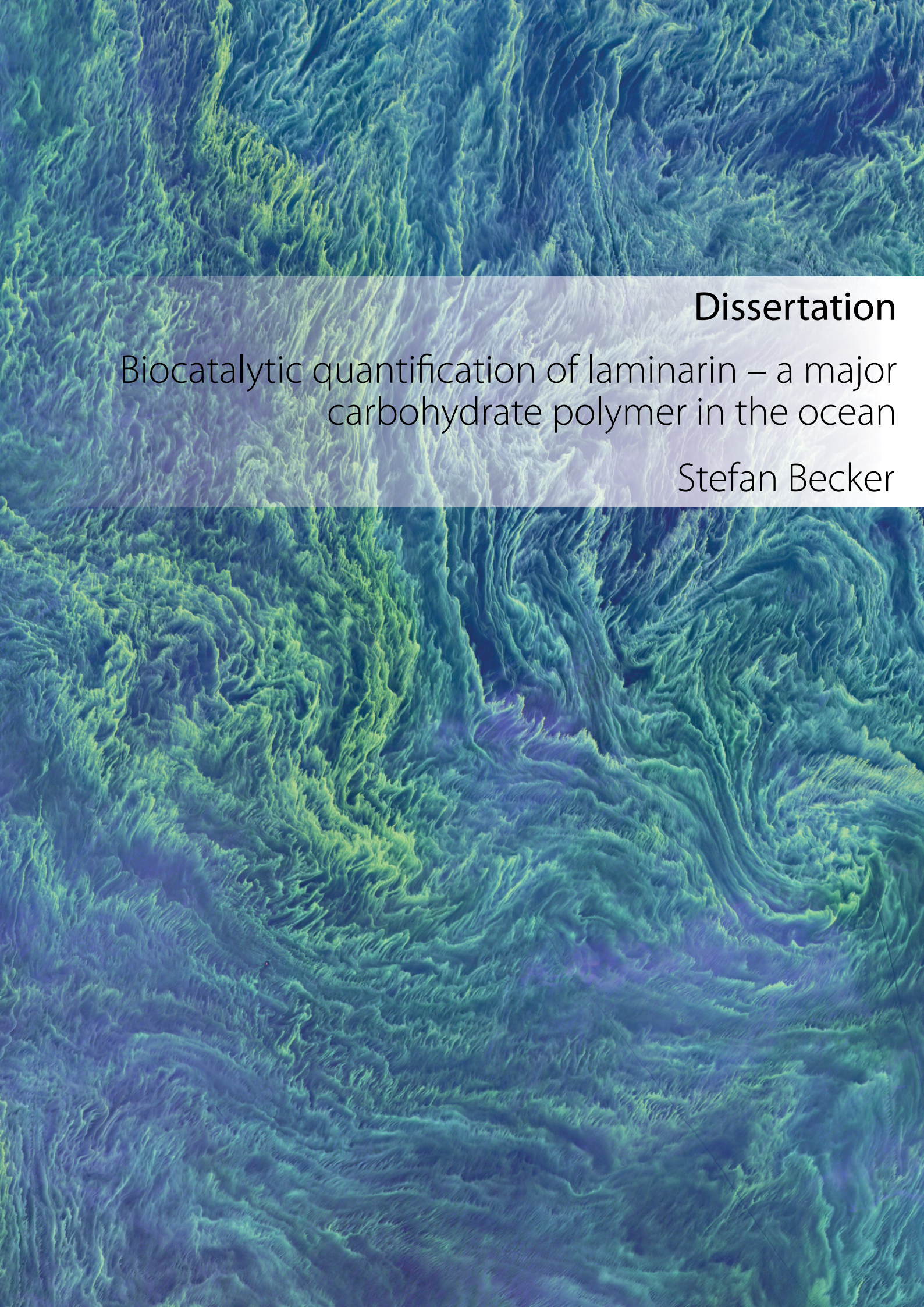


Dissertation

Biocatalytic quantification of laminarin – a major carbohydrate polymer in the ocean

Stefan Becker

Biocatalytic quantification of laminarin – a major carbohydrate polymer in the ocean

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften

– Dr. rer. nat. –

im Fachbereich Geowissenschaften der Universität Bremen

Stefan Becker

Bremen

März 2018

The cover image was taken by the Operational Land Imager (OLI) on the Landsat 8 satellite on August 11, 2015. It depicts a false-color view of a large phytoplankton bloom in the Southern Baltic Sea. The picture was rotated by 90°, so that the geographical North is now on the right side. Image courtesy of the NASA Earth Observatory.

Die vorliegende Arbeit wurde in der Zeit von April 2015 bis März 2018 in der MARUM-MPI-Brückengruppe für Marine Glykobiologie am MARUM – Zentrum für Marine Umweltwissenschaften der Universität Bremen und dem Max-Planck-Institut für Marine Mikrobiologie erarbeitet.

Autor: **Stefan Becker, M.Sc.**
MARUM – Zentrum für Marine Umweltwissenschaften
Universität Bremen
Max-Planck-Institut für Marine Mikrobiologie

Datum des Promotionskolloquiums: 22.06.2018

Gutachter: **Dr. Jan-Hendrik Hehemann**
MARUM – Zentrum für Marine Umweltwissenschaften
Universität Bremen
Max-Planck-Institut für Marine Mikrobiologie

Prof. Dr. Carol Arnosti
University of North Carolina at Chapel Hill

“Oooh, look at me, I looked up a quote!”

- Randall Munroe, xkcd.com

SUMMARY

In this work, I provide evidence that laminarin, the energy storage glucan of marine diatoms and many other algae, is a central bioenergy molecule of the ocean. This conclusion stems from the quantification of this molecule with new technology that I developed during my thesis. This biocatalytic assay is based on specific enzymes from marine microbes. This first quantification yielded an average contribution of laminarin to the carbon in algae derived organic matter of 37±19% in the environment.

The work contributes to our general understanding of the marine carbon cycle in the surface water of the ocean. In this environment, microalgae sustain approximately half of the global primary production and yield significant amounts of organic carbon in the form of polysaccharides, e.g. laminarin. So far, the role of this class of biological macromolecules in the carbon cycle is poorly understood due to the technological challenges in their analysis. The quantification of a single marine polysaccharide on a broad scale has therefore never been done. Our new approach starts to close this gap of knowledge and technology. It makes use of the enzymatic toolkit from marine microbes, that evolved specific enzymes in order to gain energy and carbon from the abundant carbohydrates. I demonstrate the advantages and the further potential of these enzymes for a faster, stereo- and sequence-specific analysis of selected polysaccharides in marine organic matter. In order to promote its further application, the method is now accessible for researchers in all fields of environmental science and can be easily applied to quantify laminarin in particulate organic matter from the marine environment.

The characterization of the microbial machinery for the degradation of laminarin was required for the method development. However, it also contributed to the in-depth investigations that were made on two strains of specialized laminarin degrading bacteria. Environmental proteomics and metagenomics in combination with cultivation experiments, uptake visualization and our biochemical and crystallographic characterizations revealed streamlined bacteria that reach high abundances during repeating algal bloom events but also dominate the laminarin turnover in a highly competitive manner.

Finally, I applied our new laminarin quantification method to environmental samples. I used particulate organic matter from two environmental time series, in different size fractions, vertical profiles, marine snow particles and a meridional transect in oceanic regions ranging from the Arctic, the North-, Central- and South-Atlantic, the coastal Pacific and the North Sea. Our measurements allow a more accurate estimation of the global annual laminarin production of 18±9 gigatons. The varying levels of laminarin indicated different bioenergetic states of the

oceanic regions and the ecological relevance of this molecule was highlighted by its sheer abundance of more than one third of the particulate organic carbon in surface waters.

ZUSAMMENFASSUNG

In dieser Arbeit zeige ich, dass Laminarin, das Glucan, welches zur Energiespeicherung in marinen Diatomeen und vielen anderen Algen dient, ein zentrales Bioenergiemolekül im Ozean ist. Diese Schlussfolgerung folgt aus der Quantifizierung dieses Moleküls mit Hilfe einer neuen biokatalytischen Methode, die ich während meiner Doktorarbeit entwickelt habe und welche auf spezifischen Enzymen aus marinen Mikroorganismen basiert. Die erste Quantifizierung ergab in Umweltproben einen durchschnittlichen Anteil von $37 \pm 19\%$ Laminarin im Verhältnis zum Kohlenstoff in der organischen Materie von Algen.

Die Arbeit trägt zu unserem allgemeinen Verständnis des marinen Kohlenstoffkreislaufs im Oberflächenwasser des Ozeans bei. In diesem Lebensraum erbringen Mikroalgen etwa die Hälfte der globalen Primärproduktion und erzeugen signifikante Mengen an organischem Kohlenstoff in Form von Polysacchariden, wie zum Beispiel Laminarin. Bis jetzt ist die Rolle dieser Klasse von biologischen Makromolekülen im Kohlenstoffkreislauf aufgrund der analytischen Herausforderungen kaum verstanden. Die Quantifizierung eines einzelnen marinen Polysaccharids wurde daher nie in einem größeren Maßstab durchgeführt. Der neue Ansatz beginnt diese Wissenslücke und das technologische Defizit zu schließen. Er nutzt die enzymatischen Werkzeuge mariner Mikroorganismen, welche spezifische Enzyme evolvierten, um Energie und Kohlenstoff aus abundanten Kohlenhydraten zu gewinnen. Ich demonstriere die Vorteile und das zukünftige Potential dieser Enzyme für eine schnellere, stereo- und sequenzspezifische Analyse ausgewählter Polysaccharide in mariner organischer Materie. Um die weitere Anwendung der Methode voran zu treiben, ist sie von nun an auch für Forschende auf anderen Gebieten der Umweltwissenschaften zugänglich und kann mühelos zur Quantifizierung von Laminarin in partikulärer organischer Materie mariner Umweltproben angewendet werden.

Die Charakterisierung der mikrobiellen Maschinerie für den Abbau von Laminarin war nicht nur für die Methodenentwicklung notwendig, sondern trug auch zu den genaueren Untersuchungen bei zweier Bakterienstämme bei, die sich auf den Laminarinabbau spezialisiert haben. Umweltproteomik und Metagenomik in Kombination mit Kultivierungsexperimenten, der Visualisierung von Substrataufnahme sowie unsere biochemischen und strukturellen Charakterisierungen zeigten hoch spezialisierte Bakterien. Diese erreichten während mehrerer Algenblüten sowohl hohe Abundanzen erreichen, als auch den Laminarinumsatz in einer sehr kompetitiven Weise dominieren.

Abschließend habe ich unsere neue Laminarin-Quantifizierungsmethode auf Umweltproben angewendet. Bei diesen handelte es ich um partikuläres organisches Material aus zwei

Langzeituntersuchungen und in verschiedenen Größenfraktionen, aus vertikalen Profilen, Proben von marinem Schnee und einem meridionalen Transekt. Die Herkunft der Proben reichte von der Arktis, dem Nord-, Mittel- und Südatlantik und pazifischen Küstengewässern bis hin zur Nordsee. Meine Messungen erlauben eine genauere Schätzung der globalen jährlichen Laminarinproduktion auf 18 ± 9 Gigatonnen. Die unterschiedlichen Laminarinwerte weisen auf unterschiedliche bioenergetische Niveaus der verschiedenen Meeresregionen hin. Die ökologische Relevanz dieses Moleküls wurde durch die absolute Abundanz von mehr als einem Drittel des partikulären organischen Kohlenstoffs in Oberflächengewässern verdeutlicht.

CONTENTS

SUMMARY VII

ZUSAMMENFASSUNG IX

ABBREVIATIONS XIII

1 Introduction 17

1.1 The marine carbon cycle in surface waters 17
1.2 Dynamics of phytoplankton primary production and the role of marine polysaccharides 19
1.3 The algal polysaccharide laminarin 21
1.4 Prokaryotic polysaccharide degradation in the microbial loop 22
1.5 Glycoside hydrolases and mechanisms of glycan degradation 24
1.6 The analytical technologies of marine glycobiology 26
1.7 Aim and outline of the thesis 28

2 Manuscripts 31

2.1 Manuscript I: Accurate quantification of laminarin in marine organic matter with enzymes from marine microbes 33
2.2 Manuscript II: Adaptive mechanisms that provide competitive advantages to marine Bacteroidetes during microalgal blooms 57
2.3 Manuscript III: Laminarin quantification in microalgae with enzymes from marine microbes 97
2.4 Manuscript IV: Laminarin is the major marine sugar polymer and indicates bioenergy states of the surface ocean 113

3 Conclusions, general discussion and outlook 131

3.1 Further method development and possible new applications 131
3.2 Specialized laminarin degraders and prospective proteinbiochemical analyses 134
3.3 Implications of the sweet ocean for future studies 135

BIBLIOGRAPHY 139

ACKNOWLEDGEMENTS 157

APPENDIX 159

Additional co-author contributions 159

ABBREVIATIONS

ABC	ATP binding cassette
AIC	Akaike's information criterion
AMT	Atlantic meridional transect
ANI	average nucleotide identity
AT	ammonium transporter
ATP	adenosine triphosphate
BIC	Bayesian information criterion
BLAST	basic local alignment search tool
BSA	bovine serum albumin
CAZyme	carbohydrate active enzyme
CBM	carbohydrate binding module
CE	carbohydrate esterase
Chl	chlorophyll
CV	column volume
DAPI	4′,6-Diamidin-2-phenylindol
DB	degree of branching
DIC	dissolved inorganic matter
DLS	dynamic light scattering
DOC	dissolved organic carbon
DOM	dissolved organic matter
DP	degree of polymerization
EC no.	enzyme commission number
EEA	extracellular enzymatic activity
EPS	extracellular polymeric substances
FACE	fluorophore-assisted carbohydrate polyacrylamide gel electrophoresis
FDR	false discovery rate
FLA	fluoresceinamine
FPLC	fast protein liquid chromatography
GF/F	glass fiber filter, grade F
GH	glycoside hydrolase
Glc	glucose
GS	glutamate synthase

GT	glycosyl transferase
HMW	high molecular weight
HP	hypothetical protein
HPAEC-PAD	high-performance anion exchange chromatography with pulsed amperometric detection
HPLC	high-performance liquid chromatography
iBAQ	intensity-based absolute quantification
IMAC	immobilized metal affinity chromatography
Lam	laminarin
LB	Luria-Bertani
LMW	low molecular weight
LOD	limit of detection
LTER	long term ecological research site
LUL	laminarin utilization loci
MFS	major facilitator superfamily
MLG	mixed-linked glucan
MOPS	3-(N-morpholino)propanesulfonic acid
MS	mass spectrometry
NMR	nuclear magnetic resonance
NSAF	normalized spectral abundance factor
PC	polycarbonate
PAGE	polyacrylamide gel electrophoresis
PAHBAH	p-hydroxybenzoic acid hydrazide
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PKD	polycystic kidney disease
PL	polysaccharide lyases
PEG	polyethylene glycol
POC	particulate organic carbon
POM	particulate organic matter
PUL	polysaccharide utilization loci
RDOC	recalcitrant dissolved organic carbon
RMSD	root-mean-square deviation
RT	room temperature
SEC	size exclusion chromatography
SD	standard deviation
SDS	sodium dodecyl sulfate

Sus	starch utilization system
TAG	triacylglycerides
TBDR	TonB-dependent receptors
TLC	thin layer chromatography
TonB	outer membrane transporter
Tris	Tris(hydroxymethyl)aminomethane

1 Introduction

1.1 The marine carbon cycle in surface waters

Investigating the marine carbon cycle in general and the cycling in surface water in specific, is highly relevant. Approximately two-thirds of the Earth's surface is covered by water, and marine algae contribute about half of the global primary production (Field et al., 1998). They convert more carbon dioxide into biomass than tropical forests. Even though the total amount of carbon in the marine system is lower than, e.g. the carbon reservoir of the entire lithosphere, it is much more reactive, meaning that the carbon turnover is much higher in this environment. On longer time scales, the carbon cycling in the marine system affects the Earth's entire climate (Ciais et al., 2013).

The global carbon cycle is a biogeochemical cycle that can be described as carbon pools in various different environments, which are interconnected by fluxes. In the marine environment, several carbon pools can be discriminated from each other. A division is between dissolved and particulate organic carbon (DOC and POC), which are both part of the dissolved and particulate organic matter (DOM and POM) pools (FIG 1.1). Carbon fluxes can be driven by the physical and the biological carbon pump and they can be described as rates of transfer of masses or volumes (Legendre et al., 2015). The difference in the partial pressure of CO_2 in water and air drives the physical carbon pump, and together with additional factors like temperature, atmospheric CO_2 dissolves in seawater and reacts to molecules referred to as dissolved inorganic carbon (DIC). The biological pump is driven by food webs in the surface oceans and it results in the sinking flux of POC (Turner, 2015). Similar to the physical carbon pump, it sequesters carbon from the atmosphere to the deep sea by the physical sinking of aggregated POM. This marine snow consists of, e.g. phytodetritus, extracellular polymeric substances (EPS) and fecal pellets, and it can export up to 3% of the annual primary production from the surface to the deep sea (Legendre et al., 2015; Ducklow et al., 2001). Independent from the biological carbon pump another concept of carbon export was introduced (Jiao et al., 2010). The sequestration of recalcitrant dissolved organic carbon (RDOC) by prokaryotes is further specified as the microbial carbon pump (Jiao and Zheng, 2011). The mechanism describes the formation and export of diverse molecules, which are recalcitrant against rapid degradation, into the deep sea (FIG 1.1). The POC and RDOC, which were

exported into the deep sea can persist in the ocean for thousands of years (Osterholz et al., 2015; Middelburg and Meysman, 2007; Sabine, 2004). These high transit times, in which no exchange with the atmosphere occurs, are due to the thermohaline circulation (Wunsch, 2002). Cold and dense surface waters are sinking and forming a deep water body that stores via this mechanism not only POC and RDOC but also DIC. On the other hand, the sunlit (euphotic) water body reaches only a depth of ~200 m and is characterized by high carbon turnover. High nutrient (eutrophic) conditions are mostly found in coastal and upwelling areas, whereas other regions are more often characterized by low nutrient (oligotrophic) conditions (Gattuso et al., 1998).

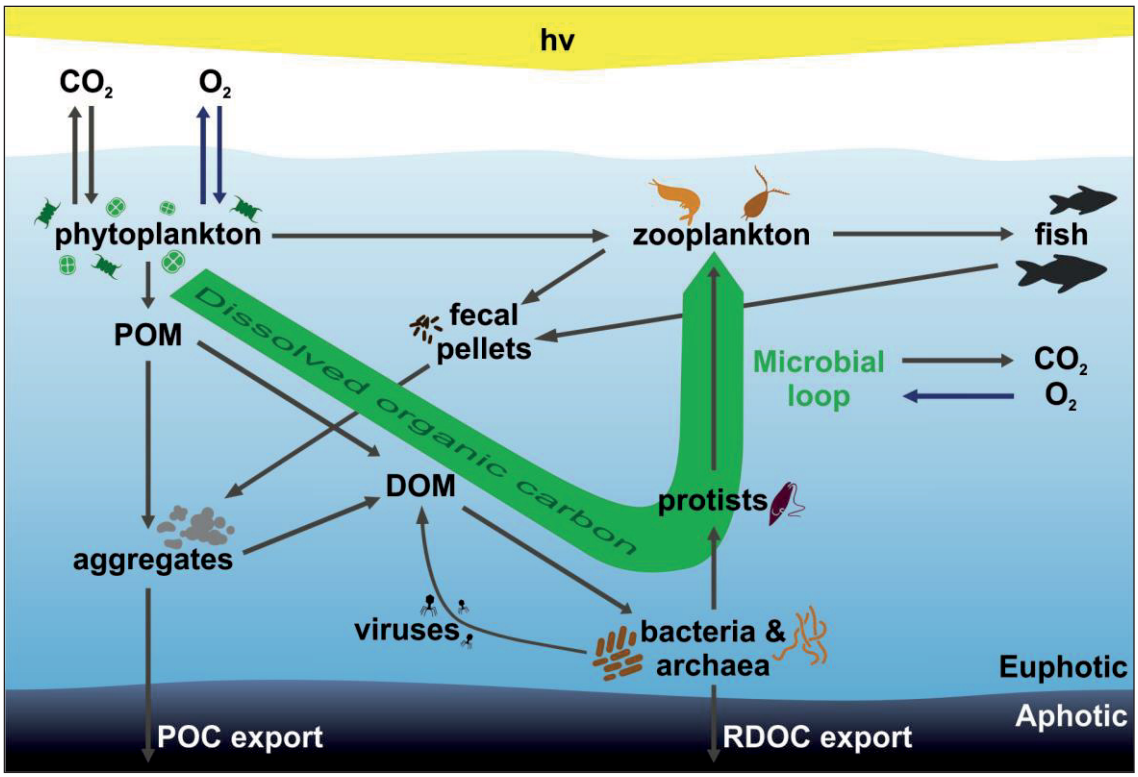


FIG 1.1 A simplified marine carbon cycle in the surface waters of the ocean, adapted from (Worden et al., 2015). Grey and blue arrows indicate carbon and oxygen fluxes. The green arrow depicts the microbial loop, where phytoplankton derived dissolved organic carbon is being respired by heterotroph prokaryotes and channeled back into the classical linear food chain via protist and zooplankton grazing.

