
864 directions: Introduction. *J. Nutr.* 130:1344S-1349S
- 865 Hand CE, Honek JF (2005) Biological chemistry of naturally occurring thiols of microbial and
866 marine origin. *J. Nat. Prod.* 68:293–308.
- 867 Hilberath T, Raffaele A, Windeln LM, Urlacher VB (2021) Evaluation of P450 monooxygenase
868 activity in lyophilized recombinant *E. coli* cells compared to resting cells. *AMB Express*
869 11:162.
- 870 Klein VJ, Irla M, López MG, Brautaset T, Brito LF (2022) Unravelling formaldehyde metabolism
871 in bacteria: Road towards Synthetic Methylophony. *Microorganisms* 10:220.
- 872 Klemetsen T, Raknes IA, Fu J, Agafonov A, Balasundaram V, Tartari G, Robertsen E,
873 Willassen NP (2018) The MAR databases: development and implementation of
874 databases specific for marine metagenomics. 46:692–699.
- 875 Koesoema AA, Standley DM, Senda T, Matsuda T (2020) Impact and relevance of alcohol
876 dehydrogenase enantioselectivities on biotechnological applications. *Appl. Microbiol.*
877 *Biotechnol.* 104:2897–2909.

- 878 Kokorin A, Parshin PD, Bakkes PJ, Pometun AA, Tishkov VI, Urlacher VB (2021) Genetic
879 fusion of P450 BM3 and formate dehydrogenase towards self-sufficient biocatalysts with
880 enhanced activity. *Sci. Rep.* 11:21706.
- 881 Kracher D, Ludwig R (2016) Cellobiose dehydrogenase: An essential enzyme for
882 lignocellulose degradation in nature - A review. *Die Bodenkultur: Journal of Land*
883 *Management, Food and Environment* 67:145–163.
- 884 Krause-Jensen D, Duarte CM (2016) Substantial role of macroalgae in marine carbon
885 sequestration. *Nat. Geosci.* 9:737–742.
- 886 Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state.
887 *J. Mol. Biol.* 372:774–797.
- 888 Lapébie P, Lombard V, Drula E, Terrapon N, Henrissat B (2019) Bacteroidetes use thousands
889 of enzyme combinations to break down glycans. *Nat. Commun.* 10: 2043.
- 890 Larroy C, Parés X, Biosca JA (2002) Characterization of a *Saccharomyces cerevisiae*
891 NADP(H)-dependent alcohol dehydrogenase (ADHVII), a member of the cinnamyl alcohol
892 dehydrogenase family. *Eur. J. Biochem.* 269:5738–5745.
- 893 Li C, Wen A, Shen B, Lu J, Huang Y, Chang Y (2011) FastCloning: a highly simplified,
894 purification-free, sequence- and ligation-independent PCR cloning method. *BMC*
895 *Biotechnol.* 11:92.
- 896 Liebschner D, Afonine P V, Baker ML, Bunkóczi G, Chen VB, Croll TI, Hintze B, Hung LW,
897 Jain S, McCoy AJ, Moriarty NW, Oeffner RD, Poon BK, Prisant MG, Read RJ, Richardson
898 JS, Richardson DC, Sammito MD, Sobolev O V, Stockwell DH, Terwilliger TC,
899 Urzhumtsev AG, Videau LL, Williams CJ, Adams PD (2019) Macromolecular structure
900 determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta*
901 *Crystallogr. Sect. D, Struct. Biol.* 75:861–877.
- 902 Lu F, Xu W, Zhang W, Guang C, Mu W (2019) Polyol dehydrogenases: intermediate role in
903 the bioconversion of rare sugars and alcohols. *Appl. Microbiol. Biotechnol.* 103:6473–
904 6481.
- 905 Mariscal R, Maireles-Torres P, Ojeda M, Sádaba I, López Granados M (2016) Furfural: A
906 renewable and versatile platform molecule for the synthesis of chemicals and fuels.
907 *Energy Environ. Sci.* 9:1144–1189.
- 908 Martens EC, Koropatkin NM, Smith TJ, Gordon JI (2009) Complex glycan catabolism by the
909 human gut microbiota: The Bacteroidetes sus-like paradigm. *J. Biol. Chem.* 284:24673–
910 24677.
- 911 Martínez-Martínez M, Coscolín C, Santiago G, Chow J, Stogios PJ, Bargiela R, Gertler C,
912 Navarro-Fernández J, Bollinger A, Thies S, Méndez-García C, Popovic A, Brown G,
913 Chernikova TN, García-Moyano A, Bjerga GEK, Pérez-García P, Hai T, Del Pozo M V.,
914 Stokke R, Steen IH, Cui H, Xu X, Nocek BP, Alcaide M, Distaso M, Mesa V, Peláez AI,
915 Sánchez J, Buchholz PCF, Pleiss J, Fernández-Guerra A, Glöckner FO, Golyshina O V.,
916 Yakimov MM, Savchenko A, Jaeger KE, Yakunin AF, Streit WR, Golyshin PN, Guallar V,
917 Ferrer M (2018) Determinants and Prediction of Esterase Substrate Promiscuity Patterns.
918 *ACS Chem. Biol.* 13:225–234.
- 919 Misset-Smits M, Van Ophem PW, Sakuda S, Duine JA (1997) Mycothiol, 1-O-(2'-[N-acetyl-L-
920 cysteinyl]amido-2'-deoxy- α -D-glucopyranosyl)-D-myo-inositol, is the factor of NAD/factor-
921 dependent formaldehyde dehydrogenase. *FEBS Lett.* 409:221–222.
- 922 Newton GL, Fahey RC (2002) Mycothiol biochemistry. *Arch Microbiol.* 178:388–394.
- 923 Newton GL, Rawat M (2019) N-methyl-bacillithiol, a novel thiol from anaerobic bacteria. *MBio*
924 10:e02634-18.