Eukaryotic microalgae are an integral part of the phytoplankton. They use photosynthesis to carry out primary production, in which carbon dioxide is converted into various organic compounds and oxygen. The most immediate product of primary production is glucose, which can be then further transformed. Marine algae, such as diatoms and dinoflagellates and also photosynthetic cyanobacteria can cause phytoplankton bloom events, where the biomass accumulates and photosynthesis rates are high. This photosynthetic biomass production is the initial step of the entire biological food web. Further processing of carbon and energy can take two different pathways. The first option would be the classical linear food chain of phytoplankton being consumed by zooplankton, e.g. copepods, and zooplankton being consumed by higher trophic

levels, e.g. fish. DOC can be released in every single step. The second option is the respiration of DOC by heterotrophic prokaryotes, which then can be channeled back into the classical food chain via protist grazing (Azam et al., 1983). This microbial loop is responsible for routing an estimated 50% of the marine primary production and DOC can also be released at every step (Azam, 1998).

1.2 Dynamics of phytoplankton primary production and the role of marine polysaccharides

As globally important primary producers, phytoplankton, such as diatoms, contribute to the production of particulate and dissolved organic carbon. Only because of the high turnover rates of phytoplankton and its products, it is possible to contribute to almost 50% of the global primary production (Ciais et al., 2013; Field et al., 1998). However, this is despite the fact that it constitutes only up to 0.4% of the total marine biomass. Diverse physicochemical conditions, i.e. nutrient supply, temperature and irradiance, lead to a very heterogeneous primary production and the aforementioned bloom and bust events (Pérez et al., 2005; Pedersen and Borum, 1996). Regions displaying high primary production in the form of temporal or constant phytoplankton blooms are mainly coastal areas or upwelling systems that are characterized by high nutrient supply and mixing of water masses (van Dongen-Vogels et al., 2012; Arístegui et al., 2009; Gattuso et al., 1998; Cloern, 1996; Sambrotto et al., 1986). The remote sensing of chlorophyll a can be used as proxy for primary production (FIG 1.2).

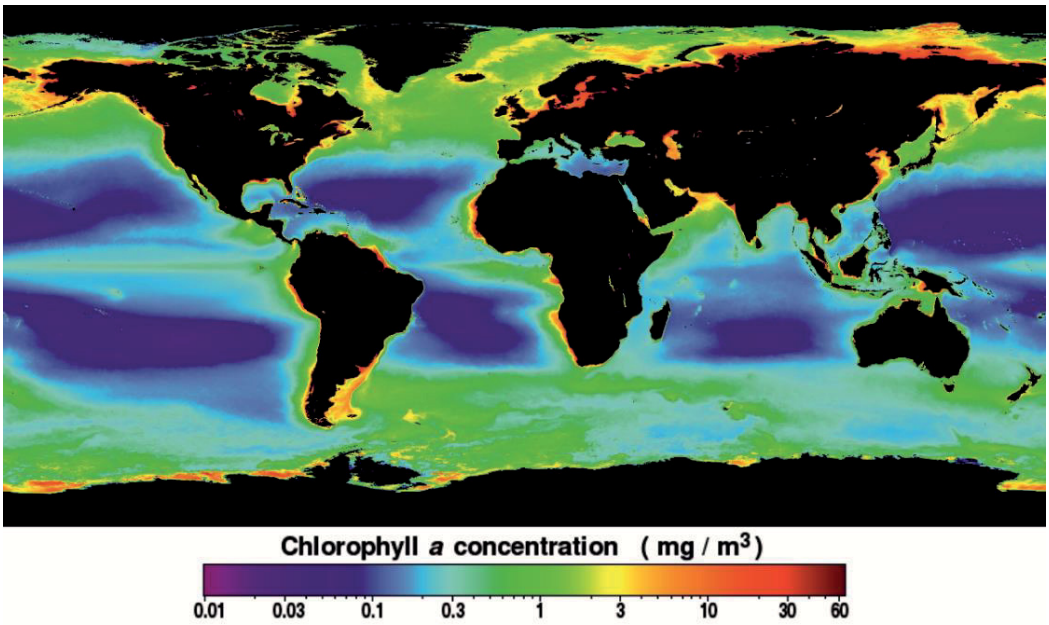


FIG 1.2 Mean chlorophyll a concentrations as a proxy of primary production from July 4, 2002 till December 31, 2017. Data was obtained from the NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group using the Moderate-resolution Imaging Spectroradiometer (MODIS) on the Aqua Satellite.

Direct and indirect measurements of both environmental bulk analyses (Benner et al., 1992) and on phytoplankton derived material (Aluwihare and Repeta, 1999; Biersmith and Benner, 1998; Biddanda and Benner, 1997; Myklestad, 1995) yielded substantial proportions of polysaccharides (up to 50% of DOM in environmental samples) and other high molecular weight (HMW) molecules (≥ 1kDa). Yet our understanding of the chemical composition of phytoplankton derived marine carbon pools is still relatively limited (Moran et al., 2016), especially in regard to the abundant polysaccharides.

This structurally diverse class of biological macromolecules consists of chains of monosaccharides linked by glycosidic bonds. They can differ in linkage type, branching, monosaccharide composition, sulfation, acetylation and methylation. This variety of different polysaccharide molecules can be also found in different marine phytoplankton groups (Painter, 1983). Polysaccharides can be used in cellular communication, as cell wall composite, adhesives or as storage compounds (Gal et al., 2016; Senni et al., 2011; Painter, 1983). In case of storage compounds, the polysaccharides mostly remain inside of the cell to serve as carbon and energy source. However, intracellular polysaccharides can also be liberated into the surrounding water by several mechanisms. They can be released passively by diffusion, by active excretion, cell lysis, as a result of e.g. viral infection, and by sloppy zooplankton feeding and egestion (Suttle, 2005; Fuhrman, 1999; Strom et al., 1997; Brussaard et al., 1995; van Boekel et al., 1992). Heterogeneous extracellular carbohydrates may even form a mucous layer around the algae (Alderkamp, Buma, et al., 2007; Hoagland et al., 1993; Hellebust, 1965).

Furthermore, the overflow hypothesis states that major parts of DOM exudation are a consequence of carbon fixation and an uncoupling of growth and photosynthesis, due to inefficiencies in cell physiology (Thornton, 2014; Fogg, 1983). Up to 50% of the organic matter produced by algae might end up as DOM (Ducklow et al., 2001; Azam et al., 1983). Culture experiments demonstrated that more DOM is being released under nutrient limitation (Obernosterer and Herndl, 1995; Zlotnik and Dubinsky, 1989; Myklestad, 1974, 1977; Hellebust, 1965; Marker, 1965; Guillard and Wangersky, 1958). The idea is that the growth of the cell depends on a large extend on a sufficient nitrogen supply for proteins, nucleic acids, etc. (Fogg, 1983). When the photosynthetic apparatus is already in place, the cell could not be able to reorganize its proteome anymore, and since light and inorganic carbon is often less limited this situation leads to an overflow once the storage capacity has been exceeded. It might not always be necessary to reorganize the proteome, but a strategy like this means only little or no cost to the algae.

1.3 The algal polysaccharide laminarin

Laminarin, which is often also referred to as laminaran, is one of the main storage compounds for the internal energy and carbon supply in marine phytoplankton and macroalgae (Painter, 1983). Macroalgae produce a type of laminarin, that contains an additional mannitol group at the reducing end of the molecule, whereas microalgae produce the so-called chrysolaminarin, that lacks this mannitol group. In the following text, it will be only referred to as laminarin.

Laminarin is a linear polysaccharide of a β-1,3-D-linked glucose backbone with a mean degree of polymerization (DP) of ~20-30 and β-1,6-linked glucose side chains (FIG 1.3). Many highly productive algae produce laminarins, that can vary in the degree of polymerization and branching (DB) (Størseth et al., 2005, 2006; McConville et al., 1986; Paulsen and Myklestad, 1978; Beattie et al., 1961). During light conditions laminarin is being produced and stored in intracellular vacuoles, where it can then be used for the continuation of the metabolism during low light conditions and even at night (Chiovitti et al., 2004; Granum et al., 2002; Vårum and Myklestad, 1984; Barlow, 1982; Handa, 1969). Even though it was never directly measured on a large scale, laminarin has been hypothesized to be one of the most abundant polysaccharides in marine systems (Alderkamp, van Rijssel, et al., 2007).

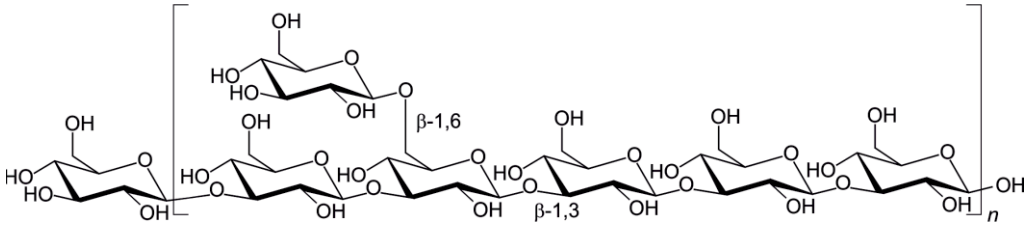


FIG 1.3 Structure of the β-glucan laminarin consisting of a linear β-1,3-D-glucose backbone and with monomeric β-1,6-linked glucose branches.

Carbohydrates have been measured not only in DOC and particles but also in sediments and sediment pore water, which indicates that not all carbohydrates are easy to digest for bacteria (Arnosti and Holmer, 1999; Benner et al., 1992; Cowie and Hedges, 1984). However, the recalcitrance of a certain polysaccharide can depend on its structural complexity and compared to other marine polysaccharides, laminarin is considered to be rather labile, relatively easy to digest and by that also less recalcitrant (Alderkamp, Buma, et al., 2007; Warren, 1996). It is thought that once it liberated from a microalgal cell, it is immediately consumed by heterotrophic bacteria without becoming a part of the DOM pool on a long term or forming polymer gel particles that are known to consist of various polysaccharides (Verdugo et al., 2004). Laminarin has a high energy density and is therefore of high value for heterotrophic bacteria. The theoretical catabolism of laminarin containing 20 glucose monomers via glycolysis, pyruvate oxidation, citric acid cycle and electron transport chain yields gross energy in the form of 640 adenosine

triphosphate (ATP) per mol (~0.18 ATP per gram). The most common lipids in diatoms and other microalgae are triacylglycerides (TAG) (d'Ippolito et al., 2015; Guschina and Harwood, 2006). As a comparison, the conversion of a single TAG molecule consisting of three unsaturated 16-C fatty acid chains via pyruvate oxidation, citric acid cycle and beta-oxidation yields 402 ATP per mol (~0.5 ATP per gram).

Due to the ubiquitous abundance of marine phytoplankton, the perception of laminarin has always been that it plays a key role in the carbon cycling in the upper water layers (Alderkamp, van Rijssel, et al., 2007), despite the lack of constrained numbers. Based on the estimation of 40% of the annual marine organic carbon production of 45–50 gigatons (1 Gt = 10^9 t) being generated by diatoms alone and an approximated carbon contribution of laminarin to that (Alderkamp et al., 2006; Granum et al., 2002; Janse et al., 1996; DM Nelson et al., 1995; Myklestad, 1974), Alderkamp et al. estimated the global annual production of laminarin by diatoms to 5-15 Gt (Alderkamp, van Rijssel, et al., 2007). As a comparison, the current annual increase of atmospheric CO_2 is 4 Gt (Ciais et al., 2013).

1.4 Prokaryotic polysaccharide degradation in the microbial loop

The carbon of released algal polysaccharides is eventually taken up by heterotrophic bacteria. The low concentrations of free monosaccharides suggests that polysaccharides are being consumed quickly (Carlson, 2002; Skoog et al., 1999; Rich et al., 1996). In surface waters, the most abundant heterotrophs belong to the Proteobacteria and Bacteriodetes (Cottrell and Kirchman, 2000a; Hagström et al., 2000; Glöckner et al., 1999) and it has been shown that they are also involved in the degradation of low and high molecular weight molecules (Fernández-Gómez et al., 2013; Alonso-Sáez and Gasol, 2007; Elifantz et al., 2005, 2007; Kirchman, 2002; Cottrell and Kirchman, 2000b; Glöckner et al., 1999; Jooste and Hugo, 1999). They are globally distributed and especially the abundance of Bacteroidetes, e.g. the genus *Formosa*, is associated with the occurrence of phytoplankton blooms (Teeling et al., 2012; Pommier et al., 2007; Pinhassi et al., 2004). Both free-living and particle attached lifestyle was observed (Pedrotti et al., 2009).

The enzymatic machinery that bacteria are using resembles the wide variety of possible polysaccharide structures. However, prokaryotic cells have different strategies to consume liberated polysaccharides. Large molecules (>600 Da) cannot be uptaken through the cell membrane via porins (Weiss et al., 1991). As a result, some bacteria deploy specialized carbohydrate active enzymes (CAZymes) outside of the outer membrane that cleave the large polysaccharide into oligosaccharides that can be uptaken by the cell. Arnosti referred to them as the 'gatekeepers' of the carbon cycle (Arnosti, 2011). However, not all of the heterotrophic bacteria produce extracellular enzymes. 'Cheating' bacteria take up hydrolysis products, which were produced by the extracellular enzymatic activity of other organisms (Allison, 2005). The

third uptake strategy is the 'selfish mechanism', where certain polysaccharides can be taken up by enzymes and binding proteins that stay attached to the cell membrane, so that the cell does not lose the substrate after initiation of the polymerization (Reintjes et al., 2017; Cuskin et al., 2015).

In Bacteroidetes, the enzymatic machinery involved in the sensing, transport and degradation of a certain polysaccharide are often clustered in genetic islands termed polysaccharide utilization loci (PUL) (Grondin et al., 2017; Martens et al., 2009; Tancula et al., 1992). The specificity of a certain hydrolyzing CAZyme is limited, which means that the enzyme is able to hydrolyze only specific glycosidic bonds (Berlemont and Martiny, 2016). This is the reason why bacteria have to evolve a set of enzymes for the complete degradation of a certain polysaccharide depending on the complexity of the molecule. A very complex example from the terrestrial world is the so called degradome of rhamnogalacturonan-II, i.e. the entirety of enzymes that are needed to degrade this complex plant polysaccharide. The gut bacterium *Bacteroides thetaiotaomicron* expresses 25 enzymes for 21 different glycosidic linkages, 13 different monomers, multiple branching and additional functional side groups (Ndeh et al., 2017) (FIG 1.4).

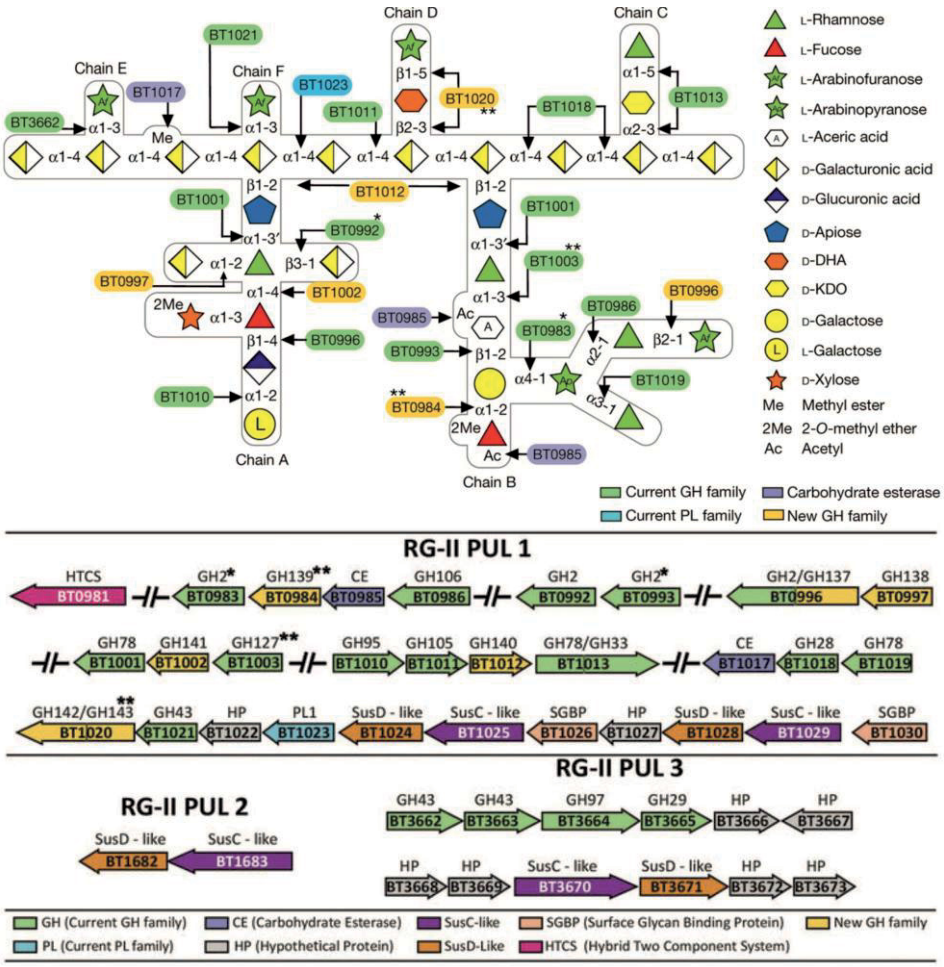


FIG 1.4 An example of specialized prokaryotic polysaccharide degradation. The Rhamnogalacturonan-II molecule, one of the most complex polysaccharides in nature, and its enzymatic degradome organized in polysaccharide utilization loci (PUL), The 25 different enzymes in the three PULs of *Bacteroides thetaiotaomicron* cleave 21 different linkages, resulting in smaller oligos, functional side groups and 13 different sugar monomers. Figure was adapted from Ndeh et al. 2007 with permission.

In the marine environment, it was shown that bacteria can specialize for certain types of polysaccharides and that for example phytoplankton blooms can induce distinct patterns of bacterial succession regarding their function and diversity (Bunse and Pinhassi, 2017; Landa et al., 2016; Kabisch et al., 2014; Teeling et al., 2012; Bauer et al., 2006).

1.5 Glycoside hydrolases and mechanisms of glycan degradation

The large chemical diversity of polysaccharides raises a wide variety of CAZymes (Lombard et al., 2014). Glycoside hydrolases (GH) are one of five different groups of CAZymes. They are the primarily responsible enzymes for the degradation of polysaccharides into mono- and oligosaccharides. Although the polysaccharide diversity is much higher compared to other classes of biological macromolecules (Laine, 1994), the number of different protein folds and catalytic mechanisms is limited. The classification of CAZymes is based on their amino acid sequence and merges them into enzyme families. The CAZy database is currently describing 152 glycoside hydrolase, 105 glycosyl transferase (GT), 28 polysaccharide lyase (PL), 16 carbohydrate esterase (GE), 83 carbohydrate binding module (CBM) and 14 families with auxiliary activity (AA) (Lombard et al., 2014).

Glycoside hydrolases catabolize laminarin degradation, which is why they were focused in this work. They hydrolyze the glycosidic linkage of glycosides. This hydrolysis eventually leads to the formation of a sugar hemiacetal and an aglycon. GHs exhibit two different catalytic mechanisms: the classical Koshland retaining mechanism and the inverting mechanism (Koshland, 1953). The retaining mechanism of β-glucosidases that all enzymes in this work are utilizing, can be divided into two steps (FIG 1.4A). They are assisted by two amino acid residues: the acid/base and the nucleophile residues that are typically glutamates or aspartates. They are located approximately 5.5 Å apart from each other (Vocadlo and Davies, 2008). In the first glycosylation step (1), the acid residue protonates the oxygen at the glycosidic bond that is going to be cleaved. At the same time, the second catalytic residue acts as a nucleophile by attacking the anomeric carbon. A glycosyl enzyme intermediate is being created via an oxocarbenium ion-like transition state. In the second deglycosylation step (2), the first residue is now acting as a base catalyst by deprotonating a water molecule during its hydrolyzation of the glycoside at the anomeric carbon. If another carbohydrate alcohol acts as an acceptor instead of a water molecule during the second step, the entire reaction is called transglycosylation. The configuration of the anomeric carbon stays the same during the retaining mechanism, whereas it is being shifted by the inverting mechanism (FIG 1.5B). This one-step mechanism (1) is carried out by residues that are further apart from each other (~10.5 Å) and it does not form an intermediate (Zechel and Withers, 2000; Qingping Wang et al., 1994).

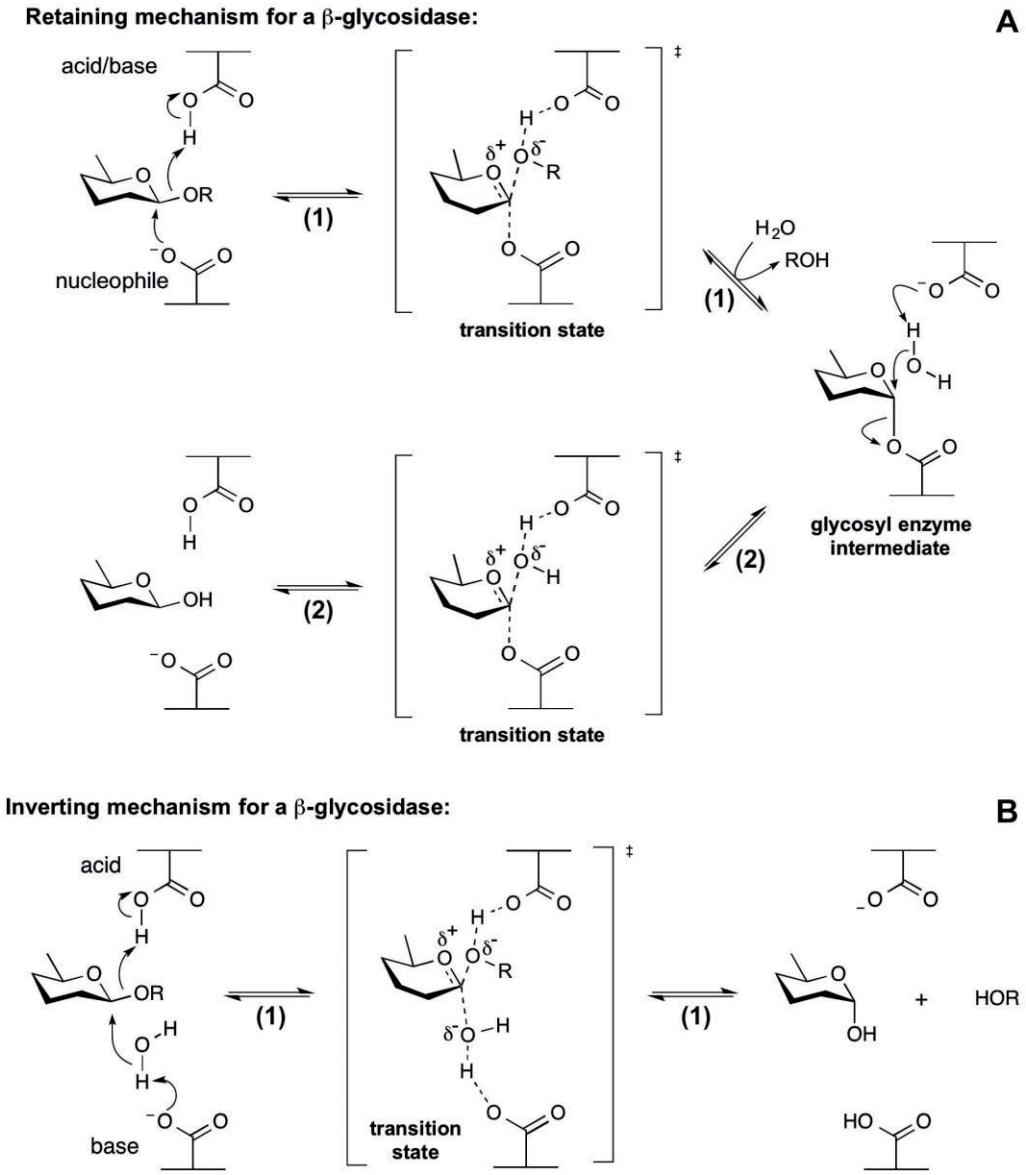


FIG 1.5 The retaining **(A)** and inverting **(B)** catalytic mechanism of glycoside hydrolases (adapted from Withers, S. and Williams, S. "Glycoside Hydrolases" in *CAZypedia*, available at URL http://www.cazypedia.org/, accessed 5 March 2018). In the retaining mechanism, the acid residue protonates the oxygen at the glycosidic bond during the first glycosylation step (1). At the same time, the second catalytic residue acts as a nucleophile by attacking the anomeric carbon and a glycosyl enzyme intermediate is being created via an oxocarbenium ion-like transition state (‡). In the second deglycosylation step (2), the first residue is acting as a base catalyst by deprotonating a water molecule during its hydrolyzation of the glycoside at the anomeric carbon. The configuration of the anomeric carbon is being retained, whereas it is being shifted by the one-step inverting mechanism.

In addition to their catalytic mechanism, glycoside hydrolases can be also distinguished by their mode of action, depending on the products of the hydrolysis reaction (FIG 1.6). Exo-acting enzymes cleave at the end of the substrate chain, whereas endo-acting enzymes cleave in the middle (Davies and Henrissat, 1995). In both modes, the enzyme dissociates from the substrate after the reaction. A pocket topology indicates an exo-activity in the three-dimensional structure of the protein. In contrast, endo-active enzymes exhibit catalytic grooves. Processive enzymes

dissociate from the substrate only after several successive hydrolysis reactions that can again be categorized as exo- or endo-active. Their topology can be either described as a groove or as a tunnel.

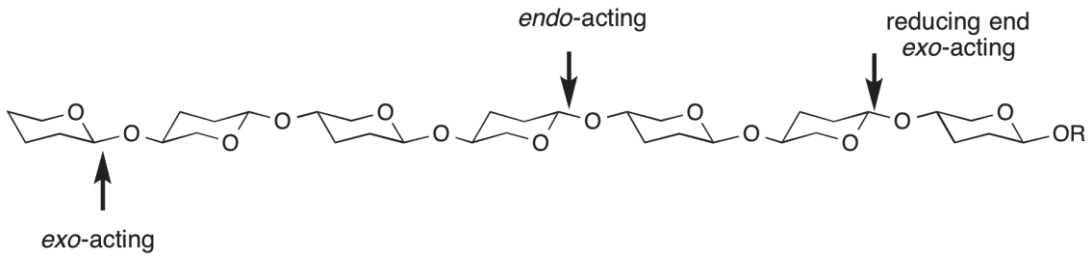


FIG 1.6 The enzymatic modes of action of glycoside hydrolases (from Withers, S. and Williams, S. "Glycoside Hydrolases" in *CAZypedia*, available at URL http://www.cazypedia.org/, accessed 5 March 2018)

A nomenclature for the enzymatic binding of a substrate was introduced by naming sub-sites away from the point of cleavage with increasing positive numbers towards the reducing end and negative numbers towards the non-reducing end (Davies et al., 1997) (FIG 1.7). The cleavage is taking place between the –1 and +1 sub-sites.

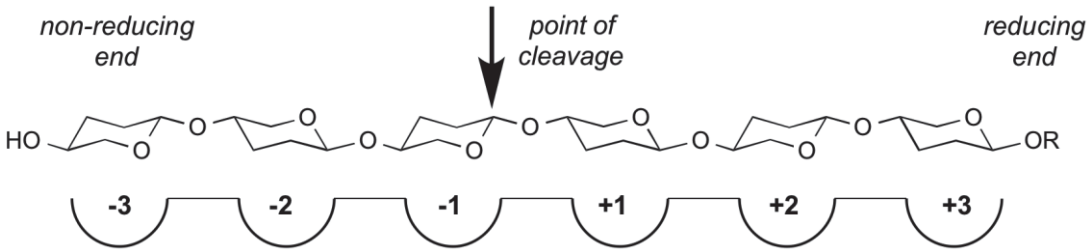


FIG 1.7 Nomenclature of polysaccharide processing enzymatic sub-sites (from Withers, S. "Sub-site nomenclature" in *CAZypedia*, available at URL http://www.cazypedia.org/, accessed 5 March 2018)

1.6 The analytical technologies of marine glycobiology

Compared to other classes of biological macromolecules, such as nucleic acids, proteins or lipids, the complexity of polysaccharide structures makes this molecule class very difficult to analyze. Polysaccharides can differ in the epimeric and anomeric configuration and ring size of their monosaccharide building blocks; they can differ in their glycosidic linkage type and the amount and position of branching; and finally they can exhibit additional functional groups that in the end might lead to two structurally different molecules with identical mass (FIG 1.8). Laine calculated that hexasaccharides can form 1.05×10^{12} theoretically unique structures (Laine, 1994). In comparison, a hexapeptide and a hexanucleotide can form only 6.4×10^{7} and $\sim 4.1 \times 10^{3}$ different

structures, respectively. The technological advances that were made in nucleotide and protein analysis over the past decades could not be transferred to the field of glycobiology. A 'carbohydrate sequencing' technology is not in sight yet.

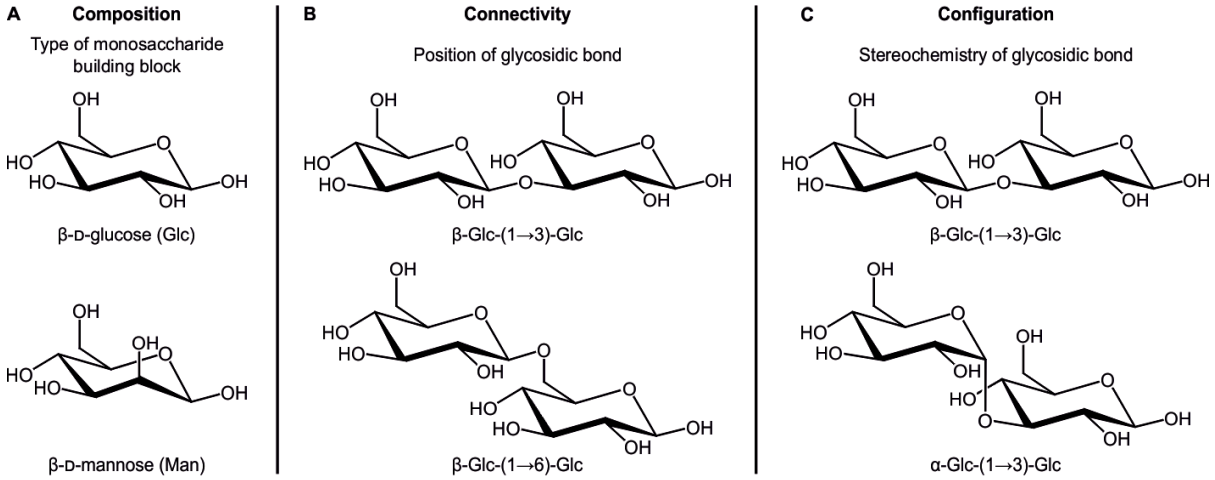


FIG 1.8 Structural features of polysaccharides adapted from Hofmann et al. 2015. (A) Polysaccharide composition is determined by monosaccharide isomers, that can differ stereochemically at only one carbon atom. **(B)** Connectivity is defined by the position of the glycosidic linkage between the monomers. **(C)** The configuration describes the stereochemistry of the anomeric carbon, which can be also involved in a linkage. The configuration can be either characterized as α or β.

There are several techniques for the direct and indirect measurement of carbohydrates. A more indirect approach is using fluorophore-labeled monomeric substrate proxies, e.g. 4-methylumbelliferyl-linked glucose. They can be used for an easily applicable and sensitive detection of extracellular CAZyme activities (Hoppe, 1983; Somville and Billen, 1983). However, the main disadvantage of these monomeric substrate proxies is that they cannot determine extracellular enzymatic activity (EEA) on polymeric substrates, especially not via endo-active enzymes, and they often do not adequately model the polymer degradation (Zhanfei Liu et al., 2010; Feller et al., 1996; Warren, 1996). Even though the substrate proxies are still widely used, fluorescently labeled polysaccharides were developed to overcome the limitations (Arnosti, 1995). The measurement requires more time and instrumentation but it yields valuable estimates for the potential EEA of a bacterial community (Ziervogel et al., 2007, 2010, Ziervogel and Arnosti, 2008, 2009; Murray et al., 2007; Arnosti et al., 2005; Arnosti, 2003).

The most common approach for the qualitative and quantitative analysis of the carbohydrates itself is acid hydrolysis. By using strong acids, the polysaccharide is hydrolyzed and the resulting monomers can be analyzed via simple reducing sugar assays or high-performance liquid chromatography (HPLC) (Engel and Händel, 2011; Lever, 1972). Although this technique is widely applied, the information derived from it is also rather limited. Linkage type, stereochemistry, branching or sequences are lost during the hydrolysis. The method is used for the analysis of monosaccharide composition and total carbohydrate amounts (McCarthy et al.,

1996; Pakulski and Benner, 1994). The application of more advanced analyses like nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry yield the highest level of detail of a carbohydrate sample (Caballero et al., 2016; Graiff et al., 2015; Hedges et al., 2001). NMR spectroscopy is still the only techniques that sufficiently resolve the stereochemistry of a certain polysaccharide. However, the large sample size, its high level of purity and the expertise that is needed to analyze environmental sample by using these sophisticated techniques renders them as a valuable source of in-depth analyses but only of very limited use for bulk measurements.

These challenges in the analysis of carbohydrates in general and marine carbohydrate especially, resulted in marine polysaccharides being severely understudied despite their importance in the marine carbon cycle. However, it also gave rise to new methodologies based on biological tools. One approach is to use carbohydrate-specific antibodies in combination with a microarray technology in order to semi-quantitatively determine the polysaccharides in a crude but concentrated sample (Vidal-Melgosa et al., 2015). This method is similar to the sandwich-like enzyme-linked immunosorbent assay (ELISA). If a certain polysaccharide needs to be detected, one has to include the respective antibody in the setup, while unknown polysaccharides are not going to be detected.

Another idea for a new stereospecific and linkage-specific method without the need for extensive purification of a natural sample is the use of enzymes that evolved specific activities on certain polysaccharides. A methodology like this is well established in the *in situ* analysis of plant and mammalian cells and also in the agro-food industry (Gilbert, 2010; Brunt et al., 1998; Ward et al., 1989; Whitaker, 1974). Similar to DNA restriction enzymes and with ready-to-use kits, one would be able to selectively cleave certain polysaccharides in a crude mixture into mono- and oligomers, which can be quantified afterwards by simple reducing sugar assay or HPLC similarly to the acid hydrolysis. A future 'carbohydrate sequencing' technology could be based on specific enzymes for a range of different polysaccharides.

1.7 Aim and outline of the thesis

The role of polysaccharides in the carbon cycle is poorly understood due to the technological challenges in the analysis of this class of biological macromolecules. The currently used acid hydrolysis for glycan quantification is highly unspecific and does not distinguish between glycans in natural samples. We know that microbes and their enzymes can do better. The enzymatic toolkit that heterotrophic bacteria evolved in order to gain energy and carbon from algal carbohydrates can be used to overcome this gap of knowledge and technology.

Therefore, the overall aim of this thesis was the development and application of a novel enzymatic method for the specific quantification of the algal storage glucan laminarin in order to shed more light on the global distribution of this specific polysaccharide and its importance in the marine carbon cycle.

The first part, *Accurate quantification of laminarin in marine organic matter with enzymes from marine microbes* focuses on the method development. The proteomic machinery that is used by marine microbes to degrade laminarin was produced in a heterologous system and characterized biochemically. The laminarin specificity of the chosen enzymes was demonstrated and compared to another less selective enzyme. The enzymatic activity was optimized and the new assay was compared to the less specific, classical acid hydrolysis based on diatom lab cultures and a small set of environmental samples.

The second part, *Adaptive mechanisms that provide competitive advantages to marine Bacteroidetes during microalgal blooms* contributed to the environmental proteomic, metagenomic and cultivation-based examination of the laminarin degradation by specialized *Formosa* strains during phytoplankton blooms. We added biochemical and crystallographic structure data of the streamlined bacterial laminarinase machinery.

The third part, *Laminarin Quantification in Microalgae with Enzymes from Marine Microbes* is a simple and easily applicable protocol of the newly developed method from the first part. By publicly and access-free sharing of both this protocol and the respective vectors of the laminarinases from the method, we made the application of the method as easy as possible for researchers from different fields.

The fourth part, *Laminarin is the major marine sugar polymer and indicates bioenergy states of the surface ocean* focuses the application of the new method on a set of samples from different oceanic regions. The samples included water samples from time series, different size fractions, vertical profiles, marine snow particles and a meridional transect. The resulting laminarin concentrations were related to certain groups of phytoplankton organisms, the total chlorophyll a and POC concentrations. It was attempted to underline the importance of the marine polysaccharide laminarin by its sheer abundance in the ocean, since the approach to measure a single polysaccharide in the ocean has never been tried before.

2 Manuscripts

2.1 Manuscript I: Accurate quantification of laminarin in marine organic matter with enzymes from marine microbes

Stefan Becker[1,2], André Scheffel[3], Martin F. Polz[4] and Jan-Hendrik Hehemann[1,2]

[1] MARUM – Center for Marine Environmental Sciences at the University of Bremen, Germany
[2] Max-Planck-Institute for Marine Microbiology, Bremen, Germany
[3] Max-Planck-Institute of Molecular Plant Physiology, Potsdam-Golm, Germany
[4] Massachusetts Institute of Technology, Cambridge, USA

Published in: Applied and Environmental Microbiology, 83(9), DOI: 10.1128/AEM.03389-16

Contributions to the manuscript (% of *SB's* contribution to the total workload):
SB (35%) and JHH designed the study. *SB* (100%) performed experimental work and data analysis. *SB* prepared figures and tables (90%). *SB* (50%) and JHH wrote the manuscript. AS and MFP were involved in the discussion and editing of the manuscript.

2.1.1 Abstract

Marine algae produce a variety of glycans, which fulfill diverse biological functions and fuel the carbon and energy demands of heterotrophic microbes. A common approach to analyze marine organic matter uses acid to hydrolyze the glycans into measurable monosaccharides. These, however, may derive from different glycans that are built with the same monosaccharides, hence this approach does not distinguish between glycans in natural samples. Here we use enzymes to digest selectively and thereby quantify laminarin in particulate organic matter. Environmental metaproteome data revealed carbohydrate-active enzymes from marine flavobacteria as tools for selective hydrolysis of the algal β-glucan laminarin. The enzymes digested laminarin into glucose and oligosaccharides, which we measured with standard methods to establish the amounts of laminarin in the samples. We cloned, expressed, purified, and characterized three new glycoside hydrolases (GHs) of *Formosa* bacteria: two are endo-β-1,3-glucanases, of the GH16 and GH17 families, and the other is a GH30 exo-β-1,6-glucanase. *Formosa* sp. nov strain Hel1_33_131 GH30 (FbGH30) removed the β-1,6-glucose side chains, and *Formosa agariphila* GH17A (FaGH17A) and FaGH16A hydrolyzed the β-1,3-glucose backbone of laminarin. Specificity profiling with a library of glucan oligosaccharides and polysaccharides revealed that FaGH17A and FbGH30 were highly specific enzymes, while FaGH16A also hydrolyzed mixed-linked glucans with β-1,4-glucose. Therefore, we chose the more specific FaGH17A and FbGH30 to quantify laminarin in two cultured diatoms, namely, *Thalassiosira weissflogii* and *Thalassiosira pseudonana*, and in seawater samples from the North Sea and the Arctic Ocean. Combined, these results demonstrate the potential of enzymes for faster, stereospecific, and sequence-specific analysis of select glycans in marine organic matter.

2.1.2 Importance

Marine algae synthesize substantial amounts of the glucose polymer laminarin for energy and carbon storage. Its concentrations, rates of production by autotrophic organisms, and rates of digestion by heterotrophic organisms remain unknown. Here we present a method based on enzymes that hydrolyze laminarin and enable its quantification even in crude substrate mixtures, without purification. Compared to the commonly used acid hydrolysis, the enzymatic method presented here is faster and stereospecific and selectively cleaves laminarin in mixtures of glycans, releasing only glucose and oligosaccharides, which can be easily quantified with reducing sugar assays.