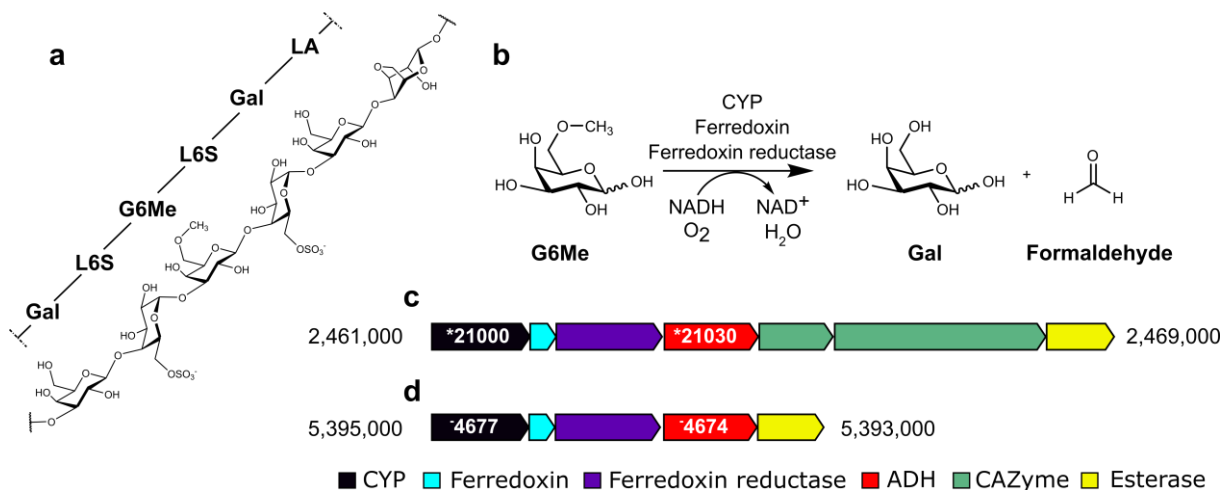
- 925 Newton GL, Rawat M, La Clair JJ, Jothivasan VK, Budiarto T, Hamilton CJ, Claiborne A,
926 Helmann JD, Fahey RC (2009) Bacillithiol is an antioxidant thiol produced in Bacilli. *Nat.*
927 *Chem. Biol.* 5:625–627.
- 928 Otwinowski Z, Minor WBT-M (1997) Processing of X-ray diffraction data collected in oscillation
929 mode. In: *Macromolecular Crystallography Part A. Methods Enzymol.* 276:307-326
- 930 Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B,
931 Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular
932 interaction networks. *Genome Res.* 13:2498-2504.
- 933 Persson B, Hedlund J, Jörnvall H (2008) Medium- and short-chain dehydrogenase/reductase
934 gene and protein families: The MDR superfamily. *Cell Mol. Life Sci.* 65:3879–3894.
- 935 Pick A, Rühmann B, Schmid J, Sieber V (2013) Novel CAD-like enzymes from *Escherichia coli*
936 K-12 as additional tools in chemical production. *Appl. Microbiol. Biotechnol.* 97:5815–
937 5824.
- 938 Priyam A, Woodcroft BJ, Rai V, Moghul I, Munagala A, Ter F, Chowdhary H, Pieniak I, Maynard
939 LJ, Gibbins MA, Moon HK, Davis-Richardson A, Uludag M, Watson-Haigh NS, Challis R,
940 Nakamura H, Favreau E, Gómez EA, Pluskal T, Leonard G, Rumpf W, Wurm Y (2019)
941 Sequenceserver: A Modern Graphical User Interface for Custom BLAST Databases. *Mol.*
942 *Biol. Evol.* 36:2922–2924.
- 943 Rao ST, Rossmann MG (1973) Comparison of super-secondary structures in proteins. *J. Mol.*
944 *Biol.* 76:241–256.
- 945 Rees DA, Conway E (1962) The structure and biosynthesis of porphyrin: a comparison of
946 some samples. *Biochem. J.* 84:411–416.
- 947 Reisky L, Préchoux A, Zühlke MK, Baumgen M, Robb CS, Gerlach N, Roret T, Stanetty C,
948 Larocque R, Michel G, Song T, Markert S, Unfried F, Mihovilovic MD, Trautwein-Schult
949 A, Becher D, Schweder T, Bornscheuer UT, Hehemann JH (2019) A marine bacterial
950 enzymatic cascade degrades the algal polysaccharide ulvan. *Nat. Chem. Biol.* 15:803–
951 812.
- 952 Reisky L, Büchenschütz HC, Engel J, Song T, Schweder T, Hehemann JH, Bornscheuer UT
953 (2018) Oxidative demethylation of algal carbohydrates by cytochrome P450
954 monooxygenases brief-communication. *Nat. Chem. Biol.* 14:342–344.
- 955 Robb CS, Reisky L, Bornscheuer UT, Hehemann JH (2018) Specificity and mechanism of
956 carbohydrate demethylation by cytochrome P450 monooxygenases. *Biochem. J.*
957 475:3875–3886.
- 958 Salentin S, Schreiber S, Haupt VJ, Adasme MF, Schroeder M (2015) PLIP: fully automated
959 protein-ligand interaction profiler. *Nucleic Acids Res.* 43:W443-7.
- 960 Sanghani PC, Stone CL, Ray BD, Pindel E V., Hurley TD, Bosron WF (2000) Kinetic
961 mechanism of human glutathione-dependent formaldehyde dehydrogenase.
962 *Biochemistry* 39:10720–10729.
- 963 Shishodia S, Zhang D, El-Sagheer AH, Brown T, Claridge TDW, Schofield CJ, Hopkinson RJ
964 (2018) NMR analyses on N-hydroxymethylated nucleobases-implications for
965 formaldehyde toxicity and nucleic acid demethylases. *Org. Biomol. Chem.* 16:4021–4032.
- 966 Sichert A, Corzett CH, Schechter MS, Unfried F, Markert S, Becher D, Fernandez-Guerra A,
967 Liebeke M, Schweder T, Polz MF, Hehemann JH (2020) Verrucomicrobia use hundreds
968 of enzymes to digest the algal polysaccharide fucoidan. *Nat. Microbiol.* 5:1026–1039.
- 969 Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M,
970 Söding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality

- 971 protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539.
- 972 Sirota FL, Maurer-Stroh S, Li Z, Eisenhaber F, Eisenhaber B (2021) Functional Classification
973 of Super-Large Families of Enzymes Based on Substrate Binding Pocket Residues for
974 Biocatalysis and Enzyme Engineering Applications. *Front. Bioeng. Biotechnol.* 9:701120.
- 975 Sützl L, Laurent CVFP, Abrera AT, Schütz G, Ludwig R, Haltrich D (2018) Multiplicity of
976 enzymatic functions in the CAZy AA3 family. *Appl. Microbiol. Biotechnol.* 102:2477–2492.
- 977 Takeda K, Matsumura H, Ishida T, Samejima M, Ohno H, Yoshida M, Igarashi K, Nakamura N
978 (2015) Characterization of a novel PQQ-dependent quinohemoprotein pyranose
979 dehydrogenase from *Coprinopsis cinerea* classified into auxiliary activities family 12 in
980 carbohydrate-active enzymes. *PLoS One* 10: e0115722.
- 981 Tayri-Wilk T, Slavin M, Zamel J, Blass A, Cohen S, Motzik A, Sun X, Shalev DE, Ram O,
982 Kalisman N (2020) Mass spectrometry reveals the chemistry of formaldehyde cross-
983 linking in structured proteins. *Nat. Commun.* 11: 3128.
- 984 Teeling H, Fuchs BM, Becher D, Klockow C, Gardebrecht A, Bennke CM, Kassabgy M, Huang
985 S, Mann AJ, Waldmann J, Weber M, Klindworth A, Otto A, Lange J, Bernhardt J, Reinsch
986 C, Hecker M, Peplies J, Bockelmann FD, Callies U, Gerdts G, Wichels A, Wiltshire KH,
987 Glöckner FO, Schweder T, Amann R (2012) Substrate-controlled succession of marine
988 bacterioplankton populations induced by a phytoplankton bloom. *Science* 336:608–611.
- 989 Thomas F, Hehemann JH, Rebuffet E, Czjzek M, Michel G (2011a) Environmental and gut
990 Bacteroidetes: The food connection. *Front. Microbiol.* 2:93.
- 991 Thomas F, Barbeyron T, Michel G (2011b) Evaluation of reference genes for real-time
992 quantitative PCR in the marine flavobacterium *Zobellia galactanivorans*. *J. Microbiol.*
993 *Methods* 84:61–66.
- 994 Trincone A (2011) Marine biocatalysts: Enzymatic features and applications. *Mar Drugs.*
995 9:478–499.
- 996 Trivić S, Leskova V, Winston GW (1999) Aldehyde dismutase activity of yeast alcohol
997 dehydrogenase. *Biotechnol. Lett.* 21:231–234.
- 998 Unsworth LD, Van Der Oost J, Koutsopoulos S (2007) Hyperthermophilic enzymes - Stability,
999 activity and implementation strategies for high temperature applications. *FEBS J.*
1000 274:4044–4056
- 1001 Vagin A, Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D Biol.*
1002 *Crystallogr.* 66:22–25.
- 1003 Vorholt JA (2002) Cofactor-dependent pathways of formaldehyde oxidation in methylotrophic
1004 bacteria. *Arch Microbiol.* 178:239–249.
- 1005 Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA,
1006 Hintze BJ, Chen VB, Jain S, Lewis SM, Arendall WB 3rd, Snoeyink J, Adams PD, Lovell
1007 SC, Richardson JS, Richardson DC (2018) MolProbity: More and better reference data
1008 for improved all-atom structure validation. *Protein Sci.* 27:293–315.
- 1009 Willson BJ, Herman R, Langer S, Thomas GH (2022) Improved furfural tolerance in
1010 *Escherichia coli* mediated by heterologous NADH-dependent benzyl alcohol
1011 dehydrogenases. *Biochem. J.* 479:1045–1058.
- 1012 Yonemitsu H, Kikuchi Y (2018) Biodegradation of high concentrations of formaldehyde using
1013 *Escherichia coli* expressing the formaldehyde dismutase gene of *Methylobacterium sp.*
1014 FD1. *Biosci. Biotechnol. Biochem.* 82:49–56.
- 1015 Yurimoto H, Kato N, Sakai Y (2005) Assimilation, dissimilation, and detoxification of

- 1016 formaldehyde, a central metabolic intermediate of methylotrophic metabolism. Chem. Rec.
1017 5:367–375.
- 1018 Zallot R, Oberg N, Gerlt JA (2019) The EFI Web Resource for genomic enzymology tools:
1019 Leveraging Protein, Genome, and Metagenome Databases to Discover Novel enzymes
1020 and Metabolic Pathways. Biochemistry 58:4169–4182.
- 1021 Zhang R, Xu Y, Xiao R (2015) Redesigning alcohol dehydrogenases/reductases for more
1022 efficient biosynthesis of enantiopure isomers. Biotechnol. Adv. 33:1671–1684.
- 1023 Zheng YG, Yin HH, Yu DF, Chen X, Tang XL, Zhang XJ, Xue YP, Wang YJ, Liu ZQ (2017)
1024 Recent advances in biotechnological applications of alcohol dehydrogenases. Appl.
1025 Microbiol. Biotechnol. 101:987–1001.
- 1026 Zhu Y, Thomas F, Larocque R, Li N, Duffieux D, Cladière L, Souchaud F, Michel G, McBride
1027 MJ (2017) Genetic analyses unravel the crucial role of a horizontally acquired alginate
1028 lyase for brown algal biomass degradation by *Zobellia galactanivorans*. Environ. Microbiol.
1029 19:2164–2181
- 1030 Zobell CE (1941) Studies on marine bacteria. I. The cultural requirements of heterotrophic
1031 aerobes. J. Mar. Res. 4:41–75

1032

1033 **Figures**

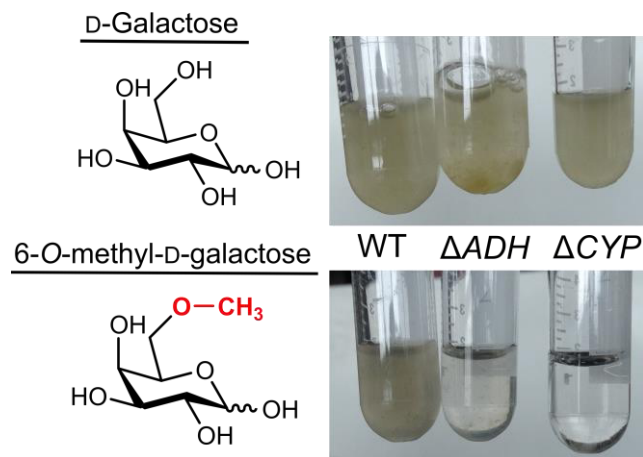


1034

1035 **Fig. 1** Porphyrans contain 6-O-methyl-D-galactose, which can be metabolized by marine
1036 bacteria via oxidative demethylation. **a**) Porphyrans, the common name of the galactan of red
1037 algae of the genus *Porphyra*, consist of chains composed mainly of the alternating
1038 monosaccharide units 4-linked- α -L-galactose-6-sulfate (L6S) and 3-linked- β -D-galactose (Gal)
1039 or 3,6-anhydro- α -L-galactose (LA). Furthermore, O-methylation of D-galactose results in the
1040 formation of 6-O-methyl-D-galactose (G6Me). **b**) The oxidative demethylation of G6Me is
1041 catalyzed by a cytochrome P450 monooxygenase in combination with its redox partners
1042 ferredoxin and ferredoxin reductase, producing D-galactose and formaldehyde in equimolar
1043 amounts. **c**) In *Formosa agariphila* KMM 3901^T and **d**) *Zobellia galactanivorans* Dsij^T, genes
1044 encoding for the key enzymes of oxidative demethylation are located in close proximity to a
1045 gene encoding for a zinc-dependent alcohol dehydrogenase. *BN863_, for example *21030
1046 refers to locus tag BN863_21030 for *F. agariphila* while zgal_, for example zgal_4674 refers to
1047 locus tag zgal_4674 for *Z. galactanivorans*.