2.1.3 Introduction

Marine algae are thought to contribute ~50% of the global primary production, converting more carbon dioxide into biomass through photosynthesis than tropical forests (Field et al., 1998). Major outputs of carbon fixation are glycans, which can amount to 80% of the algal biomass

(Myklestad, 1974) and have a variety of functions in energy and carbon storage, as cell wall building material, in cell communication, as adhesives, and as organic templates for biomineralization processes (Gal et al., 2016; Senni et al., 2011; Painter, 1983). Among the most abundant glycans produced by algae are β-glucans, including laminarin, which is a linear polysaccharide of ~20 to 30 β -1,3-linked glucose residues with β-1,6-linked side chains consisting of a single glucose molecule. Laminarin is produced by many algal species that are characterized by high productivity and rapid turnover (Menshova et al., 2014; HJ Shin et al., 2009; Chizhov et al., 1998; Read et al., 1996; Painter, 1983; Usui et al., 1979; Maeda and Nishizawa, 1968; Percival and Ross, 1951). Consequently, laminarin may play a major role in the marine carbon cycle. Diatoms alone are thought to produce ~40% of the 45 to 50 gigatons (1 gigaton = 10^9 tons) of organic carbon annually in the sea (Granum and Myklestad, 2002; Granum et al., 2002; Painter, 1983). Alderkamp et al. estimated that the global annual production of laminarin, also known as chrysolaminarin, by these microalgae amounts to 5 to 15 gigatons (Alderkamp, van Rijssel, et al., 2007). To place this into perspective, the current annual increase in atmospheric CO_2 is 4 gigatons (Ciais et al., 2013). Hence, new methods are required to better monitor the turnover of this abundant glycan in the sea and to improve our understanding of the impact of laminarin on the marine carbon cycle.

Structural and quantitative analysis of glycans in marine samples such as marine snow (sinking aggregated detritus), fecal pellets, algal cell walls, particulate organic matter (POM), and dissolved organic matter remains complicated, because in seawater glycans occur as complex mixtures in which individual components may occur at low concentrations (Hedges et al., 2001). Glycans are built from different sugar monomers, which can be linked in various ways and configurations into linear or branched macromolecules, frequently hosting substitutions with diverse chemical groups (Laine, 1994). A common method for compositional glycan analysis involves acid hydrolysis to identify and to quantify the monomers with reducing sugar assays, chromatographic assays, or spectroscopy (Engel and Händel, 2011). However, information concerning linkage types, sequences, and stereochemistry remains difficult to obtain with this commonly used method. This problem especially applies to marine samples, in which the abundance of polysaccharides is often too low to permit chromatographic purification for more sophisticated nuclear magnetic resonance (NMR) or mass spectrometry analysis, which would allow structural elucidation of the types of glycans in the mixture (Hedges et al., 2001). Therefore, new stereospecific and linkage-specific approaches are required to identify polysaccharides in natural samples, preferably with little or no need for additional purification. Enzymes evolved to overcome the structural diversity of glycans present in nature; in plant and mammalian systems, enzymes are used to dissect and to analyze the structures of complex cell wall and mammalian glycans *in situ* (Gilbert, 2010; Ward et al., 1989). Moreover, enzymes are less prone to side reactions, which are common with acid hydrolysis, leading to conversions of monosaccharides

(Mäki-Arvela et al., 2011). Enzymes may also be powerful tools for the analysis of ecologically relevant marine glycans, which often have different chemical structures than plant glycans and thus require different sets of enzymes.

Bacteria depolymerize algal polysaccharides with glycoside hydrolases (GHs) and polysaccharide lyases (PLs), which are classified in the Carbohydrate-Active Enzymes (CAZy) database into over 130 GH families and over 20 PL families (Lombard et al., 2014). Depolymerization of polysaccharides, including laminarin, involves systems of enzymes with different specificities and modes of action. Laminarinases are classified into endo-β-1,3-glucanases (laminarinases) (EC 3.2.1.6 and EC 3.2.1.39), which cleave randomly within the chain, hydrolyzing the β-1,3 backbone into glucose and oligosaccharides, and exo-β-1,3-glucanases (EC 3.2.1.58), which cleave at the non-reducing ends, hydrolyzing laminarin oligosaccharides into glucose (FIG 2.1.1). Endo-acting laminarinases belong to different glycoside hydrolase families, mainly GH16, GH17, GH55, GH64, and GH81, and exo-acting laminarinases belong to the GH3 family. Exo- and endo-acting enzymes specific for the β-1,3 linkage have been described (Hehemann et al., 2014); however, the enzymes responsible for cleaving the β-1,6-linked side chains in laminarin remain unknown, to the best of our knowledge. The side chains can limit the activity of endo-acting enzymes, owing to steric hindrance, and thereby limit the complete hydrolysis of laminarin (Labourel et al., 2015). Therefore, an enzyme system for complete hydrolysis of laminarin into measurable glucose would require a mixture of endo- and exo-acting enzymes with different specificities, accounting for the two types of linkages present in laminarin.

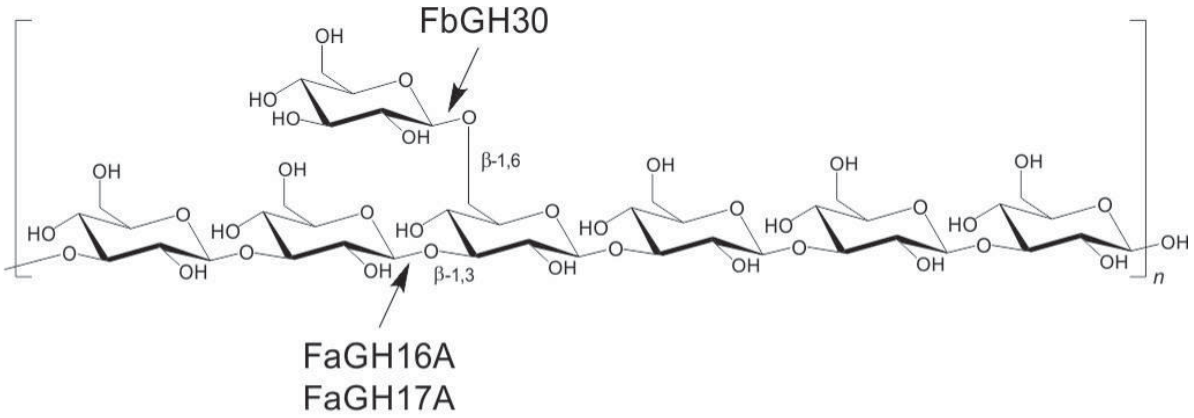


FIG 2.1.1 Laminarin structure and enzyme activities. The linear storage glucan consists of a β-1,3-D-glucose polysaccharide with β-1,6-linked monomer side chains. The characterization of three enzymes from marine Bacteroidetes, belonging to the GH16, GH17, and GH30 families, presented here showed that the GH30 enzyme hydrolyzed the β-1,6-linked side chain and the GH17 and GH16 enzymes hydrolyzed the β-1,3-D-linked main chain of laminarin.

Metagenome and proteome data on marine blooms of algae and bacteria represent a source of ecologically relevant enzymes and support the notion that laminarin represents a crucial carbon and energy source for prokaryotic heterotrophs (Teeling et al., 2012). During spring microalgal

blooms in the North Sea (2009), diatoms supported the growth of heterotrophic *Bacteroidetes*, which consumed glycans and other molecules produced by the algae. Protein expression data showed that putative laminarinases were among the proteins with the highest levels of expression, which suggested that laminarin represented a major food resource for marine *Bacteroidetes*. The GH16 and GH30 families were among the most highly expressed proteins, which suggests that these enzymes may mediate the turnover of laminarin from microalgae and therefore may represent suitable tools for marine glycan analysis.

In *Bacteroidetes*, the carbohydrate-active enzymes involved in the depolymerization of glycans are often clustered in genetic islands termed polysaccharide utilization loci (PUL) (Martens et al., 2009; Tancula et al., 1992). We identified a putative PUL for laminarin catabolism in *Formosa agariphila*, and a homolog of this PUL was upregulated by laminarin in *Gramella forsettii* (Mann et al., 2013). These PUL contain orthologues of the GH16 and GH30 enzymes that were highly expressed during the North Sea algal bloom in 2009; in addition, they contain GH17 enzymes. Therefore, we focused on these conserved PUL of marine *Formosa* spp. and characterized putative GH16, GH17, and GH30 enzymes, showing that they are active laminarinases. In subsequent experiments, we used the enzymes to quantify laminarin in diatom species grown in the laboratory. We also used the enzymes to measure laminarin in environmental samples by digesting particulate organic matter collected on filters from North Sea and Arctic Ocean waters. Our results demonstrate that carbohydrate-active enzymes have strong potential for the specific quantification of select types of glycans in marine organic matter.

2.1.4 Results

2.1.4.1 Heterologous expression and initial characterization of FaGH16A, FaGH17A and FbGH30 enzymes

We cloned and expressed three enzymes, i.e., GH16 and GH17 enzymes from a genetic island of *Formosa agariphila* KMM 3901T (Mann et al., 2013) and a GH30 enzyme from *Formosa* sp. nov strain Hel1_33_131 (Hahnke et al., 2015); both species are involved in laminarin catabolism. Two endo-acting laminarinases from *Zobellia galactanivorans*, with ~44% and 43% sequence identity with GH16A of *F. agariphila* (FaGH16A), have recently been identified and characterized. Although these findings suggest that FaGH16A may be a suitable tool for the analysis of laminarin (Labourel et al., 2014, 2015), the enzymes also hydrolyzed related glucans with β-1,4 linkages, showing that they are of limited specificity. We chose to include a GH16 enzyme in our analysis to illustrate the activity of a promiscuous laminarinase, as a control, and to highlight the advantages of specific enzymes. Enzymes with higher specificity than the GH16 enzyme may be needed for highly selective laminarin analysis. We hypothesized that such alternative enzymes might be from the GH17 and GH30 families. To our knowledge, GH17 enzymes have been described only from terrestrial microbes and plants. They are endo-type enzymes, which are highly specific for non-decorated stretches of β-1,3-glucans (Varghese et al., 1994). This suggests that orthologous genes

from marine bacteria may have a similar function. The GH30 family contains enzymes with known β-1,6-glucanase activity (Oyama et al., 2002; Reese et al., 1962) hence, we reasoned that this enzyme may hydrolyze the β-1,6-linked side chains of laminarin in an exo-type manner and may be required, together with an endo-type enzyme, for efficient depolymerization of laminarin.

After immobilized metal affinity chromatography and size exclusion chromatography, the typical yields of purified enzyme were ~20 mg liter-1 for FaGH16A and ~50 mg liter-1 for FaGH17A. Because the GH30 enzyme from *F. agariphila* was targeted to inclusion bodies without detectable amounts of soluble protein, we chose to clone and to produce an orthologous enzyme from *Formosa* sp. nov. strain Hel1_33_131 (i.e., FbGH30), which was isolated from Helgoland in the North Sea and shares 70% identity with FaGH30. FbGH30 was produced as a soluble protein in Escherichia coli, and overexpression yielded ~30 mg liter-1 of purified protein (FIG 2.1.2A). Size exclusion chromatography and SDS-PAGE data revealed molecular masses of 59 kDa (FaGH16A), 42 kDa (FaGH17A), and 55 kDa (FbGH30), which are close to the theoretical values of 58.5 kDa (FaGH16A), 44.8 kDa (FaGH17A), and 54.7 kDa (FbGH30), suggesting that all three enzymes occur as monomers in solution. This notion was supported by dynamic light scattering (DLS) measurements, confirming that the two enzymes used for analytical assays remained monodisperse and stable for at least 4 months at 4°C. Over a period of 120 days, FaGH17A and FbGH30 maintained hydrodynamic radii of 3.63 ± 0.23 nm (mean ± standard deviation [SD]) and 3.78 ± 0.12 nm, respectively.

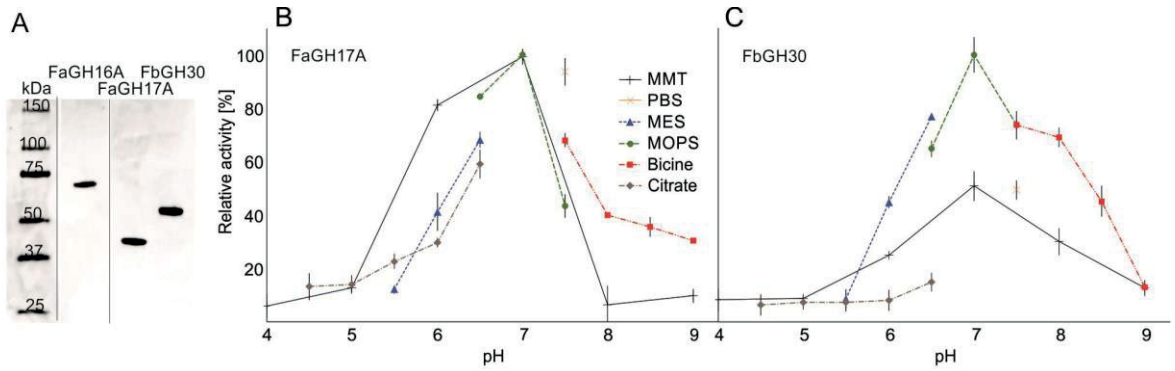


FIG 2.1.2 Recombinant glycoside hydrolases from marine Bacteroidetes showing greatest activity in MOPS buffer at neutral pH. **(A)** SDS-PAGE analysis of purified enzymes. One gel is shown. Unimportant lanes were intentionally omitted, as indicated by vertical lines. Approximately 0.5 μg of each protein was loaded on the gel. All proteins were run on the same gel, and lanes were spliced for clarity. **(B and C)** Enzymatic rates were measured with 0.1% (wt/vol) laminarin, which was hydrolyzed by 100 nM (~5.0 μg ml^{-1}) purified FaGH17A (B) or FbGH30 (C) at 37°C for 30 min in 50 mM buffer. The greatest activity rate was observed in MOPS buffer at pH 7.0 and was set as the 100% reference value. MES, morpholineethanesulfonic acid; MMT, malic acid-MES-Tris base.

2.1.4.2 Temperature stability and pH optima of FaGH17A and FbGH30.

Melting curve measurements with dynamic light scattering revealed the mesophilic character of FaGH16A, FaGH17A, and FbGH30. The melting curves showed that FaGH17A started to aggregate

at ~40°C, FbGH30 at ~42°C, and FaGH16A at ~55°C, as measured with three different protein concentrations (Supplementary FIG 2.1.1). Because FaGH17A and FbGH30 showed higher specificity for laminarin (see below), we concentrated our analysis of stability and activity in different buffer systems on these two enzymes and tested the activity of FaGH17A and FbGH30 with different organic and inorganic buffers and pH values (covering a pH range of 4 to 9), using laminarin as the substrate. Bell-shaped activity profiles were obtained for both enzymes, with the highest enzymatic activities at a pH optimum of 7 (FIG 2.1.2B and C). Both enzymes had the highest activities in 50 mM 3-(N-morpholino)propanesulfonic acid (MOPS) buffer. Therefore, a reaction temperature of 37°C and a buffer system of MOPS or phosphate-buffered saline (PBS), with a pH of 7, were chosen for subsequent experiments.

2.1.4.3 Activity measurements of FaGH16A, FaGH17A, and FbGH30 with glucan polysaccharides and oligosaccharides

Activity measurements with defined substrates revealed that FaGH17A, FbGH30, and FaGH16A were active with laminarin, while having different specificities and modes of action. To test their modes of action, we used fluorophore-assisted carbohydrate polyacrylamide gel electrophoresis (FACE) (Jackson, 1990). We digested laminarin and analyzed samples taken during the time course of the reaction. FaGH16A and FaGH17A hydrolyzed laminarin into oligosaccharides of different sizes, creating a ladder-type profile (Supplementary FIG 2.1.2A), which is typical for endo-acting glycoside hydrolases (Labourel et al., 2015). Time course analysis of the products released by FbGH30 revealed the accumulation of glucose during all phases of the enzyme assay (Supplementary FIG 2.1.2B), a pattern that is typical for an exo-acting enzyme (Labourel et al., 2014, 2015).

To map the specificity of the enzymes, we used different β-glucans from terrestrial plants and marine algae and measured the reaction products with reducing sugar assays. We tested activity with β-1,4-linked carboxy-methyl cellulose, β-1,3/β-1,4-mixed-linked lichenan, β-1,3/β-1,4-mixed-linked glucan (MLG) from barley, and laminarin (FIG 2.1.3A). Hydrolysis was quantified with the p-hydroxybenzoic acid hydrazide (PAHBAH) assay (Lever, 1972), which detects the reducing end, i.e., the aldehyde group. Figure 2.1.3A shows that FaGH16A released the greatest amount of reducing sugar equivalents (0.23 ± 0.006 mg ml^{-1} [mean ± SD]) with laminarin as the substrate, 3.6-fold more than FaGH17A (0.064 ± 0.001 mg ml^{-1}) and 2.8-fold more than FbGH30 (0.083 ± 0.006 mg ml^{-1}) ($P \leq 0.0001$). FaGH16A also hydrolyzed lichenan (0.058 ± 0.004 mg ml^{-1}) and mixed-linked glucan from barley (0.099 ± 0.006 mg ml^{-1}), which contain β-1,3 and β-1,4 linkages. FaGH17A and FbGH30 did not release significant amounts of reducing sugar equivalents when these glucans were used as the substrates ($P \leq 0.0001$). These results pointed to FaGH17A and FbGH30 being more specific than FaGH16A, which appeared to have greater activity with laminarin, however.

To validate these results, we used different glucans (oligosaccharides and polysaccharides) as substrates and analyzed the reaction products by thin-layer chromatography (TLC) (Supplementary FIG 2.1.3). The substrates included laminarin oligosaccharides (biose, triose, and tetraose of the β-1,3 series), cellooligosaccharides (biose and tetraose of the β-1,4 series), mixed-linked cellotetraose with internal β-1,3 linkages at different positions, gentiobiose (a β-1,6-linked biose), and β-1,3/β-1,4-mixed-linked glucan oligosaccharides (FIG 2.1.3B). Of the three enzymes, FaGH16A hydrolyzed most substrates, digesting laminarin and β-1,3 oligosaccharides into glucose and laminaribiose. It also hydrolyzed MLG, which was not hydrolyzed by the other enzymes. FaGH16A hydrolyzed mixed- linked cellotetraoligosaccharides with an internal β-1,4 linkage, an activity that was shown previously for the GH16 laminarinases from *Zobellia galactanivorans* (Labourel et al., 2014, 2015). FaGH17A was exclusively active with laminarin polysaccharides and oligosaccharides, producing glucose and laminaribiose as the major products, which confirmed the results described above. These experiments confirmed that the highly selective activity of terrestrial enzymes in the GH17 family for β-1,3-glucans (Chen et al., 1995; Varghese et al., 1994; Woodward and Fincher, 1982) could be extended to this marine GH17 enzyme. The β-1,3-trisaccharide was the smallest substrate hydrolyzed into glucose and laminaribiose by FaGH17A. FbGH30 hydrolyzed laminarin and produced glucose as the major product, but it did not show activity with the other substrates containing β-1,3 or β-1,4 linkages. Instead, FbGH30 was the only enzyme that hydrolyzed gentiobiose, a β-1,6-linked disaccharide. In aggregate, these experiments indicated that FaGH16A is a promiscuous enzyme that cleaves laminarin and mixed-linked glucans, FaGH17A is a specific endo-β-1,3-glucanase, and FbGH30 is a specific exo-β-1,6-glucanase.

2.1.4.4 Increased hydrolysis of laminarin with mixtures of enzymes

The preferences of FaGH16A, FaGH17A, and FbGH30 for different types of linkages suggested that combinations of these enzymes might improve laminarin hydrolysis and consequently the sensitivity of the analytical assays. We tested this by measuring the activity of enzyme mixtures with laminarin as the substrate in the PAHBAH assay, which is sensitive and feasible and has a limit of detection of 2 nM (Moretti and Thorson, 2008). As noted previously, FaGH16A had the highest activity, compared to FaGH17A and FbGH30. In contrast, combinations of FaGH17A with FaGH16A increased reducing sugar concentrations (FIG 2.1.3C). This effect increased when FbGH30 was added to either enzyme. Together, FaGH16A and FbGH30 increased the yield by a factor of 1.9 ± 0.04 ($P \leq 0.0001$) and FaGH17A and FbGH30 increased the yield by a factor of 6.2 ± 0.43, compared to FaGH17A alone. Because the mixture of FaGH17A and FbGH30 achieved the same total yield as FaGH16A and FbGH30 but showed greater specificity, we chose this enzyme mixture for the enzymatic analysis of laminarin in microalgae.

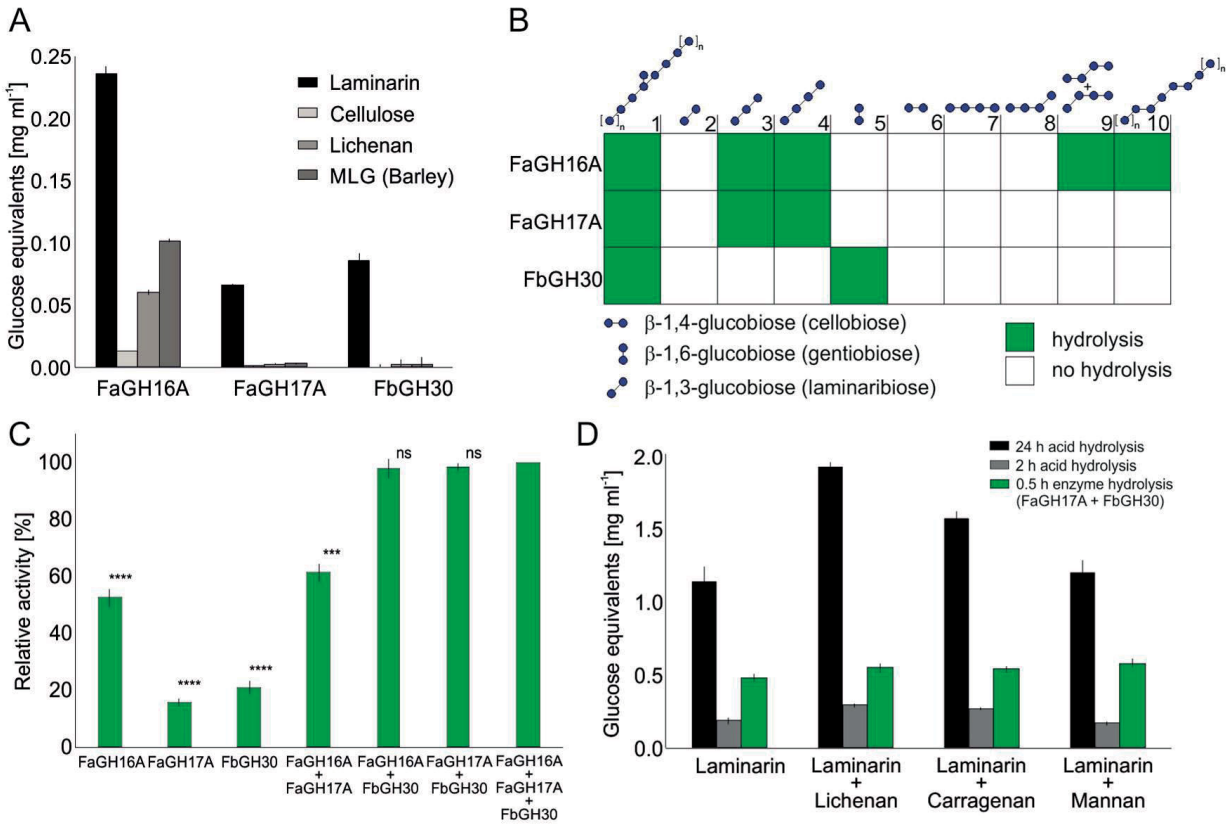


FIG 2.1.3 Glycoside hydrolases showing different levels of specificity for laminarin and related glucans. **(A)** Activity tests with glucan polysaccharides and oligosaccharides containing β-1,3, β-1,4, and β-1,6 linkages, based on the PAHBAH reducing sugar assay. Shown are mean and standard deviation (SD) values from three technical replicates. **(B)** Enzyme specificity tested with defined oligosaccharide substrates, using thin-layer chromatography. The results are presented as a heatmap (Supplementary FIG 2.1.3). The substrates at 0.1% (wt/vol) were hydrolyzed for 30 min at 37°C by 100 nM (~5 µg ml-1) purified enzyme in PBS buffer at pH 7.5. **(C)** Mixtures of FbGH30 and FaGH17A or FbGH30 and FaGH16A, showing greater activity than the individual enzymes. The highest activity level with all three enzymes was set to 100%, and all other samples were compared to that value. Laminarin at 0.1% (wt/vol) was hydrolyzed for 30 min at 37°C by 100 nM (~5 µg ml-1) of each purified enzyme in PBS buffer at pH 7.5, and hydrolysis was measured with the PAHBAH assay. Shown are mean and SD values from three technical replicates. ****, $P \leq 0.0001$; ***, $P \leq 0.001$, independent two-sample Student's t test. ns, not significant. **(D)** Comparison of hydrolysis yields of enzymatic, partial acid, and total acid hydrolysis of different polysaccharides. Lichenan, carrageenan, and mannan at 0.1% (wt/vol) were added to 0.1% (wt/vol) laminarin and were hydrolyzed for 30 min at 37°C with 100 nM purified enzyme (~5 µg ml^{-1} of FaGH16A, FaGH17A, or FbGH30) in 50 mM MOPS buffer. Boiling for 5 min at 100°C stopped the reaction. Partial acid hydrolysis was conducted for 2 h at 20°C with 50 mM H_2SO_4. Total acid hydrolysis was carried out for 24 h at 100°C with 1 M HCl. The reaction mixtures were analyzed with the PAHBAH assay. All experiments were carried out in triplicate. Shown are mean and SD values from three technical replicates.

Next, we investigated how the enzyme mixture performed in comparison with acid when other polysaccharides were present. To evaluate the ability of the enzymes to hydrolyze laminarin in the presence of other polysaccharides, we incubated different glycan mixtures with enzymes or acid and compared the product yields. The mixture of FaGH17A and FbGH30 was incubated with laminarin alone or with added reference polysaccharides, i.e., lichenan, carrageenan, and mannan. The laminarin concentration was kept constant; hence, the product concentrations obtained with the enzymes after 30 min of hydrolysis remained the same in all tested mixtures. This result indicated that only laminarin and none of the other substrates was cleaved by the enzyme mixture (FIG 2.1.3D). Interestingly, the moderate acid hydrolysis protocol that is commonly used for

laminarin quantification in microalgae (Granum and Myklestad, 2002) delivered lower values than the enzyme method even after2hof hydrolysis, which showed that enzymes react faster. Moreover, the acid-derived values varied when other polysaccharides were present, owing to the nonspecific hydrolysis of mannan, lichenan, and carrageenan, which showed that acid is less specific. Acid hydrolysis for 24 h completely hydrolyzed the laminarin into glucose; the yield of reducing ends with enzymes was ~50% of the total acid hydrolysis, indicating that the enzyme mixture cleaved about 50% of the glycosidic linkages in laminarin (Supplementary FIG 2.1.4). However, this longer acid hydrolysis also strongly hydrolyzed the other polysaccharides, leading to highly inflated sugar-reducing signals, compared to the enzyme method. This experiment highlights the superior specificity of enzymes, which hydrolyzed laminarin more quickly and more accurately than the commonly used acid hydrolysis approach.

2.1.4.5 Enzymatic quantification of laminarin in diatom batch cultures and environmental samples

We quantified laminarin in microalgae with the enzyme mixture (FaGH17A and FbGH30) and with the conventional acid-based method (Granum and Myklestad, 2002). The two species *Thalassiosira weissflogii* and *Thalassiosira pseudonana* displayed typical growth behavior (FIG 2.1.4A). Repeated extraction of the retentate with MOPS buffer did not increase the total yield, which confirmed that one step is generally sufficient to extract laminarin (Chiovitti et al., 2004). To account for the fact that the enzymes hydrolyzed only about 50% of the glycosidic linkages, we prepared a calibration curve with laminarin that had been digested under the same conditions as the natural samples, which revealed a linear relationship between laminarin concentrations and reducing sugar signals. This calibration curve enabled the quantification of laminarin despite its incomplete hydrolysis (Supplementary FIG 2.1.5). The correlation coefficients for both the acid hydrolysis and the enzymatic hydrolysis were higher than 0.99.

The conventional acid hydrolysis and the new enzymatic approach gave comparable glucan quantities over the time course of the growth experiment. The acid method gave higher values than the enzyme method, i.e., 8 to 23 pg and 0.3 to 3.9 pg of glucan per cell for *T. weissflogii* and *T. pseudonana*, respectively, compared to 6 to 20 pg and 0.4 to 3.5 pg of glucan per cell, respectively, via enzymatic digestion. While the cell densities and laminarin concentrations correlated well for *T. weissflogii* during the growth experiment (FIG 2.1.4A and B), this was not the case for *T. pseudonana* (FIG 2.1.4A and C). At the beginning of the growth period and during the exponential growth phase, cell numbers increased independently of the laminarin concentrations, which was noted previously for this diatom (Myklestad, 1974). TAB 2.1.1 compares previously established glucan quantities in diatoms derived by acid hydrolysis with values obtained in our experiments, showing that they are in a similar range as the enzyme-based measurements.

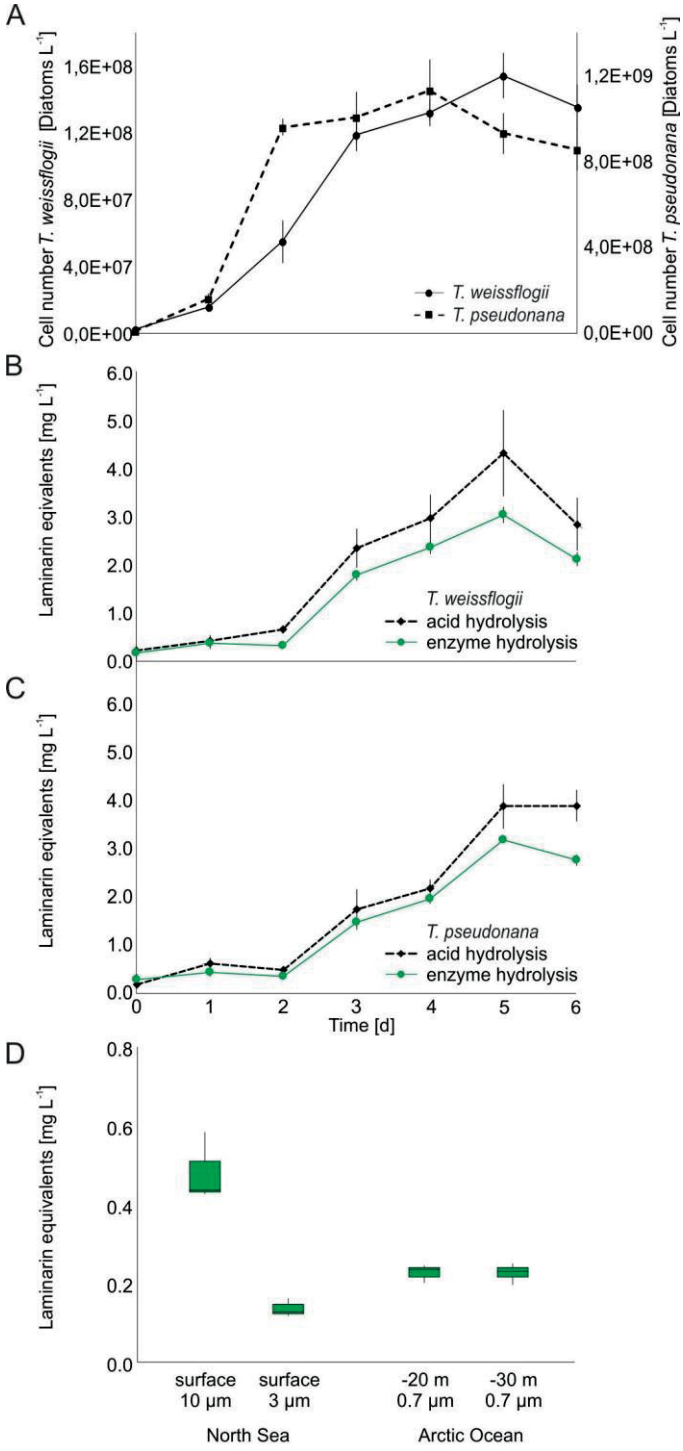


FIG 2.1.4 Laminarin quantification in *Thalassiosira weissflogii* and *T. pseudonana* laboratory cultures and in environmental samples. **(A)** Numbers of diatoms cultured in modified ESAW/HESNW medium at 15°C, with a 12-h/12-h light/dark cycle, for 6 days. **(B)** Quantification of laminarin extracted from *T. weissflogii* by acid (dashed black line) or enzyme (solid green line) hydrolysis. **(C)** Quantification of laminarin extracted from *T. pseudonana* by acid (dashed black line) or enzyme (solid green line) hydrolysis. **(D)** Laminarin contents of particulate organic matter, which was concentrated by filtering seawater from the North Sea near Helgoland (spring 2009) and from the Arctic Ocean near Svalbard (summer 2015) and enzymatically hydrolyzed. Water-soluble extracts were hydrolyzed for 30 min at 37°C with 100 nM (~5.0 mg ml^{-1}) FaGH17A and FbGH30, with 1 mg ml^{-1} BSA, in 50 mM MOPS buffer at pH 7. Alternatively, the extracted polysaccharides were hydrolyzed twice for 2 h at 20°C with 50 mM H_2SO_4, with shaking at 1,500 rpm. The products were quantified with the PAHBAH reducing sugar assay. The calibration curve was prepared with laminarin hydrolyzed with enzyme or acid, and the reducing sugar signals were measured as described above (Supplementary FIG 2.1.5).

We also analyzed the laminarin content in particulate organic matter (POM) obtained from two marine sampling sites (FIG 2.1.4D). POM was obtained by filtering seawater during a diatom-dominated algal bloom in the North Sea near the island of Helgoland and in the Fram Strait in the Arctic Ocean. The retentate from the North Sea corresponds to a time point shortly after the spring bloom of diatoms peaked in a longer time series (Teeling et al., 2012). The proteome data set for bacterioplankton from the same day showed that orthologues of the laminarinases that we used here for the assay were highly expressed *in situ*. Accordingly, we were able to determine laminarin concentrations in the POM fraction by using the enzyme assay. The 10 µm POM fraction of the Helgoland samples collected from 1 liter of seawater contained over 0.48 ± 0.09 mg and the 3 µm fraction contained over 0.13 ± 0.02 mg of laminarin (FIG 2.1.4D). Acid hydrolysis of these two samples yielded a slightly greater signal (data not shown). This discrepancy is similar to what we observed with the cultivated diatoms. The enzyme mixture hydrolyzed laminarin in samples from the Arctic Ocean, providing nearly the same signals for the two water depths; about 0.22 ± 0.03 mg of laminarin was contained within the POM collected from 1 liter of seawater. Altogether, these data show that these enzymes can detect and quantify laminarin in POM without additional purification.

TAB 2.1.1 Glucan amounts of various diatom species, related to dry weight and growth phase

Species	Sampling growth phase	Medium[a]	Temperature (°C)	Light cycle (light h/dark h)	Irradiance (μmol photons m^{-2} s^{-1})	Glucan wt (pg $cell^{-1}$)	Hydrolysis method	Reference
T. weissflogii	Mean over 60 days	F/2	20	24/0	300	45.2	Acid	Li 1979
T. pseudonana	Late exponential	F/2	20	12/12	75	2.4	Acid	Brown 1991
Skeletonema marinoi	Late exponential	F/2	20	12/12	75	2.5	Acid	Brown 1991
Chaetoceros debilis	Late stationary	F/2	18	24	200	15.8	Acid	Størseth et al. 2006
Chaetoceros affinis	Exponential	F/10	13	14/10		40.0	Acid	Myklestad and Haug 1972
S. marinoi	Stationary	F/10	13	14/10	200	375.0	Acid	Granum et al. 2002
T. weissflogii	Late exponential	NEPCC	15	12/12	140	16.0	Acid	This study
T. pseudonana	Late exponential	NEPCC	15	12/12	140	1.3	Acid	This study
T. weissflogii	Late exponential	NEPCC	15	12/12	140	15.0	Enzymatic	This study
T. pseudonana	Late exponential	NEPCC	15	12/12	140	1.5	Enzymatic	This study

[a] Details for the F/2 and F/10 media can be found in (Guillard and Ryther, 1962); details for the NEPCC medium can be found in (Harrison et al., 1980).

2.1.5 Discussion

Highly specific, stable, and active enzymes that do not hydrolyze other polysaccharides enable the selective hydrolysis and quantification of laminarin even in crude algal samples. The enzymes are stereospecific and function in high concentrations of salt, which limits the sensitivity of alternative methods commonly used to analyze marine organic matter, such as proton NMR or mass spectrometry (Kelly et al., 2002). Three enzymes, from the GH16, GH17, and GH30 families, were characterized, revealing various specificities and different potentials as tools for the analysis of marine glucans. Compared to FaGH16A, with broad activity, FaGH17A showed narrow specificity for β-1,3-glucan, which may be due to the enzyme architecture. Active sites of GH17 enzymes are known to be longer than those found in GH16 enzymes (Qin et al., 2015; Jeng et al., 2011) and to

require longer stretches of non-decorated β-1,3-glucan chain for productive binding, a relationship that was observed previously among agarases of the marine bacterium *Zobellia galactanivorans* (Hehemann, Correc, et al., 2012). The longer active sites of GH17 enzymes increase the selectivity for non-decorated β-1,3-glucans and decrease the activity of GH17 enzymes with β-1,6-branched laminarin and β-1,3/1,4-mixed-linked glucans; the shorter active site of GH16 enables hydrolysis of mixed-linked glucans and native laminarin, as indicated by greater activity with those substrates. Labourel et al. (Labourel et al., 2015) proposed an alternative hypothesis, i.e., GH16 laminarinases may be able to cleave β-1,6-decorated laminarin due to the presence of small pockets lining the active sites, which can host the side chain glucose residues. These pockets and shorter active sites may explain the greater activity with decorated laminarin, compared to GH17 enzymes.

Combining FbGH30 with FaGH17A was key for efficient hydrolysis of laminarin and significantly improved the final yield. It should be noted that the improved hydrolysis with FbGH30 is a consequence of its ability to remove the β-1,6-glucose side chains. After this debranching, FaGH17A can cleave the non-decorated backbone; together, the two enzymes achieve increased laminarin hydrolysis yields. The TLC data showed that FbGH30 was active with gentiobiose, which has only one β-1,6 linkage, while it was inactive with oligosaccharides with β-1,3 and β-1,4 linkages, indicating that FbGH30 is a highly stereospecific enzyme. The only β-1,6 linkage in laminarin is the side chain, which indicates that FbGH30 removes the side chains of the polysaccharide and in this way opens new recognition sites for the GH17 enzyme. The comparative analysis of these enzymes and their different specificities emphasizes that a sensitive choice of carbohydrate-active enzymes is paramount for the quantitative analysis of marine glycans.

The enzyme mixture of FaGH17A and FbGH30 is more specific and faster and can quantify laminarin despite not hydrolyzing it completely. Thin-layer chromatography experiments revealed that the GH17 enzyme did not cleave laminaribiose (FIG 2.1.3B), indicating that the enzymes did not completely hydrolyze laminarin; comparison with the product yield obtained with acid hydrolysis revealed that the enzymes hydrolyzed about 50% of the glycosidic linkages. This result implies that additional enzymes, such as β-glucosidases, are required to complete the hydrolysis of the remaining laminaribiose and larger oligosaccharides. However, we hydrolyzed different laminarin concentrations with the enzymes, producing a linear calibration curve to account for incomplete hydrolysis. This calibration curve established the amounts of laminarin in diatom extracts and environmental samples. Moreover, our results showed that the acid hydrolysis method based on the method described by Granum and Myklestad (Granum and Myklestad, 2002) produced lower values than the enzyme method even after 2 h of acid hydrolysis, showing that enzyme hydrolysis can be faster. Increasing the reaction time, acid concentration, or temperature led to nonspecific hydrolysis of co-occurring polysaccharides. This

nonselective hydrolysis of other polysaccharides or other compounds may explain the slightly higher values measured with the acid method. Although we cannot exclude the possibility that the enzymes may actually underestimate the laminarin content, they evolved to be specific and therefore it is clear that they provide a more conservative measure of laminarin, especially when polysaccharides and other labile compounds are present in the sample.

Applying the enzyme method to environmental samples demonstrated the successful quantification of laminarin. The North Sea retentate included material from the same time and location in the algal bloom at which the orthologues of the laminarinases used here for the assay were highly expressed in the metaproteome data set. We measured concentrations between 0.13 ± 0.02 and 0.48 ± 0.09 mg liter^{-1} in the particulate organic matter fraction, revealing that laminarin was indeed the elicitor of enzyme expression in the North Sea. To put this into perspective, van Oijen et al. measured total carbohydrate concentrations in Antarctic Ocean water samples by using acid hydrolysis (van Oijen et al., 2003). Compared to chlorophyll concentrations that are usually measured during a spring bloom in the North Sea (9.3 g liter^{-1}), this water body contained a significantly lower chlorophyll concentration (0.7 µg liter^{-1}), and smaller amounts of biomass yielded lower carbohydrate concentrations of up to 15 µg liter^{-1}. In contrast, Pakulski and Benner showed similar or higher concentrations of total carbohydrates in the surface water of the North Atlantic Ocean, the equatorial Pacific Ocean, the Gulf of Mexico, and the Antarctic Ocean (Pakulski and Benner, 1994).