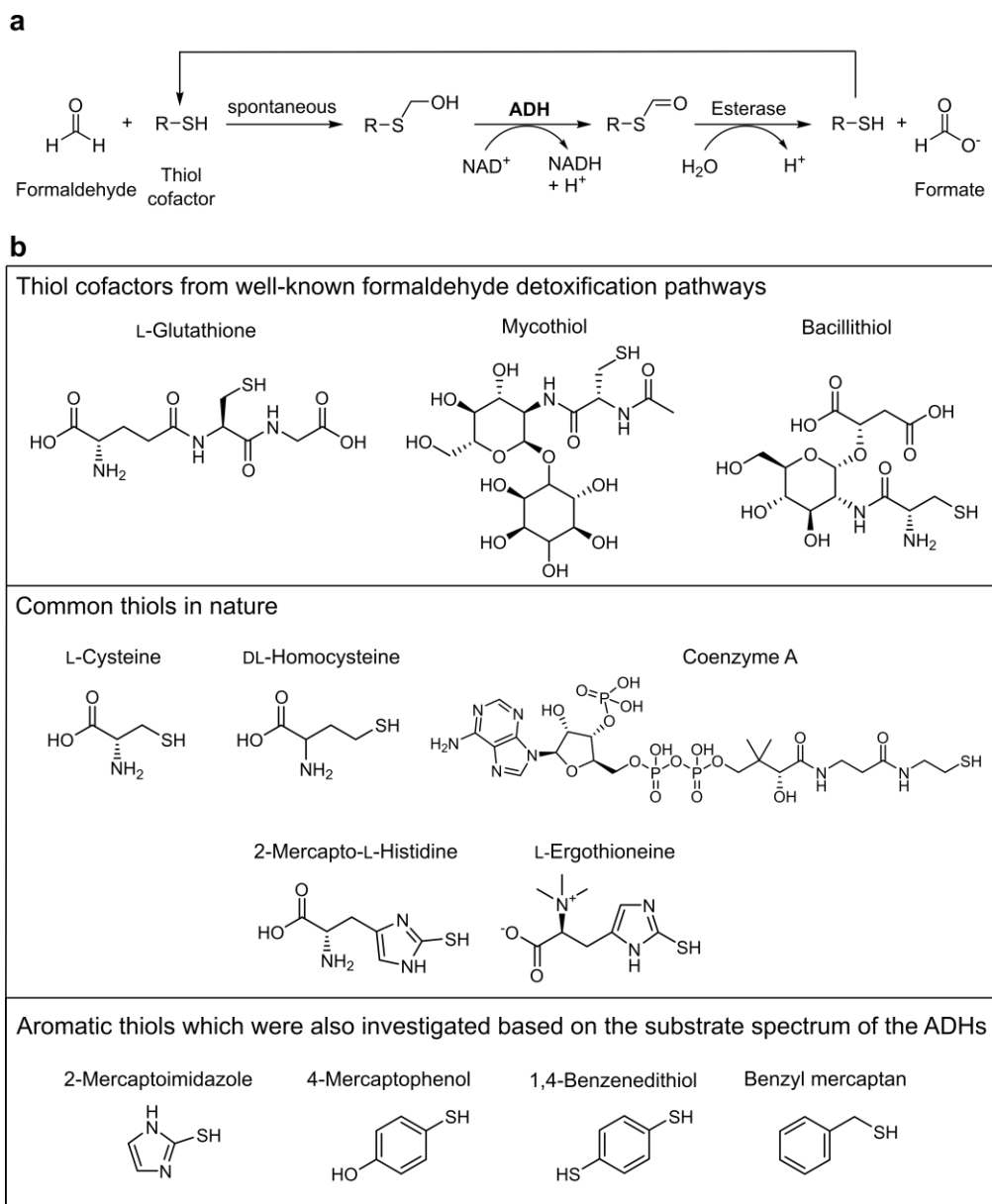
1048

1049



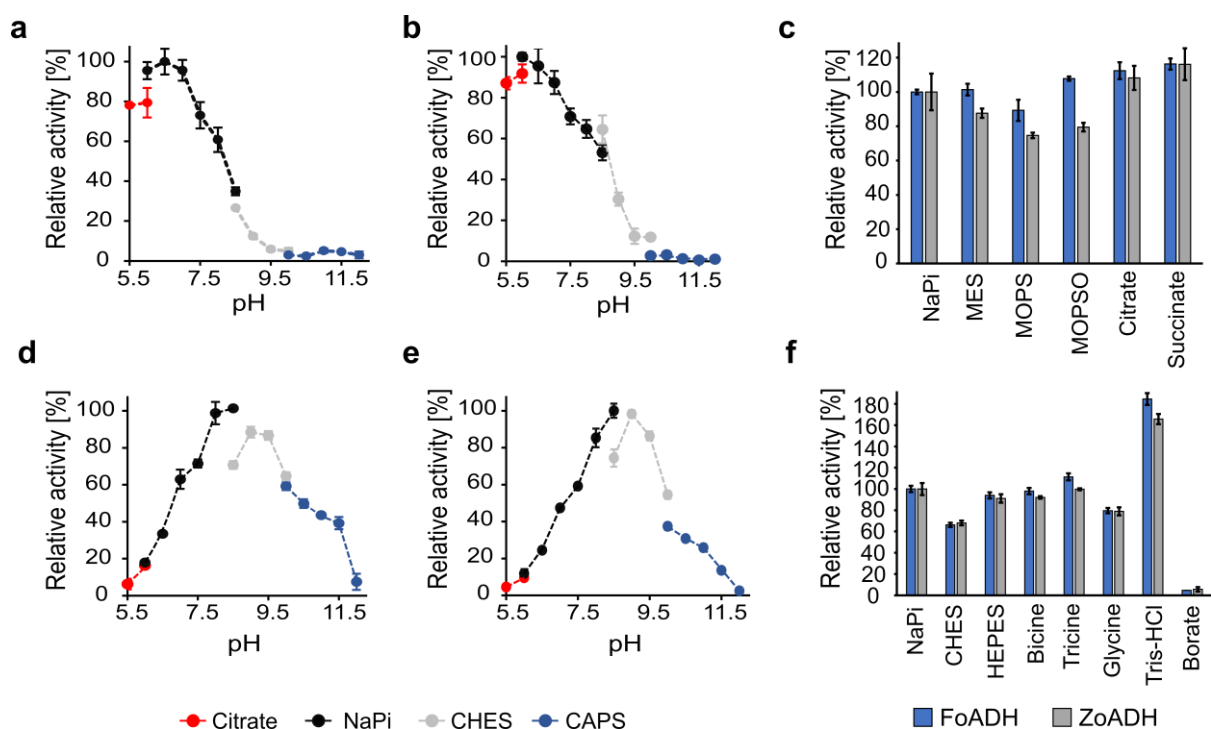
1050

1051 **Fig. 2** Knockout of the ADH gene in *Z. galactanivorans* leads to impaired growth on G6Me.
1052 Different *Z. galactanivorans* strains (wild type (WT), gene knockout ADH (ΔADH), and gene
1053 knockout CYP (ΔCYP)) were incubated in minimal medium amended with D-galactose or
1054 G6Me for 3 days at RT.