Moran and colleagues recently stated that, in light of an anthropogenically altered carbon cycle (Moran et al., 2016), a more thorough understanding of organic matter, with its fluxes and intertwined relationships between molecules and metabolizing organisms, requires advancing technologies in different fields of environmental sciences. One important question addresses which molecules represent the largest conduits of carbon flux through the labile marine dissolved organic matter pool. Glycans belong to the most abundant molecules in surface waters, where they are produced by microalgae, but their concentrations seem to decrease rapidly with time and depth; this has been inferred by measuring the glucose concentrations of acid-hydrolyzed samples (Keith and Arnosti, 2001; Borch and Kirchman, 1997; McCarthy et al., 1996). Those studies, together with radiocarbon dating of glycan-rich dissolved organic matter (Repeta and Aluwihare, 2006), support the notions that heterotrophs rapidly consume glycans and they belong to the most reactive macromolecules. It is possible that laminarin belongs to these abundant but quickly metabolized compounds, which fuel catabolic activity in ocean surface waters. This enzymatic method might be a suitable approach to address these questions by unambiguously measuring concentrations of β-glucans, such as laminarin, in dissolved and particulate organic matter.

2.1.6 Materials and Methods

2.1.6.1 Cloning

The putative laminarinase genes (FAGH16A locus tag, BN863_18740; FaGH17A locus tag, BN863_18720) were amplified from the genomic DNA of *Formosa agariphila* (KMM 3901T) (GenBank accession number HG315671.1), which was obtained from the Leibniz Institute DMSZ. Primers were designed to amplify the coding regions corresponding to the catalytic module and a polycystic kidney disease (PKD) domain (Bycroft, 1999) in the case of FaGH16A (FaGH16A-forward, CTGGTGCCGCGCGGCAGCCATATGGCTAGCGAAGACGATGTGTCCATAG; FaGH17Aforward, CTGGTGCCGCGCGGCAGCCATATGGCTAGCTTAGGATGTAACAATAAAACTCAGATTGCG; FaGH16A-reverse, ATCTCAGTGGTGGTGGTGGTGGTGCTCGAGTTATTGGTAAACTCTTAC; FaGH17Areverse, ATCTCAGTGGTGGTGGTGGTGGTGCTCGAGTTAATTTTTAATTGGTGGCA-CTGCAACG). The PCR was performed using standard PCR amplification protocols. The primers were designed to include ~30-bp overhangs complementary with the pET28a(+) vector, for Gibson assembly (Gibson et al., 2009), which results in an N-terminal hexahistidine tag in the recombinant protein. A restriction site was included to allow for alternative restriction enzyme-based cloning. The NEB Gibson cloning kit was used, according to the manufacturer's instructions. Colonies were screened for positive clones by using T7 primers and colony PCR. Because the GH30 enzyme from *F. agariphila* was targeted to inclusion bodies without detectable amounts of soluble protein, we chose to clone and to produce an orthologous enzyme (FbGH30) from *Formosa* sp. nov. strain Hel1_33_131, which was isolated from Helgoland in the North Sea (Hahnke et al., 2015) and shares 70% identity with FaGH30. The FbGH30 gene (locus tag, FORMB_24730) was synthesized and cloned into a pET28a(+) vector by the company GenScript. The vectors with insert were purified from 2-ml Luria-Bertani (LB) broth precultures using the Qiagen Miniprep kit. For protein expression, the vectors were used to transform E. coli BL21(DE3). The cells were stored at -80°C until further use. For common use, we deposited the expression vectors for FaGH17A (Addgene plasmid 86462) and FbGH30 (Addgene plasmid 86463) in the Addgene plasmid repository (http://www. addgene.org).

2.1.6.2 Overexpression and purification

Proteins were expressed in E. coli BL21(DE3) cells by using 1 liter of autoinduction medium (Studier, 2005), which was inoculated with 1 ml of an overnight culture at 37°C in LB medium supplemented with 50 µg ml^{-1} kanamycin. Bacteria were grown for 3 days at 20°C in ZYP autoinduction medium with 100 µg ml^{-1} kanamycin, with shaking at 150 rpm. Cells were harvested by centrifugation at 4,500×g for 25 min at 4°C, and the pellet was stored at -20°C. After thawing, the cells were suspended in 15 ml IMAC buffer A (20 mM Tris-HCl [pH 8.0], 500 mM NaCl, 40 mM imidazole). Resuspended cells were disrupted in a high-pressure homogenizer (Constant Systems) at 1.5 kbar. After the addition of 10 mg DNase (Sigma) and mixing for 10 min at 20°C,

the lysate was centrifuged at 16,000×g for 45 min at 4°C. The supernatant was loaded onto a 5 ml HiTrap IMAC HP column (GE Healthcare), which was charged with 1 column volume (CV) of 500 mM $NiSO_4$ and equilibrated with 5 CVs of IMAC buffer A. After sample injection, the column was washed with IMAC buffer A (15 CVs), and the protein was eluted with 5 CVs of a linear gradient from IMAC buffer A to IMAC buffer B (20 mM Tris-HCl [pH 8.0], 500 mM NaCl, 500 mM imidazole), at a flow rate of 5 ml min^{-1}, at 20°C. The 1 ml fractions were analyzed by SDS-PAGE. Fractions corresponding to a band at the expected size were pooled and concentrated in an ultrafiltration stirred cell (Amicon, Millipore), on a polyethersulfone membrane with a 10 kDa cutoff value (Millipore). An aliquot of 1 to 2 ml was loaded onto a 120-ml HiPrep 16/60 Sephacryl S-200 HR column that had been equilibrated previously with 3 CVs of SEC buffer (20 mM Tris-HCl [pH 7.5]) at 20°C. The protein was eluted with 1 CV of SEC buffer, and the fractions were analyzed by SDS-PAGE. Fractions with a single band at the expected size and at an elution time corresponding to monomeric or dimeric protein were combined and concentrated. The concentration and hydrodynamic radius were determined by using a BioSpectrometer (Eppendorf) and dynamic light scattering (DLS).

2.1.6.3 Dynamic light scattering

A DynaPro plate reader II (Wyatt Technology), in combination with 384-well plates (Aurora Microplates), was used for measurements of hydrodynamic radii via dynamic light scattering. The 0.2 µm-filtered protein samples were measured in triplicate, after removal of bubbles by centrifugation at 2,000×g for 2 min in the plate. Samples were measured in SEC buffer (20 mM Tris-HCl [pH 7.5]). Ten measurements with 5 s acquisition time were taken from every well at 25°C. For analysis and calculation of diffusion coefficients and hydrodynamic radii via the Stokes-Einstein equation, Dynamics software (v.7.1.9.3; Wyatt Technologies) was used. For determination of melting curves, we added one drop of paraffin oil to each well, to prevent evaporation. The temperature protocol increased the temperature stepwise from 25 to 80°C, at a rate of 0.25°C min^{-1}.

2.1.6.4 Comparative hydrolysis of different polysaccharides with FaGH16A, FaGH17A, and FbGH30 enzymes.

To evaluate the substrate specificity of the enzymes, laminarin from *Eisenia bicyclis* (Carbosynth), carboxy-methyl cellulose, lichenan, and mixed-linked glucan from barley (MLG) (all at 0.1% [wt/vol], from Megazyme) were hydrolyzed for 30 min at 37°C with 100 nM purified enzyme (~5 µg ml^{-1} FaGH16A, FaGH17A, or FbGH30) in 1×phosphate-buffered saline (PBS) buffer. The reaction was stopped by boiling the samples for 5 min at 100°C. The reaction mixtures were used for reducing sugar assays and thin-layer chromatography.

In order to compare the hydrolysis yields of enzymatic hydrolysis and partial and total acid hydrolysis, different polysaccharides, including laminarin from *E. bicyclis* (Carbosynth), lichenan (Megazyme), carrageenan (Sigma), and mannan (Megazyme) (all at 0.1% [wt/vol]), were

hydrolyzed for 30 min at 37°C with 100 nM purified enzyme (~5 µg ml^{-1} of FaGH16A, FaGH17A, or FbGH30) in 50 mM MOPS buffer. The reaction was stopped by boiling the samples for 5 min at 100°C. Partial acid hydrolysis was conducted in the same manner as for the algal samples, i.e., hydrolysis twice for2hat 20°C with 50 mM H_2SO_4, with shaking at 1,500 rpm (Granum and Myklestad, 2002), followed by centrifugation for 15 min at 12,000 rpm and collection of the supernatant. Total acid hydrolysis was carried out for 24 h at 100°C with 1 M HCl, with shaking at 1,500 rpm. The reaction mixtures were measured with the reducing sugar assay described below. All experiments were carried out in triplicate.

2.1.6.5 Quantification of product production with PAHBAH reducing sugar assay

One milliliter of a freshly prepared 9:1 mixture of reagent A (0.3 M 4-hydroxybenzhydrazide, 0.6 M HCl) and reagent B (48 mM trisodium citrate, 10mMCaCl2, 0.5M NaOH) was added to 0.1 ml of sample. The mixture was heated for 5 min at 100°C. Absorbance was determined at 410 nm using a BioSpectrometer (Eppendorf). In order to calculate the amount of released reducing ends as glucose reducing end equivalents, a calibration curve with glucose (0, 0.03125, 0.0625, 0.125, 0.25, 0.5, 1, and 2 mg ml^{-1}) was determined for every experiment. Calibration with laminarin was conducted by hydrolysis and measurement of laminarin from *E. bicyclis* (Carbosynth) at different concentrations (0, 0.03125, 0.0625, 0.125, 0.25, 0.5, 1, and 2 mg ml^{-1}). For the samples that were used to establish the calibration curve, enzyme- or acid-based hydrolysis was carried out in the same manner as for the samples from the wild.

2.1.6.6 Product profiling with thin-layer chromatography

In order to produce defined band-like spots, we spotted three spots of 0.4 µl of the hydrolysis products of the following substrates, each at a concentration of 1 mg ml^{-1}, directly next to each other on a silica gel 60 TLC plate (Merck-Millipore). We used laminarin from *E. bicyclis* (Carbosynth), laminaribiose, laminaritriose, laminaritetraose, laminaripentaose, laminarihexaose, cellobiose, cellotetraose, mixed-linked cellotetraose 1 (cellotriosyl-β-1,3-D-glucose), mixed-linked cellotetraose 2 (cellobiosyl-β-1,3-D-cellobiose plus glucosyl-β-1,3-D-cellotriose), mixed-linked glycan from barley (all from Megazyme), and gentiobiose (AppliChem). Ethyl acetate/acetic acid/methanol/formic acid/water (8:4:1:1:1) was used as the solvent, as reported by Jeng et al. (Jeng et al., 2011). The plate was developed for 10 min at 100°C with 10% sulfuric acid in ethanol.

2.1.6.7 Product profiling with fluorophore-assisted carbohydrate electrophoresis

Laminarin from *E. bicyclis* (0.1% [wt/vol]; Carbosynth) was hydrolyzed for 30 min at 37°C with 100 nM purified enzyme (~5 µg ml^{-1} of FaGH16A, FaGH17A, or FbGH30) in 1× PBS buffer. Aliquots of 200 µl were taken at 0 s, 5 min, 20 min, 40 min, and 60 h. The reaction was stopped by boiling the sample for 5 min at 100°C, and the samples were dried in a vacuum concentrator plus (Eppendorf). Derivatization and electrophoresis were performed as described previously

(Jackson, 1990). Two microliters of 0.15 M 8-aminonaphthalene-1,3,6-trisulfonic acid (ANTS) and 5 µl of 1 M $NaBH_3CN$ were added to the dried oligosaccharides. The reaction mixture was incubated overnight at 37°C and again dried under vacuum. The oligosaccharides were resuspended in 20 µl of 25% glycerol, and 5 µl aliquots were loaded onto a standard 36% tris-glycine-acrylamide gel without SDS or a 4 to 20% Tris-HCl Ready Gel (Bio-Rad). Electrophoresis was performed at 200 V at 4°C using running buffer without SDS (192 mM glycine, 25 mM Tris [pH 8.5]).

2.1.6.8 Algal growth, extraction, and hydrolysis

The diatoms *Thalassiosira weissflogii* and *T. pseudonana* were cultivated in triplicate batch cultures with modified ESAW/HESNW culture medium (Harrison et al., 1980). They were kept in T75 suspension cell culture flasks at a constant temperature of 15°C, with a 12-h/12-h light/dark cycle, without stirring. A total volume of 250 ml was inoculated with 7 ml of a 25-ml culture that had been grown for 7 days. The growth medium was inoculated with 600,000 cells of *T. weissflogii* and 2,000,000 cells of *T. pseudonana*. Illumination was applied with five Sylvania F36W/GRO tubes, resulting in an irradiance of ~140 µmol photons m^{-2} s^{-1}. Fifteen milliliters of the cultures was harvested every day. Cultivation was continued until the stationary phase was reached. Diatom abundance was measured by cell counting using a Sedgewick-Rafter counting chamber. The algae were harvested by filtration at 200 mbar on a GF/F glass microfiber filter and were stored at -20°C until extraction. Glycans were extracted from one-half of the filter with 50 mM MOPS buffer at 60°C for 60 min, with shaking at 1,500 rpm (adapted from the method described by Zha et al. (Zha et al., 2012)), followed by centrifugation for 15 min at 12,000 rpm and collection of the supernatant. A second extraction was performed to confirm the extraction efficiency. The samples were hydrolyzed for 30 min at 37°C with 100 nM purified enzyme (~5 µg ml^{-1} of FaGH17A or FbGH30), after the addition of 1 mg ml^{-1} bovine serum albumin (BSA), in 50mM MOPS buffer. The reaction was stopped by boiling the samples for 5 min at 100°C. The glycans on the other one-half of the filters were extracted and hydrolyzed twice for 2 h at 20°C with 50 mM H_2SO_4 , with shaking at 1,500 rpm (Granum and Myklestad, 2002), followed by centrifugation for 15 min at 12,000 rpm and collection of the supernatant. All experiments were carried out in triplicate. Both enzymatically hydrolyzed and acid-hydrolyzed samples were used for reducing sugar assays.

2.1.6.9 Sampling and processing of environmental samples

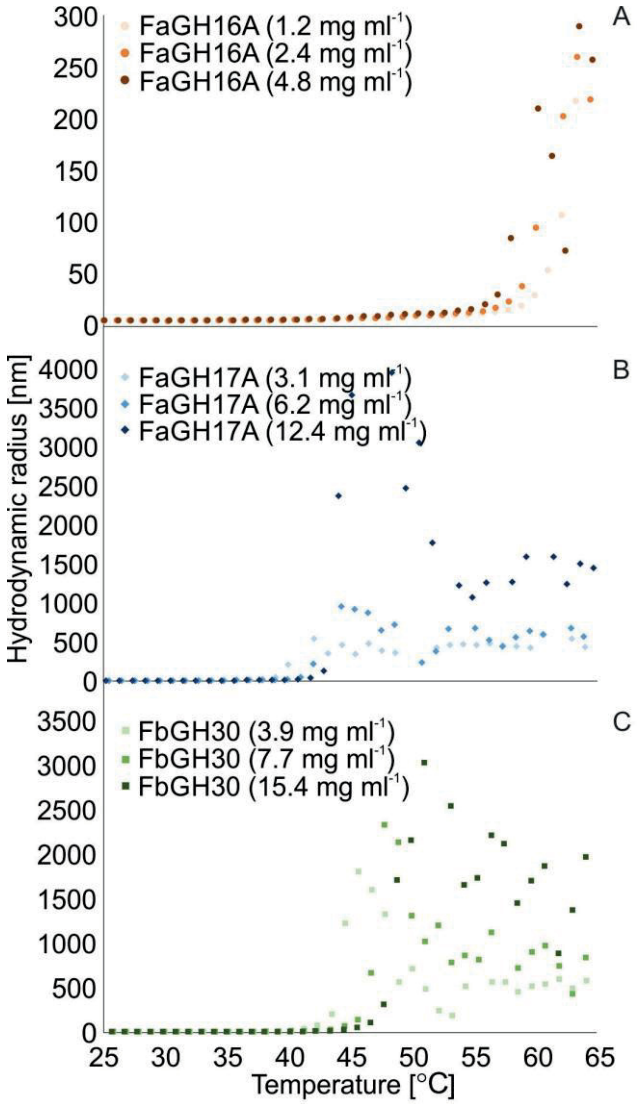
We used environmental samples derived from two different sites. In 2009, samples were collected during a North Sea spring algal bloom that was dominated by diatoms (Wiltshire and Dürselen, 2004). Details about the sampling were described by Teeling et al. (Teeling et al., 2012). In short, samples were taken at the long-term ecological research site near the North Sea island of Helgoland (Kabeltonne station, at 54°11'03''N, 7°54'00''E) on 31 March 2009. During the campaign, biomass from the sea surface was sampled by using sequential filtration through

142 mm diameter, 10 µm- and 3 µm pore-size, TCTP polycarbonate filters (Millipore), with peristaltic pumping. The filters were frozen at -80°C. The filters were cut into pieces of equal size, and triplicates were treated in the same way as the laboratory cultures. Additionally, we analyzed samples obtained during the expedition PS93 cruise (summer 2015) to the Fram long-term ecological research site near Svalbard in the Arctic Ocean. Via a rosette water sampler on board the research vessel Polarstern, samples were taken at station EG-IV (78°50'07''N, 2°47'95''W) on 30 July 2015. Two liters of sampled water was filtered onto a GF/F glass microfiber filter and stored at -20°C until extraction. The samples were extracted for 60 min at 60°C with 5 ml of 50 mM MOPS buffer, with shaking at 1,500 rpm, followed by centrifugation for 15 min at 12,000 rpm and collection of the supernatant. The samples were hydrolyzed for 30 min at 37°C with 100 nM purified enzyme (~5 µg ml^{-1} FaGH17A or FbGH30), with 1 mg ml^{-1} BSA, and the reaction was stopped by boiling the samples for 5 min at 100°C. The samples were measured using the PAHBAH reducing sugar assay.

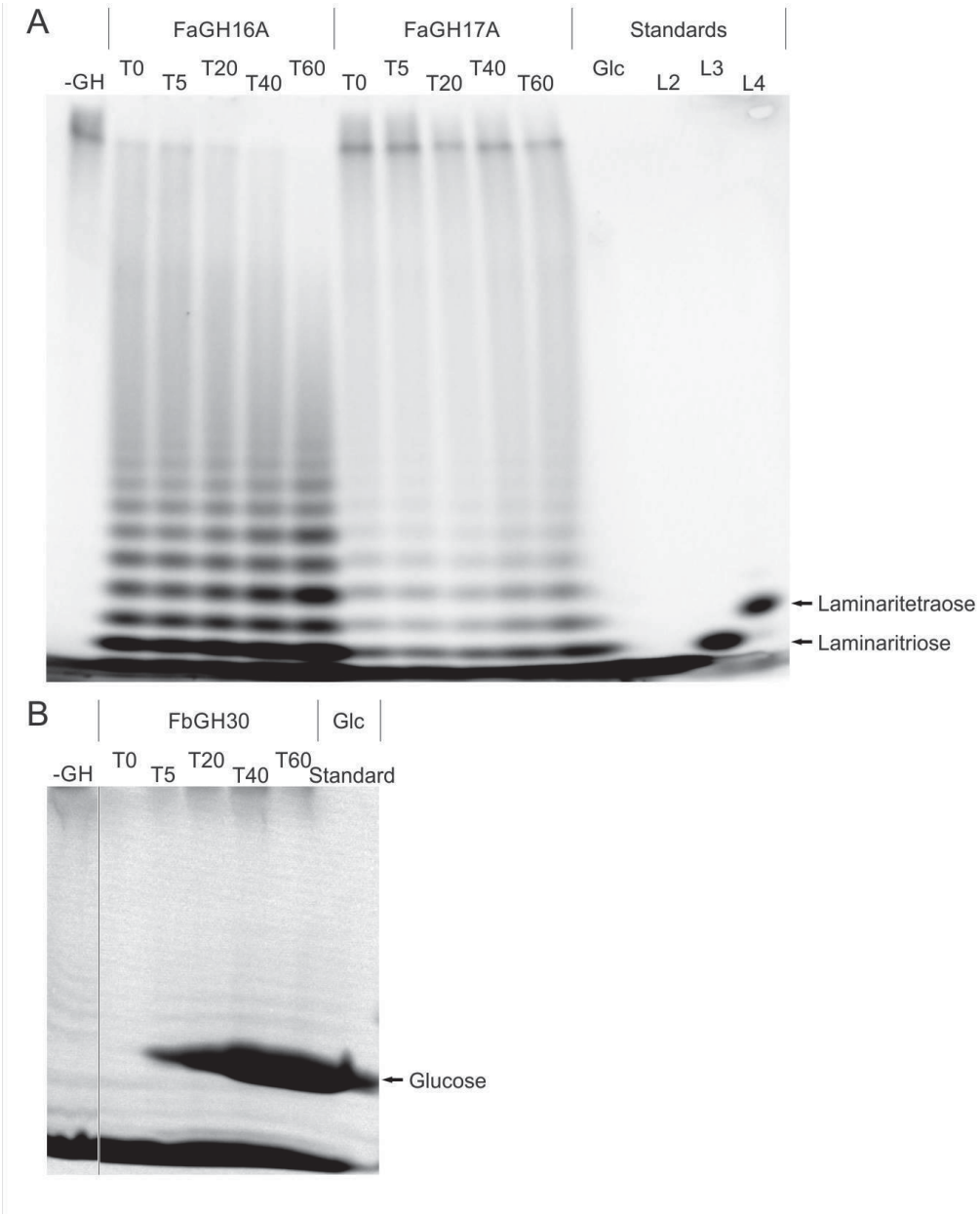
2.1.7 Acknowledgements

This research was supported by the Deutsche Forschungsgemeinschaft (grant HE 7217/1-1 to J.-H.H.) and by the Max Planck Society. This work was partially supported by a grant from the U.S. Department of Energy (grant DE-SC0008743 to M.F.P.). We thank Rudolf Amann for reading and improving the manuscript and providing us with samples from the North Sea. We thank Hanno Teeling for supplying us with the FbGH30 sequence and Carol Arnosti for providing critical insight. Furthermore, we thank Morten Iversen for providing samples from the Arctic Ocean and the crew of the RV Polarstern for helping with sample acquisition.

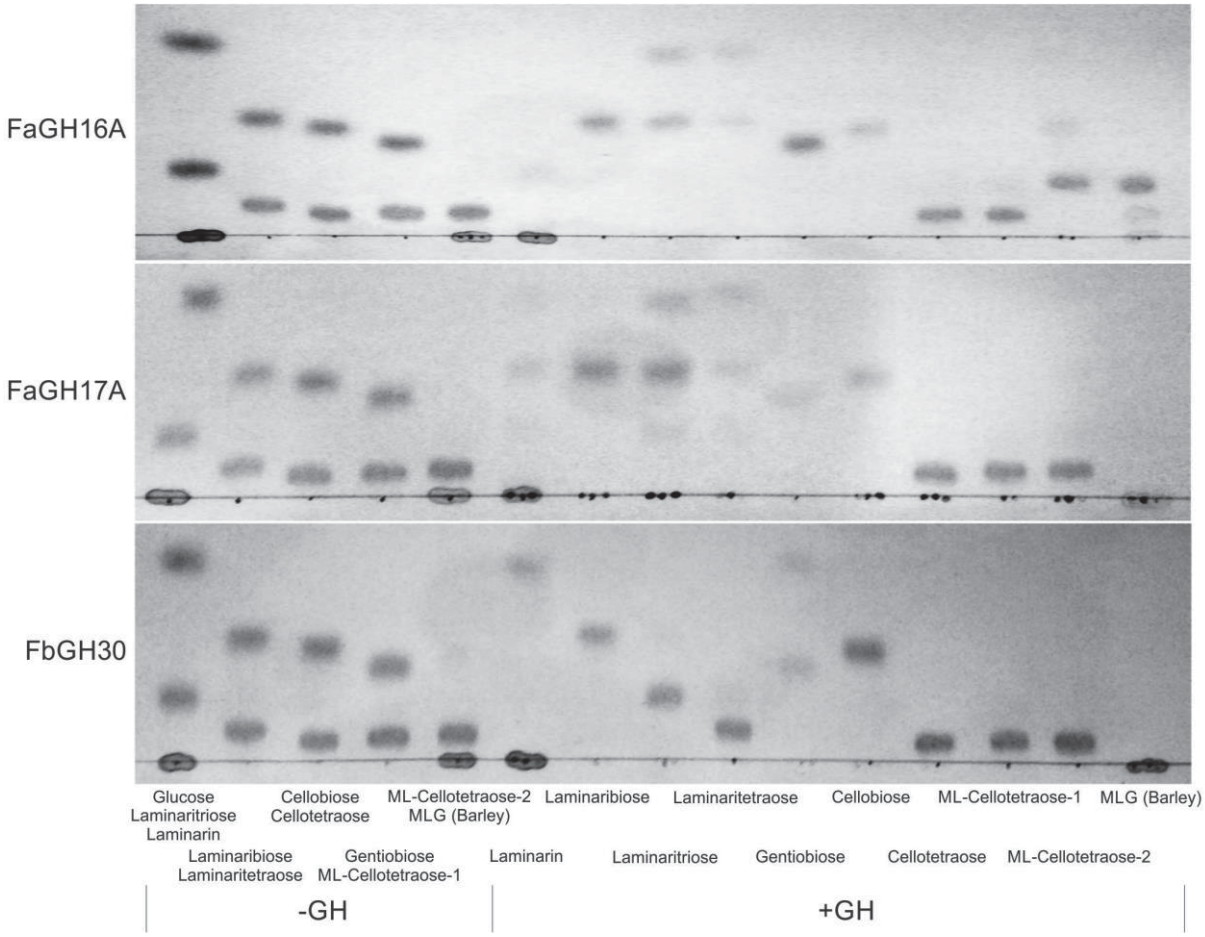
2.1.8 Supplementary material



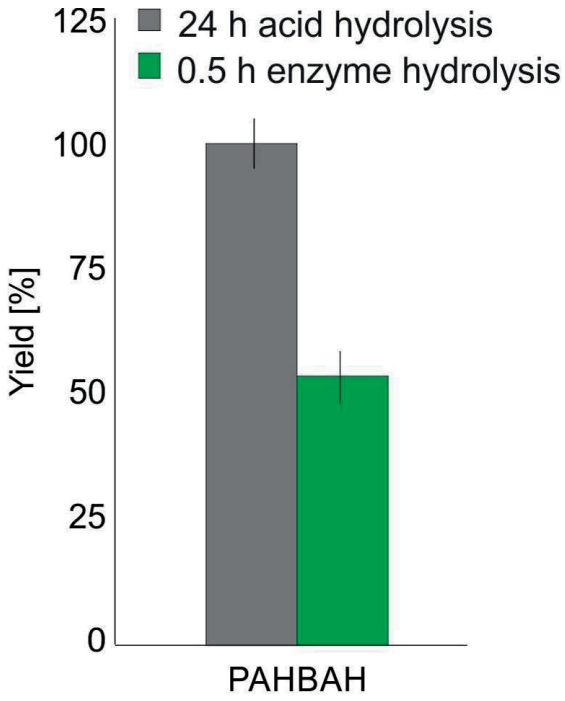
Supplementary FIG 2.1.1 Temperature stability of FaGH16A, FaGH17A and FbGH30 measured by dynamic light scattering. The melting curves showed the mesophilic character of all three enzymes. A: FaGH17A started to denature at ~40°C. B: FbGH30 started to denature at ~42°C. C: FaGH16A started to denature at ~55°C. Each measurement comprised ten times 5 s acquisitions. The temperature was increased at a rate of 0.25°C per min, from 25 to 80°C.



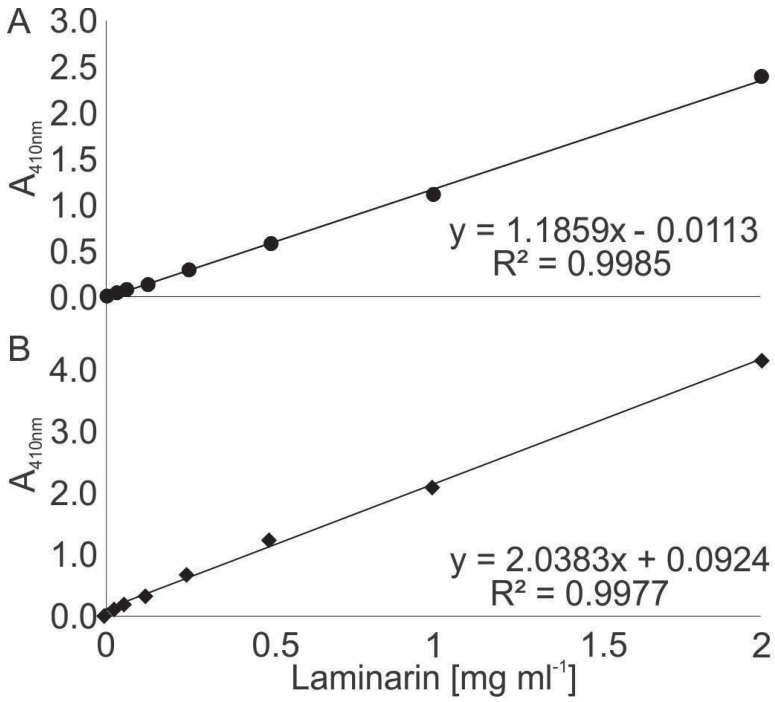
Supplementary FIG 2.1.2 Fluorophore assisted carbohydrate electrophoresis of hydrolysis products revealed the mode of action of FaGH16A, FaGH17A and FbGH30. Hydrolysis of laminarin was monitored over 60 min (To-T60). Untreated Glucose (Glc), Laminaribiose (L2), Laminaritriose (L3) and Laminaritetraose (L4) were added as a reference. **(A)** On a 4-20% acrylamide gel FaGH16A and FaGH17A exhibited an endo-acting activity by creating a ladder type profile. **(B)** On a 36% acrylamide gel FbGH30 produced only glucose, even during initial phases of the kinetic, which is typical for an exo-acting enzyme. 0.1% (w/v) laminarin was hydrolyzed by 100 nM (~5 µg ml^{-1}) purified enzyme at 37°C for 30 min in PBS buffer at pH 7.5. The reaction was stopped by heating for 10 min at 100°C. Derivatization and electrophoresis were performed as previously described (Jackson, 1990).



Supplementary FIG 2.1.3 Enzyme specificity tested with defined oligosaccharides substrates using thin layer chromatography. FaGH16A released laminaribiose and glucose from laminarin, laminaritriose and -tetraose. This enzyme also hydrolyzed mixed linked cellotetraose and mixed linked β-glucan from barley suggesting it can hydrolyze ß-1,3-linkages adjacent to β-1,4-linkages. FaGH17A showed a similar product profile of laminaribiose and glucose released from laminarin, laminaritriose and laminaritetraose. However, this enzyme did not cleave the mixed linked cellotetraose. FbGH30 released glucose from laminarin and hydrolyzed gentiobiose, a β-1,6 linked glucose disaccharide into glucose. In these experiments 0.1% (w/v) of the substrates were hydrolyzed by 100 nM (~5 µg ml^{-1}) purified enzyme at 37°C for 30 min in PBS buffer at pH 7.5.



Supplementary FIG 2.1.4 Yield of laminarin enzyme hydrolysis (~51%) compared to total acid hydrolysis (100%). 0.1% (w/v) of laminarin was hydrolyzed either by 100 nM (~5 µg ml^{-1}) purified enzyme at 37°C for 30 min in MOPS buffer at pH 7.5. Acid hydrolysis was conducted with 1 M HCl for 24 h at 100°C.



Supplementary FIG 2.1.5 PAHBAH calibration curves of laminarin hydrolysis. Comparison of acid hydrolysis **(A)** and enzyme hydrolysis using FaGH17A and FbGH30 **(B)**.

2.2 Manuscript II: Adaptive mechanisms that provide competitive advantages to marine Bacteroidetes during microalgal blooms

Frank Unfried[1,2,3], Stefan Becker[2,4], Craig S. Robb[2,4], Jan-Hendrik Hehemann[2,4], Stephanie Markert[1,3], Stefan E. Heiden[1], Tjorven Hinzke[1,3], Dörte Becher[3,5], Greta Reintjes[2], Karen Krüger[2], Burak Avci[2], Lennart Kappelmann[2], Richard L. Hahnke[2], Tanja Fischer[2], Jens Harder[2], Hanno Teeling[2], Bernhard Fuchs[2], Tristan Barbeyron[6], Rudolf Amann[2] and Thomas Schweder[1,3]

[1] Pharmaceutical Biotechnology at the University of Greifswald, Germany
[2] Max-Planck-Institute for Marine Microbiology, Bremen, Germany
[3] Institute of Marine Biotechnology, Greifswald, Germany
[4] MARUM – Center for Marine Environmental Sciences at the University of Bremen, Germany
[5] Institute for Microbiology at the University of Greifswald, Germany
[6] Station Biologique de Roscoff, France

Under review at: The ISME Journal

Contributions to the manuscript (% of *SB's* contribution to the total workload):
TS, RA and JHH designed the study. FU, HT, BA, LK, SH, KK, CSR, BF and TB performed computational analyses. *SB* (20%), FU, SM, DB, GR, TF and JH were involved in experimental work and the respective data analysis. *SB* cloned, purified and crystallized the enzymes and performed the functional protein studies. TS, JHH and RA wrote the manuscript, *SB* (5%) contributed minor parts of the text and figures.

2.2.1 Abstract

Polysaccharide degradation by heterotrophic microbes is a key process within Earth's carbon cycle. Here, we use environmental proteomics and metagenomics in combination with cultivation experiments and biochemical characterizations to investigate the molecular details of *in situ* polysaccharide degradation mechanisms during microalgal blooms. For this, we use laminarin as a model polysaccharide. Laminarin is an ubiquitous marine storage polymer of marine microalgae and is particularly abundant during phytoplankton blooms. In this study, we show that highly specialized bacterial strains of the Bacteroidetes phylum repeatedly reached high abundances during North Sea algal blooms and dominated laminarin turnover. These genomically streamlined bacteria of the genus *Formosa* have an expanded set of laminarin hydrolases and transporters that belonged to the most abundant proteins in the environmental samples. In vitro experiments with cultured isolates allowed us to determine the functions of in situ expressed key enzymes and to confirm their role in laminarin utilization. It is shown that laminarin consumption of *Formosa* spp. is paralleled by enhanced uptake of diatom-derived peptides. This study reveals that genome reduction, enzyme fusions, transporters and enzyme expansion, as well as a tight coupling of carbon- and nitrogen metabolism provide the tools, which make *Formosa* spp. so competitive during microalgal blooms.

2.2.2 Introduction

Phytoplankton blooms produce large quantities of beta-glucans, such as laminarin, a soluble β-1,3-glucan with β-1,6 side chains. The breakdown of these polysaccharides by heterotrophic microbes is a central part of the marine carbon cycle. Diatoms alone are estimated to produce about 5 to 15 Gt of laminarin per year as their storage compound, making it a major food resource for heterotrophic marine organisms (Alderkamp, van Rijssel, et al., 2007). Bacterial laminarinase activities are abundant in ocean surface waters, but also within deeper parts of the water column and in sediments (Arnosti et al., 2005; Keith and Arnosti, 2001). This suggests laminarin-degrading bacteria and their laminarinases are common across the oceans. How bacteria compete for this abundant labile energy substrate is therefore of relevance for a better understanding of the marine carbon cycle. Although partially studied with model organisms in the laboratory (Labourel et al., 2014, 2015; Xing et al., 2015; Kabisch et al., 2014), the enzymes used for laminarin degradation by microbes in the wild remain largely unknown or uncharacterized.

For complete degradation of one polysaccharide, microbes must have an adapted glycolytic pathway that contains multiple enzymes, which individually address each of the different glycosidic linkages and structural compositions present in the macromolecule. The genes of glycan-degrading pathways cluster in operons named polysaccharide utilization loci (PUL). Recent works suggest that each polysaccharide requires a corresponding PUL (Grondin et al., 2017). Horizontal gene transfer, vertical inheritance and gene loss distribute PULs

asymmetrically among genomes of microbes, creating the molecular basis for polysaccharide resource partitioning (Hehemann et al., 2010, 2016, 2017). This might explain the occurrence of diverse bacterial communities in the human gut (Ndeh et al., 2017; Cockburn and Koropatkin, 2016; Kaoutari et al., 2013; Martens et al., 2011) or in the oceans, whose members rely on different degradation products of the same polysaccharide to co-exist (Needham and Fuhrman, 2016; Teeling et al., 2012, 2016; Buchan et al., 2014). However, it remains unclear whether the degradation of complex carbohydrates is a community effort or mainly driven by highly specialized individual strains. Furthermore, how microbes effectively compete for the same polysaccharide resource, such as the abundant laminarin, is currently unknown.

In previous studies, we reported the high abundance (up to 24% of all bacteria) of the flavobacterial genus *Formosa* during diatom-dominated spring blooms off the North Sea island Helgoland (Teeling et al., 2012, 2016). Furthermore, high laminarin concentrations were measured at the same sampling site (Stefan Becker et al., 2017). Together, these findings suggested that *Formosa* spp. are prominent candidates for the recycling of laminarin during spring microalgae blooms.

In this study, we explored molecular strategies, which provide competitive advantages to the genus *Formosa* during microalgal blooms in general and for laminarin utilization in particular. We examined two strains, *Formosa* Hel3_A1_48 (referred to as strain A) and *Formosa* Hel1_33_131 (strain B), both of which were isolated from the same sampling location (Hahnke et al., 2015), and which are representative of two distinct taxonomical clades found during phytoplankton blooms (Chafee et al., 2017). The combination of high-resolution metaproteomics and metagenomics of spring bloom water samples with the detailed proteomic and biochemical characterization of the respective PUL in a cultured model strain (*Formosa* B) allowed us to show that a specialized enzyme repertoire represents one of the adaptive mechanisms that provide a competitive advantage in substrate exploitation. Using laminarin as a model substrate, we demonstrate how a microalgal glycan resource can promote the enrichment of individual dominating taxa from an initially diverse microbial community with similar metabolic functions. Our data indicate that *Formosa* B tightly couples glycan utilization with the uptake of nitrogen compounds. This suggests that a balanced carbon and nitrogen diet is required for competitive laminarin utilization during phytoplankton blooms.

2.2.3 Materials and Methods

2.2.3.1 Growth experiments and physiological characterization

The investigated strains *Formosa* sp. Hel1_33_131 (*Formosa* strain B) and *Formosa* sp. Hel3_A1_48 (*Formosa* strain A) were isolated by dilution cultivation during a spring and a summer phytoplankton bloom, respectively, from surface water near the North Sea island Helgoland in the German Bight (Hahnke et al., 2015). Growth experiments were performed in a

modified HaHa medium (Hahnke et al., 2015)(with 0.1 g L^{-1} peptone, 0.1 g L^{-1} casamino acids, 0.1 g L^{-1} yeast extract, 200 µM NH_4Cl, and 16 µM KH_2PO_4) with defined carbon sources as substrates at 12°C during gentle shaking at 55 rpm. For the proteome analyses, described below, D-glucose and laminarin (L9634, Sigma-Aldrich Chemie GmbH, Taufkirchen, Germany) were used as carbon sources (concentrations: 2 g L^{-1}). In addition, the utilization of chitin (SAFSC9213, VWR) was tested in this medium and these cultures were used as a control condition for the in vitro proteome analyses with glucose and laminarin. All growth experiments were carried out in triplicates. Cells were harvested by centrifugation (15 min; 9500 × g; 4°C), and the resulting pellets and supernatants were stored at -80°C until use.

2.2.3.2 Genome sequencing, assembly and annotation

For genome sequencing of the strains *Formosa* A (Hel3_A1_48) and B (Hel1_33_131) DNA was extracted according to the protocol of (Zhou et al., 1996). Sequencing was performed at LGC Genomics (Berlin, Germany) using the 454 GS FLX Ti platform (454 Life Sciences, Branfort, CT, USA) using standard shotgun libraries. Draft genomes were assembled with Newbler v2.6 for Hel3_A1_48 from 640,093 reads (406,983,286 bp) and for Hel1_33_131 from 636,323 reads (410,253,204 bp), yielding 2,025,184 bp (77 contigs) and 2,727,763 bp (61 contigs), respectively. The remaining gaps were closed by PCR and Sanger sequencing, yielding circular assemblies of 2,016,454 bp for Hel3_A1_48 (*Formosa* A) and 2,735,158 bp for Hel1_33_131 (*Formosa* B). Gene prediction and annotation (including the phylogeny-guided CAZyme annotations provided in Supplementary TAB 2.2.2) were performed as described previously (Mann et al., 2013). Further bioinformatic analyses are described in supplementary information. Annotated genome sequences were submitted to NCBI's GenBank with the accession numbers CP017 259.1 for *Formosa* sp. Hel3_A1_48 (*Formosa* strain A) and CP017260.1 for *Formosa* sp. Hel1_33_131 (*Formosa* strain B).

2.2.3.3 Proteome analyses

The soluble intracellular proteome, the enriched membrane-associated proteome and the soluble extracellular proteome was characterized from exponentially growing cells of *Formosa* strain B. Details of the protein extraction and subproteome enrichment can be found in supplementary information.