1055

1056 **Fig. 3** Thiol-dependent detoxification of formaldehyde catalyzed by an ADH and an esterase.
 1057 **a)** Principle of thiol-dependent detoxification of formaldehyde and **b)** investigated thiols.

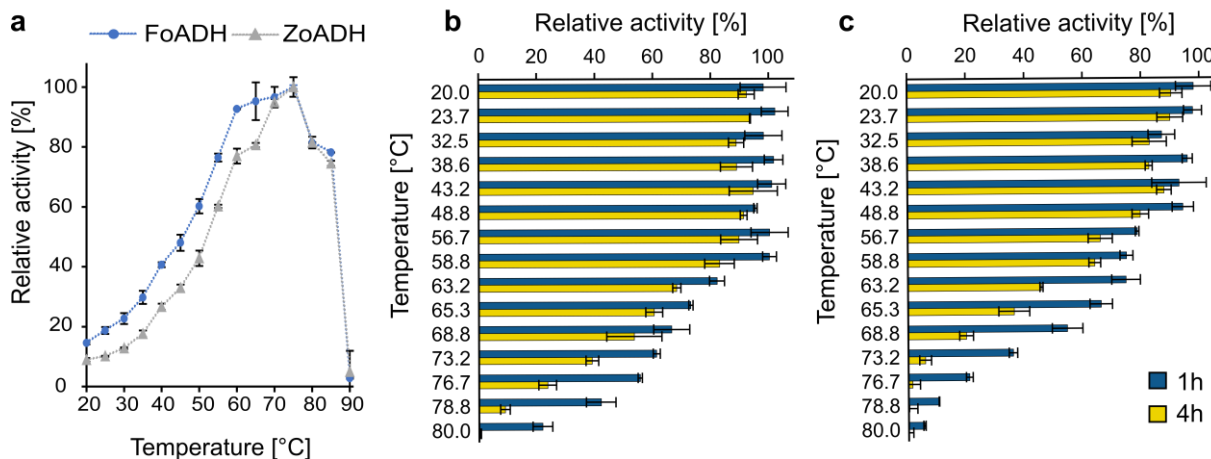


1058

● Citrate ● NaPi ● CHES ● CAPS

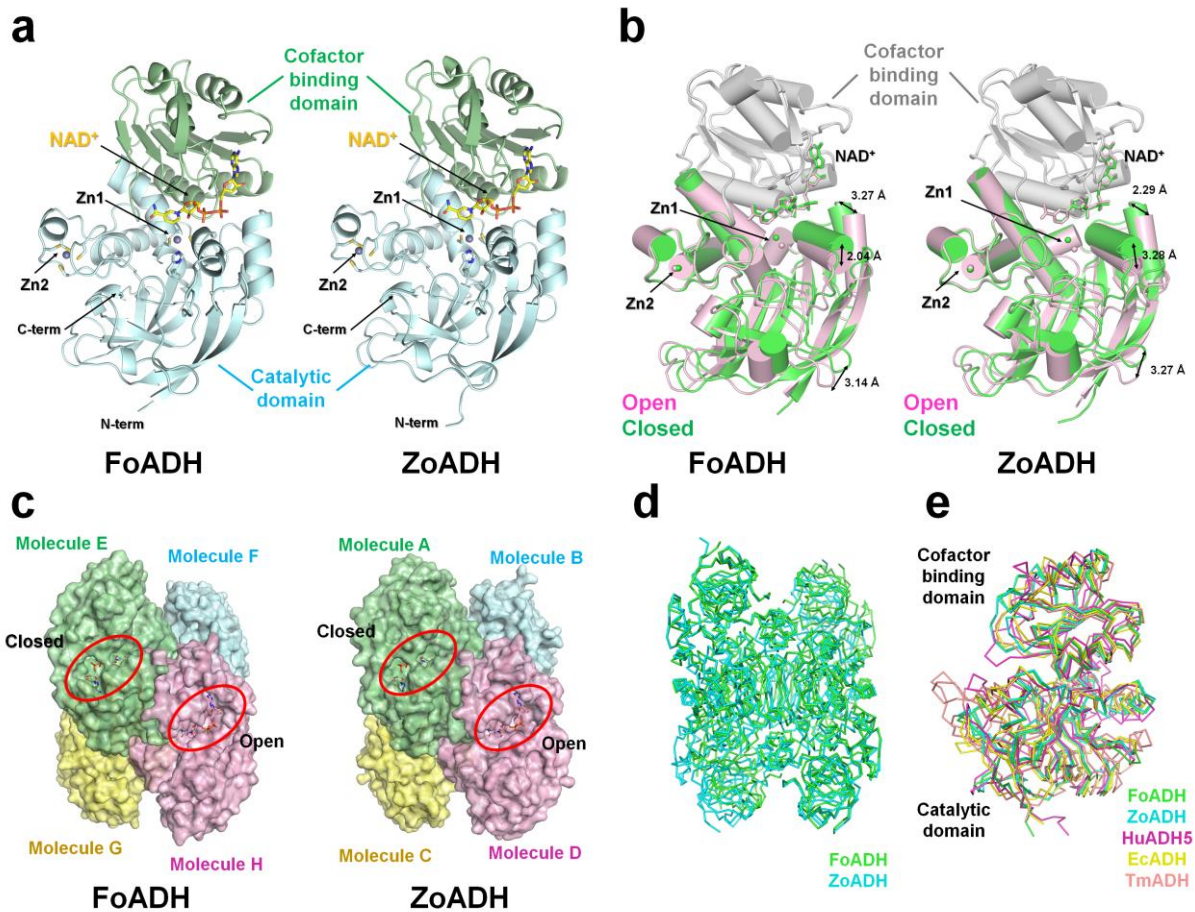
■ FoADH ■ ZoADH

1059 **Fig. 4** Influence of pH and buffer components on the ADH activity. pH optimum for the
 1060 reduction reaction of **a)** FoADH and **b)** ZoADH as well as the pH optimum for the oxidation
 1061 reaction catalyzed by **d)** FoADH and **e)** ZoADH. **c)** Reduction of benzaldehyde and **f)** oxidation
 1062 of benzyl alcohol by the ADHs at the respective pH optima using various buffers. A pH of 6.5
 1063 was employed for the reduction reaction and a pH of 8.5 for the oxidation reaction; all buffers
 1064 had a concentration of 50 mM. Since some buffers including Bicine, Tricine, Tris, MOPSO and
 1065 HEPES contain hydroxyl groups, a falsified activity due to turnover of these substances was
 1066 excluded by a measurement without additional substrate. However, no activity was observed
 1067 for any buffer component. All measurements (**a-f**) were performed under following conditions:
 1068 a final substrate concentration of 10 mM benzyl alcohol or benzaldehyde, 3.5% (v/v) DMSO
 1069 and 0.5 mM NAD⁺ or NADH was used. The reaction was started by the addition of ADH at a
 1070 final enzyme concentration of 0.1 mg mL⁻¹. The measurement was performed at 25 °C in the
 1071 respective buffers with concentrations of 50 mM. The maximum relative activity (100%)
 1072 corresponds to the measurements in the 50 mM NaPi buffers pH 6.5 for reduction and pH 8.5
 1073 for oxidation reactions. All measurements were performed as triplicates, the mean is given and
 1074 the error bars indicate the standard deviation.



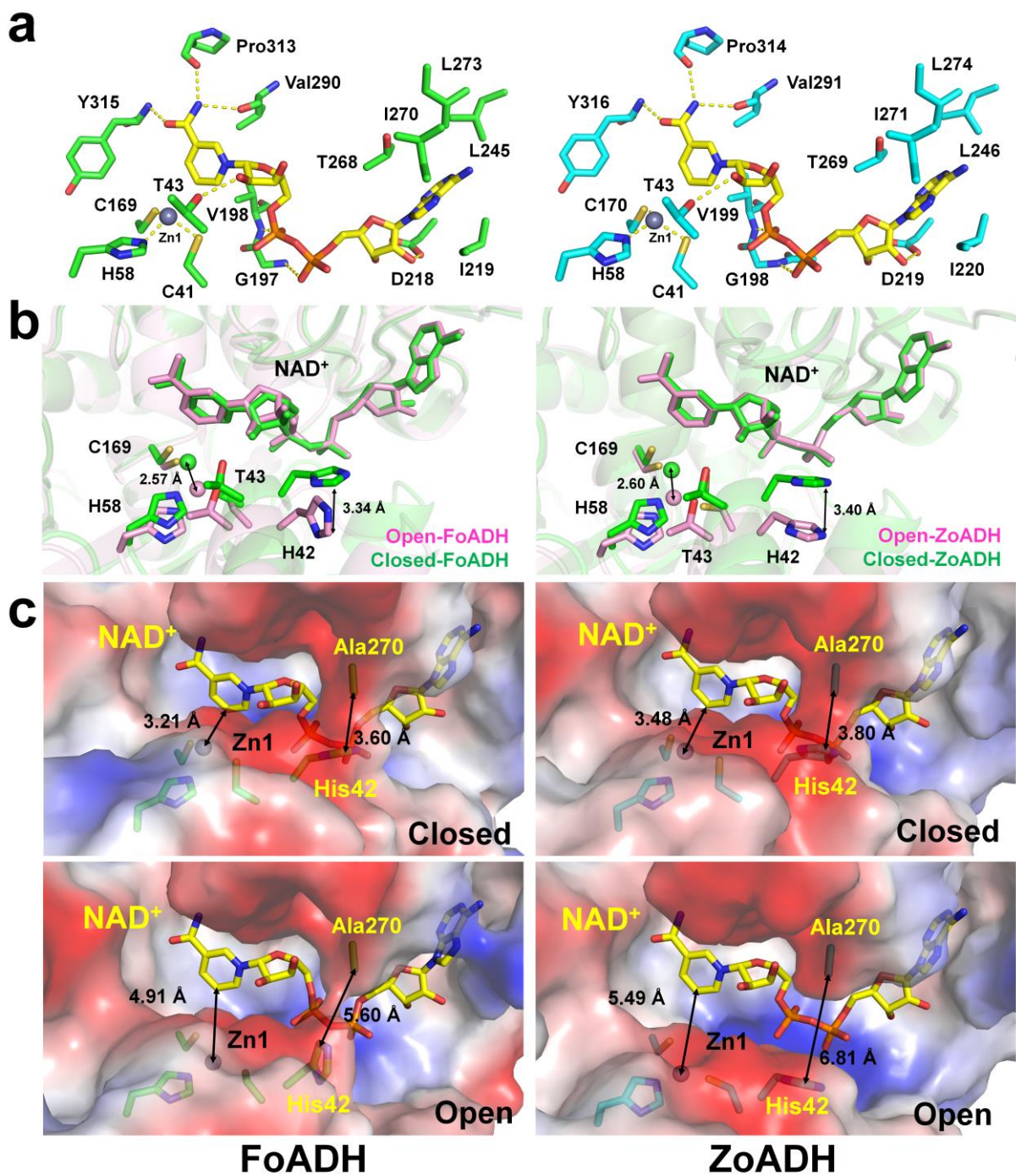
1075

1076 **Fig. 5** Temperature optimum and thermostability of the ADHs. **a)** Influence of temperature on
 1077 enzyme activity. The measurement was performed at various temperatures ranging from 20 to
 1078 80 °C for 10 min. The maximum relative activity (100%) corresponds to the measurement at
 1079 75 °C for both enzymes. Influence of temperature on enzyme stability for **b)** FoADH and **c)**
 1080 ZoADH. The enzymes with a concentration of 1 mg mL⁻¹ were incubated at different
 1081 temperatures between 20 and 80 °C for 1 or 4 hours, followed by the determination of residual
 1082 activity. The measurement was performed at 40 °C. The maximum relative activity (100%)
 1083 corresponds to a control incubated on ice for 1 or 4 h. All measurements (**a-c**) were performed
 1084 under following conditions: a final substrate concentration of 10 mM benzyl alcohol and 0.5
 1085 mM NAD⁺ was used. The reaction was started by the addition of ADH at a final enzyme
 1086 concentration of 0.1 mg mL⁻¹. The measurements were performed in a 50 mM NaPi buffer
 1087 pH 7.5. All measurements were performed as triplicates, the mean is given and the error bars
 1088 indicate the standard deviation.



1089

1090 **Fig. 6** Crystal structures of FoADH and ZoADH. **a)** Monomer structures of ZoADH and FoADH. Catalytic and cofactor domain are indicated by cyan and green, respectively. NAD⁺ and zinc ions are indicated by yellow stick and grey sphere, respectively. **b)** Superimposition of closed (green) and open conformation between catalytic and cofactor-binding domains of ZoADH and FoADH monomers. The superimposed cofactor-binding domain of ZoADH and FoADH are indicated as grey cartoon. **c)** Tetrameric formation of ZoADH and FoADH. **d)** Superimposition of tetrameric formation of FoADH (green) and ZoADH (cyan). **e)** Superimposition of monomer structure of FoADH (green) and ZoADH (cyan) with all-trans-retinol dehydrogenase ADH4 from *Homo sapiens* (pink, PDB code: 3COS), uncharacterized zinc-type alcohol dehydrogenase-like protein YdjJ from *E. coli* (wheat, 5vm2), scyllo-inosose 3-dehydrogenase from *Thermotoga maritima* (3IP1, yellow).

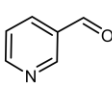

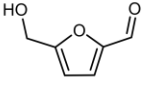

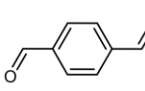
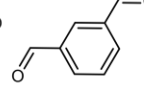
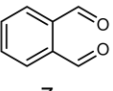
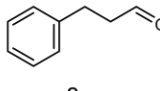
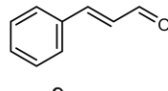
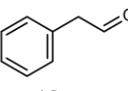
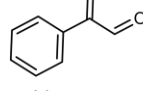
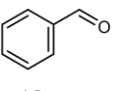
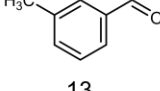
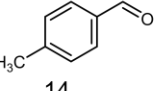
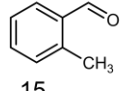
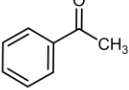
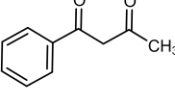

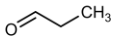
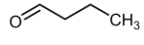
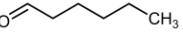


1101

1102 **Fig. 7** Active sites of FoADH and ZoADH. **a)** Interaction of ZoADH and FoADH with NAD⁺ and
 1103 zinc ion at the Zn1 site. **b)** Superimposition of active site of open and closed conformation of
 1104 FoADH and ZoADH. **c)** Comparison of electrostatic surface structure of open and closed
 1105 conformations of FoADH and ZoADH

1106 **Tables**

1107 **Table 1** Initial substrate screening of the ADH in the reduction direction revealed that they
 1108 preferentially convert aromatic aldehydes. Substrates were employed at a final concentration
 1109 of 10 mM. For NADH a concentration of 0.5 mM was used. The reaction contained 3.5% (v/v)
 1110 DMSO. The reaction was conducted in a 50 mM succinate buffer pH 6.5 at an incubation
 1111 temperature of 70 °C. All measurements were performed as triplicates, the mean and the
 1112 standard deviation is given.

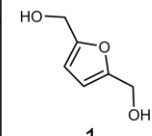
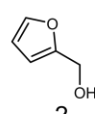
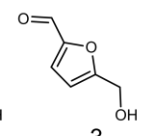
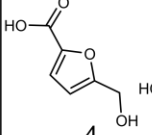
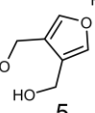
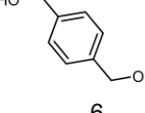
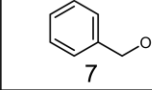
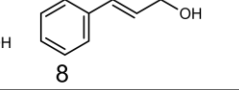
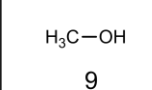
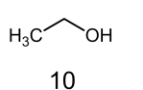

Abbr.	Substrates	Specific activity [U/mg]		
		FoADH	ZoADH	
1	Pyridine-3-carbaldehyde	64.09 ± 2.39	54.85 ± 4.34	
2	Furfural	47.77 ± 1.19	44.78 ± 2.07	
3	5-(Hydroxymethyl)furfural	44.81 ± 2.16	38.29 ± 2.47	
4	Thiophene-3-carbaldehyde	37.45 ± 4.19	29.32 ± 2.59	
5	Terephthalaldehyde	30.05 ± 3.60	27.34 ± 4.68	
6	Isophthalaldehyde	26.71 ± 1.47	36.94 ± 3.68	
7	Phthalaldehyde	n.d. ^{a)}	n.d.	
8	Hydrocinnamaldehyde	26.00 ± 2.23	30.13 ± 2.32	
9	Cinnamaldehyde ^{b)}	12.57 ± 0.90	11.20 ± 0.11	
10	Phenylacetaldehyde	n.d.	n.d.	
11	Phenylglyoxal	n.d.	n.d.	
12	Benzaldehyde	5.14 ± 0.09	12.49 ± 0.73	
13	<i>m</i> -Tolualdehyde	6.30 ± 1.06	7.32 ± 0.32	
14	<i>p</i> -Tolualdehyde	4.71 ± 0.32	5.60 ± 0.44	
15	<i>o</i> -Tolualdehyde	0.04 ± 0.01	0.04 ± 0.01	
16	Acetophenone	n.d.	n.d.	
17	Benzoylacetone	n.d.	n.d.	
18	Formaldehyde	n.d.	n.d.	
19	Propionaldehyde	1.58 ± 0.10	0.92 ± 0.36	
20	Butyraldehyde	3.62 ± 1.04	4.02 ± 0.48	
21	Caproaldehyde	7.07 ± 0.64	6.62 ± 0.30	

a) n.d.: not detected

b) Due to a high background absorption of the component, a substrate concentration of 1mM was employed.

1113

1114 **Table 2** Both ADHs possess minor specific activities for the oxidation of alcohols.
 1115 Formaldehyde was also tested in a possible oxidation reaction to exclude thiol-independent
 1116 formaldehyde dehydrogenase activity. Substrates were employed at a final concentration of
 1117 10 mM. For NAD⁺ a concentration of 0.5 mM was used. The reaction contained 3.5% (v/v)
 1118 DMSO. The reaction was conducted in a 50 mM NaPi buffer pH 8.5 at an incubation
 1119 temperature of 70 °C. All measurements were performed as triplicates, the mean and the
 1120 standard deviation is given.

Abbr.	Substrate	Specific activity [mU/mg]		
		FoADH	ZoADH	
1	2,5-Bis(hydroxymethyl)furan	490 ± 50	290 ± 90	
2	Furfuryl alcohol	310 ± 10	194 ± 51	
3	5-(Hydroxymethyl)furfural	n.d. ^{a)}	n.d.	
4	5-(Hydroxymethyl)furan-2-carboxylic acid	n.d.	n.d.	
5	3,4-Bis(hydroxymethyl)furan	n.d.	n.d.	
6	1,4-Benzenedimethanol	86 ± 1.5	51 ± 3.0	
7	Benzyl alcohol	81 ± 0.7	53 ± 0.1	
8	Cinnamyl alcohol ^{b)}	11 ± 3.1	6 ± 2.9	
9	Methanol	n.d.	n.d.	
10	Ethanol	n.d.	n.d.	
11	Formaldehyde	n.d.	n.d.	

a) n.d.: not detected

1121 b) Due to a high background absorption of the component, a substrate concentration of 1mM was employed.

1122 **Table 3** Influence of various substances on the enzyme activity of both ADHs. The ADH was
 1123 incubated with the respective component for 1 h at RT prior to measurement. The maximum
 1124 relative activity (100%) corresponds to the measurement for the control, which contained no
 1125 additives. All measurements were performed under following conditions: a final substrate
 1126 concentration of 10 mM benzyl alcohol and 0.5 mM NAD⁺ was used. The reaction was started
 1127 by the addition of ADH at a final enzyme concentration of 0.1 mg mL⁻¹. The measurements
 1128 were performed in a 50 mM HEPES buffer pH 8.5 at 25 °C. All measurements were performed
 1129 as triplicates, the mean and the standard deviation is given.

Chemical	Conc. [mM]	Relative activity [%]	
		FoADH	ZoADH
None	-	100 ± 1.3	100 ± 9.3
KCl	10	121 ± 1.9	154 ± 7.9
CaCl ₂	10	188 ± 6.3	306 ± 9.9
MgCl ₂	10	219 ± 7.2	328 ± 11.7
NiCl ₂	10	891 ± 13.8	1004 ± 6.6
CoCl ₂	10	1280 ± 38.1	1242 ± 25.4
MnCl ₂	10	1394 ± 58.5	973 ± 45.0
ZnCl ₂	10	n.d. ^{a)}	n.d.
CuCl ₂	10	n.d.	n.d.
FeCl ₃	10	n.d.	n.d.
EDTA	25	61 ± 1.9	61 ± 7.8
DTT	10	48 ± 4.6	85 ± 3.9
2-ME	10	19 ± 1.7	28 ± 4.1

1130

a) n.d.: not detected

List of publications

- **Marine polysaccharides: Occurrence, enzymatic degradation and utilization**
M. Bäumgen*, T. Dutschei* and U. T. Bornscheuer, *ChemBioChem* **2021**, 22, 2247-2256.
- **A new carbohydrate-active oligosaccharide dehydratase is involved in the degradation of ulvan**
M. Bäumgen*, T. Dutschei*, D. Bartosik, C. Suster, L. Reisky, N. Gerlach, C. Stanetty, M. D. Mihovilovic, T. Schweder, J. H. Hehemann and U. T. Bornscheuer, *J. Chem. Biol.* **2021**, 297, 101210
- **Metabolic engineering enables *Bacillus licheniformis* to grow on the marine polysaccharide ulvan**
T. Dutschei, M. K. Zühlke, N. Welsch, T. Eisenack, M. Hilkmann, J. Krull, C. Stühle, S. Brott, A. Dürwald, L. Reisky, J.-H. Hehemann, D. Becher, T. Schweder and U. T. Bornscheuer, *Microb. Cell Fact.* **2022**, 21, 207 (under revision)
- **Utilization of a diverse range of xylan structures from marine *Bacteroidetes***
T. Dutschei, I. Beidler, D. Bartosik, J.-M. Seeßelberg, M. Teune, M. Bäumgen, S. Q. Ferreira, J. Heldmann, F. Nagel, J. Krull, L. Berndt, K. Methling, M. Hein, D. Becher, P. Langer, M. Delcea, M. Lalk, M. Lammers, M. Höhne, J. H. Hehemann, T. Schweder and U. T. Bornscheuer, *Environ. Microbiol.* **2022** (submitted)
- **Connecting algal polysaccharide degradation to formaldehyde detoxification**
S. Brott, F. Thomas, M. Behrens, K. Methling, D. Bartosik, T. Dutschei, M. Lalk, G. Michel, T. Schweder, and U. T. Bornscheuer, *ChemBioChem* **2022** 23, e202200269
- **Unique alcohol dehydrogenases involved in algal sugar utilization by marine bacteria**
S. Brott, K. H. Nam, F. Thomas, T. Dutschei, L. Reisky, M. Behrens, H. C. Grimm, G. Michel, T. Schweder, and U. T. Bornscheuer, *Appl. Microbiol. Biotechnol.* **2022** (submitted)

Acknowledgments

At first, I wish to thank Uwe Bornscheuer for the nice time, support and great opportunity to work in his team on this topic. The long time I worked for you let me develop as much as a scientist as personally. I was always honoured to work for you.

Additionally, I wish to give a special thanks to Thomas Schweder for coordinating the POMPU consortium. The consortium enabled interdisciplinary work, which kept on inspiring me in my scientific development and led to the resulting complex studies.

I also want to thank the German Research foundation (DFG) for the funding of the POMPU FOR2406 and acknowledging the consortium.

Good work means teamwork, so I thank all members of the “sugar team” at the University of Greifswald in the Institute of Biochemistry, the POMPU-consortium and all associated partners.

I wish to thank Lukas and Marcus for teaching me the necessary skills to guide me through my studies and the inspiring discussions. Thanks to Stefan for always accepting me as I am and the great co-work. Thanks to Michelle for the nice overlap time and that you are the enthusiast scientist as you are. Thanks to Jana for the nice time and trips which led to explore a different angle of science and life. Additionally, I want to thank my bachelor students Carlo and Jonas for their great support. Thanks as well to Soraia, Julia and Meike for their contribution with their works.

In addition, I give special thanks to the members of the AG Schweder: Marie, Daniel, Irena, Norma, Max and Alex, for the fruitful cooperation. I enjoyed working with you, the interesting talks and evenings. This broadens up my view to look beyond.

I want to thank Felix for the biophysical and bioanalytical discussions, and always be there for a talk.

Many thanks to the exciting exchange and discussions with Joris, Silvia, Nicola, Jaagni, Jan Hendrik from Bremen, Christian and Christoph from Vienna and Francois and Gurvan from France.

The right working environment can impact who you become, so I am thankful to work with all members of the AG Bornscheuer, especially Ina, Dominique, Thomas, Hannes, Aileen, Henrik T, Henrik M, Chao, Eva, Lara, Celine, Lina, Benji, Yannik, Amanda and Katja.

Besides, I want to thank my family and friends to support me on that journey.