Peptides were subjected to a reversed phase C18 column chromatography on a nano ACQUITY UPLC (Waters Corporation, Milford, MA, USA) and separated as described by (Otto et al., 2010). Mass spectrometry (MS) and MS/MS data were recorded using an online-coupled LTQ-Orbitrap classic mass spectrometer (Thermo Fisher Scientific Inc., Waltham, MA, USA). We searched MS spectra against a target-decoy protein sequence database including sequences of *Formosa* B (Hel1_33_131) and of common laboratory contaminants.

Data analysis was performed with the Andromeda engine integrated in MaxQuant (Cox and Mann, 2008) and peptide level FDR (false discovery rate) was set to 0.01 (1%). Only proteins that could be detected in at least 2 out of 3 replicates were counted as identified. The automatically calculated iBAQ values (intensity-based absolute quantification; i.e. peak area divided by the sum of all theoretical peptides) were used to manually calculate riBAQ values (relative iBAQ; giving the relative protein abundance in % of all proteins in the same sample, (JB Shin et al., 2013)) for semi-quantitative comparisons between samples from different nutrient conditions. Tests for differential expression were performed using Perseus (Tyanova et al., 2016) v. 1.6.1.1 with Welchs t-test (permutation-based FDR 0.05).

The mass spectrometry proteomics data are available through the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (Vizcaíno et al., 2012) with the dataset identifier PXD007934 (Reviewer account details: Username: reviewer19024@ebi.ac.uk Password: QlZTbLvW).

2.2.3.4 Biochemical enzyme characterizations

The cloning of the FbGH17A gene (locus tag FORMB_24720) and the FaGH17B gene (locus tag FORMB_24740) is described in supplementary information. Cloning of the FbGH30 gene is described by (Stefan Becker et al., 2017). Detailed information on the overexpression, enzyme refolding and purification of these proteins can be found in Supplementary Information. For enzyme characterizations Laminarin from *Laminaria digitata* (0.1% [w/v]; Sigma) was hydrolyzed over the course of 60 min at 37°C with 100 nM purified enzyme (~5 µg mL^{-1} of FbGH30, FbGH17A, or FbGH17B) in 50 mM MOPS buffer at pH 7. The preparation and purification of debranched laminarin as well as the determination of kinetic parameters of the three enzymes FbGH30, FbGH17A and FbGH17B acting on native and debranched laminarin is explained supplementary information. High performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) was applied for qualitative product analysis of the enzyme reactions (see supplementary information).

2.2.3.5 Protein crystallization and structure solution

Crystals of FbGH17A were obtained by hanging drop vapor diffusion of the protein with 12.3 mg mL^{-1} mixed 1:1 with a "well solution" (0.03 M $MgCl_2$, 0.1 M MOPS (pH 7), 9% PEG8K) supplemented with 15% ethylene glycol (Supplementary FIG 2.2.10). The crystals were cryoprotected prior to freezing in the "well solution" supplemented with ethylene glycol to a final concentration of 30%. Crystals were frozen by flash freezing in liquid nitrogen in nylon loops. X-ray diffraction data were collected at the DESY P11 beamline. The structure was solved by molecular replacement using PHASER in the phenix suite (PD Adams et al., 2010; McCoy et al., 2007) using the pdb 4wtp (Qin et al., 2015). The model was built using BUCCANEER (Cowtan,

2006) and Coot, refined in REFMAC5 (Murshudov et al., 2011), and validated and deposited with pdb code 6FCG.

2.2.4 Results

2.2.4.1 Genome properties and phylogeny

We sequenced and annotated the genomes of the *Formosa* strains A and B. Both have a single chromosome with a GC content of 36.4% and 36.6%, respectively. With 2,016,454 bp (strain A) and 2,735,158 bp (strain B) they possess small genomes compared to other marine polysaccharide degrading *Flavobacteriia* (Barbeyron et al., 2016; Kabisch et al., 2014; Mann et al., 2013; Bauer et al., 2006). Strain A has 1,913 predicted genes including 1,866 CDS, 40 tRNAs genes and 2 rRNA operons (identical 5S, 16S, 23S rRNA genes), while the strain B genome encodes 2,675 predicted genes with 2,628 CDS, 39 tRNAs genes and 2 rRNA operons (identical 5S, 16S, 23S rRNA genes).

Phylogenetic analyses based on 16S rRNA gene sequences indicate that the *Formosa* strains A and B are representatives of two previously uncultured clades of the genus *Formosa*. These occur not only in the North Sea, but also in surface waters from coastal and open ocean sites throughout the world (Supplementary FIG 2.2.1). Of the 33 full-length *Formosa* 16S rRNA sequences obtained from 2009 spring bloom bacterioplankton (Teeling et al., 2012), 16 were more than 99% identical to *Formosa* sp. Hel1_33_131 (*Formosa* B) (Supplementary FIG 2.2.1).

2.2.4.2 In situ abundance and relevance of *Formosa* strain A and B

We investigated the in situ abundance of the *Formosa* strains A and B by recruiting *Formosa* reads from the 44 metagenomes of the years 2009–2012 from Helgoland bacterioplankton samples (Teeling et al., 2016). At the ≥95% average nucleotide identity (ANI) threshold, the strain A and B genomes recruited up to 0.28% and 2.94% of individual metagenomic reads in 2009, 0.04% and 0.99% for 2010, 0.03% and 0.98% for 2011, and 0.02% and 0.21% for 2012 (Supplementary TAB 2.2.1) respectively. The mapped reads covered up to 91%, 99%, 97% and 94% of the strain B genome from 2009 to 2012, respectively, and only up to 58% of the strain A genome in 2012 (FIG 2.2.1 and Supplementary TAB 2.2.1). This suggests that strain B was recurrent and abundant during the spring bloom events, while strain A was likely more representative for late summer blooms reaching highest abundances of mapped reads in September 2009. Reads mapped to the *Formosa* strain B genome with 70-93% ANI suggest the presence of other closely related *Formosa* spp. during the spring blooms of 2009 to 2012 that reached up to 6.84%, 1.2%, 5.1% and 1.8% of the metagenome reads, respectively (FIG 2.2.1). Altogether, these results indicate that strain B is one of the representatives of the recurrent *Formosa* clade during North Sea spring microalgae blooms.

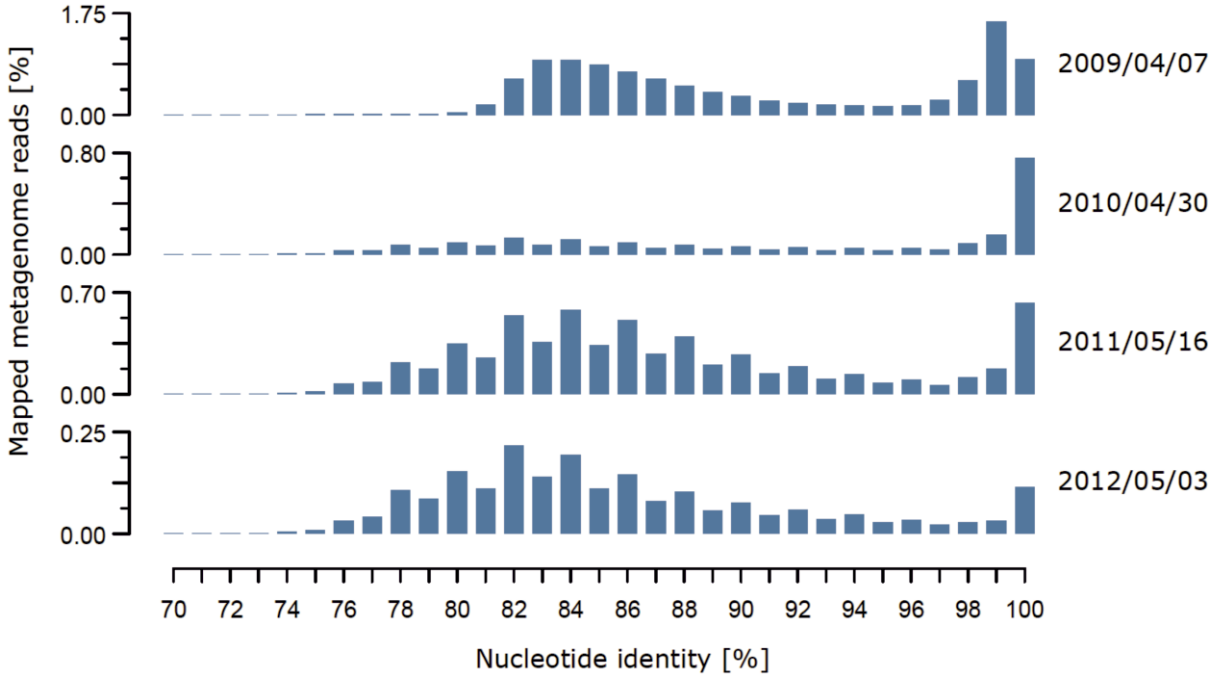


FIG 2.2.1 Relevance of *Formosa* strain B and related species during four spring phytoplankton blooms indicated by the percentage of metagenomics reads mapped at different nucleotide identities. Reads recruited at ≤ 93% nucleotide identity represent other *Formosa* spp. that are abundant during the spring bloom events at Helgoland from 2009 until 2012. Dates on the right indicate the four metagenomes (i.e., time points), which produced highest mapping coverage with the *Formosa* strain B genome in their respective years. For a summary of all 44 metagenomes (up to 18 time points per year) and their mapping results see Supplementary TAB 2.2.1.

2.2.4.3 *Formosa* genomes encode PULs for laminarin degradation

Genome annotation suggested that the *Formosa* strains A and B are specialized polysaccharide degraders which concentrate their genetic potential on a small set of sugars. *Formosa* A contains 7 PULs (Supplementary FIG 2.2.2) and 28 glycoside hydrolases (Supplementary TAB 2.2.2), while *Formosa* B contains 6 PULs (Supplementary FIG 2.2.3) and 21 glycoside hydrolases (Supplementary TAB 2.2.2). This is a very small repertoire, even compared to other marine Bacteroidetes isolated from algal blooms (Xing et al., 2015). These small carbohydrate-active enzyme (CAZyme) repertoires contrast particularly with those of generalist polysaccharide degraders isolated from macroalgae, such as *Formosa agariphila* (Mann et al., 2013), which have broad polysaccharide degrading capacity. *F. agariphila* has, for example, a genome size of 4.48 Mbp, and 84 glycoside hydrolases in 13 PULs (Supplementary TAB 2.2.2).

To functionally characterize laminarin-specific PULs of the *Formosa* strains A and B, we searched the genomes for enzymes belonging to known laminarinase-containing families (Supplementary TAB 2.2.2). We found putative laminarinases of the families 217 GH30, GH17, GH16, GH3 and GH2 located in close proximity to TonB-dependent receptors (TBDR) and SusD-like proteins, which are indicators of polysaccharide utilization loci (PULs) (Tang et al., 2012; Sonnenburg et al., 2010). Our results suggest that there are three putative laminarin-specific genomic PULs in both *Formosa* strains (Supplementary FIG 2.2.2–5).

The laminarin PULs 1 and 2 of *Formosa* A and B revealed a high synteny with PULs from other bacteroidetal strains (Supplementary FIG 2.2.4) from North Sea surface water (Panschin et al., 2016; Hahnke et al., 2015; Bauer et al., 2006). This points to a potential for competition between those groups, but also suggests that this part of the laminarin utilization machinery is highly conserved. However, the *Formosa* B PULs 1 and 2 are enlarged with laminarinases and transporters that are partially not present in the other bacteria. Moreover, the entire PUL 3 of *Formosa* B is missing in these other strains (Supplementary FIG 2.2.5). Instead, *Formosa* B's PUL 3 shows syntenies to PULs of other marine *Flavobacteriia*, which do, however, not possess the PULs 1 and 2 (Supplementary FIG 2.2.5). All three *Formosa* B PULs were completely covered by metagenomic contigs of the spring bloom in 2009 and 2010 (FIG 2.2.2A and Supplementary TAB 2.2.3-4), and partially covered in the metagenomes of 2011 and 2012 (Supplementary TAB 2.2.4). This illustrates the strong selection pressure imposed by laminarin on this pathway during four consecutive annual spring phytoplankton blooms in the North Sea.

2.2.4.4 Metaproteomic identification of *Formosa*-specific enzymes and transporters during microalgal blooms

We examined the presence of polysaccharide degradation- and consumption-related proteins of the *Formosa* strains A and B in the in situ metaproteomes of spring blooms in 2009 and 2010 (Supplementary TAB 2.2.5). The proteome analysis of the planktonic bacterial fraction sampled during the spring bloom on April 7th, 2009 uncovered 46 proteins from *Formosa* strain A and 361 proteins from *Formosa* strain B. Remarkably, several marker proteins from the putative laminarin specific *Formosa* B PULs were highly abundant 241 (FIG 2.2.2B and Supplementary TAB 2.2.5A) in the metaproteome samples. This analysis identified 13 proteins of the PULs 1, 2 and 3 (see also Supporting Information) and thus indicated that a significant proportion of *Formosa* B's putative laminarin PULs were expressed in situ during the spring bloom in 2009. Although the metaproteome analysis of 2010 uncovered fewer proteins from both *Formosa* strains, three marker proteins of PUL 1 from *Formosa* B were detected in the environmental samples (supplementary information and Supplementary TAB 2.2.5B).

Besides glycoside hydrolases and laminarin-specific transporter proteins, we also observed several *Formosa* B proteins in the environmental metaproteome samples of 2009, which are involved in the central catabolism of the monosaccharide glucose, the product of laminarin hydrolysis (FIG 2.2.2B, supplementary information and Supplementary TAB 2.2.5A). This includes nearly all glycolytic enzymes as well as a putative glycogen synthase of *Formosa* B. These data indicate that the *Formosa* B strain substantially contributed to laminarin degradation and turnover during a diatom driven phytoplankton bloom.

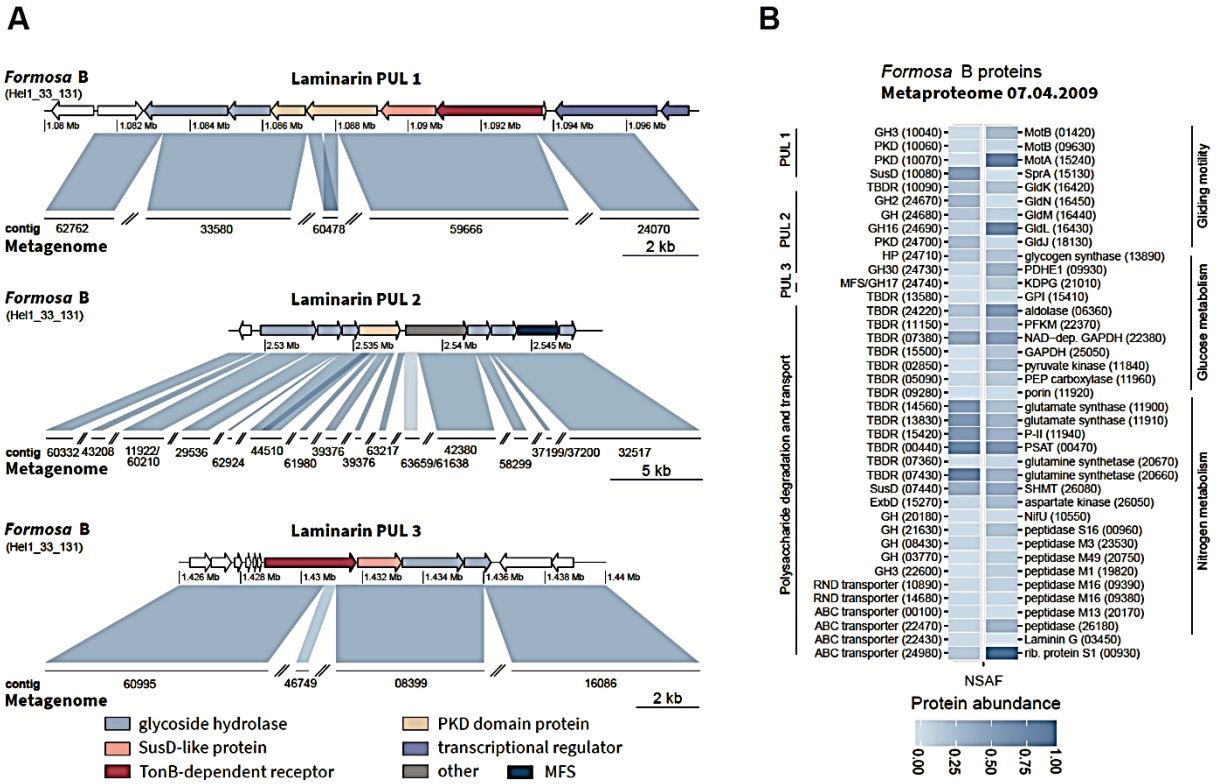


FIG 2.2.2 Relevance Detection of *Formosa* strain B laminarin PULs in the Helgoland spring bloom metagenome and metaproteome in 2009 (Teeling et al., 2012). **(A)** Synteny between the laminarin PULs of *Formosa* sp. Hel1_33_131 and partial PUL sequences in the metagenomes from 07.04.2009. The sequence comparisons were performed with Bl2seq (BLASTn, E-value 1e-5). Sequence similarities are depicted by blue hues for direct comparisons. Darker colors correspond to higher identities. **(B)** Heatmap of the relative abundance of *Formosa* strain B proteins (displayed as normalized spectral abundance factor values, NSAF%) detected in the metaproteome from 07.04.2009. Displayed are selected proteins, which likely play a role in polysaccharide or protein utilization. A highly abundant ribosomal protein of *Formosa* strain B is also displayed as a reference to illustrate the high abundance of polysaccharide utilization-specific proteins during the bloom condition. Gene locus tag numbers are given in parenthesis. TBDR, TonB-dependent receptor; SusD, SusD-family protein; PKD, PKD-domain containing protein; GH, glycoside hydrolase; MFS, major facilitator superfamily; HP, hypothetical protein.

2.2.4.5 Laminarin elicits the expression of specific polysaccharide utilization loci in *Formosa*

The ability of *Formosa* B to quickly react and take up laminarin could be directly visualized using FLA laminarin incubations. *Formosa* B accumulated high amounts of FLA-laminarin after just 5 min of incubation (FIG 2.2.3A). Additionally, the halo-like staining pattern showed that the FLA-laminarin was imported into the periplasm of the cells by a "selfish" uptake mechanism (Reintjes et al., 2017; Cuskin et al., 2015). Selfish substrate uptake is dependent on the presence of SusCD-like transporters and secures an enrichment of substrate in the periplasmic space without diffusive loss (Reintjes et al., 2017).

To elucidate the metabolism of *Formosa* B on laminarin and to verify whether laminarin specifically controls the expression of the genomically and metaproteomically detectable PULs we performed cultivation experiments with this bacterium with purified laminarin as growth substrate. Growth curves of *Formosa* B in HaHa medium with laminarin, glucose and only protein

extracts are exemplarily shown in FIG 2.2.3B. We used proteomics to record the global protein expression patterns with these substrates. We investigated (i) the soluble intracellular proteome, (ii) the enriched membrane proteome and (iii) the extracellular proteome (see supplementary information and Supplementary TAB 2.2.6A-C). These comparative analyses showed that although glucose is the monomer of laminarin, the utilization of either carbon source led to quite different proteomic signatures in different functional protein categories, such as in nucleotide, lipid and coenzyme metabolism as well as in carbohydrate metabolism and transport (Supplementary FIG 2.2.6). About 100 genes were significantly upregulated or exclusively induced by laminarin in *Formosa* B (Supplementary TAB 2.2.7). Of all three substrates, laminarin elicited the strongest expression of the three laminarin PULs of *Formosa* B (FIG 2.2.3C), which is indicative of specific and tightly controlled expression. The SusD-like protein (FORMB_10080) of PUL 1 and the GH16 (FORMB_24690) of PUL 2 were exclusively expressed with laminarin but not with the other substrates. Furthermore, the expression of PUL 3 was exclusively induced by laminarin and not detectable with glucose or only peptone (FIG 2.2.3C). The specific response of the *Formosa* PULs to laminarin and not to glucose implies that the three-dimensional structure of laminarin might be the key to induce the expression of these PULs.

2.2.4.6 Biochemical analysis of laminarinases expressed by *Formosa* spp.

To functionally characterize further PUL-encoded proteins and to map the laminarin degradation pathway, we cloned and biochemically analyzed laminarinases that could be identified as abundant proteins in the in situ (spring bloom metaproteome) and in vitro (*Formosa* B) proteomes. We cloned and examined the genes encoding FbGH17A (locus tag: FORMB_24720), FaGH17B (locus tag: FORMB_24740) and FbGH30 (locus tag: FORMB_24730). As all three proteins are encoded in a single gene cluster, we hypothesized that these enzymes might work together in spatial proximity. To test this hypothesis, we conducted a series of biochemical experiments, which revealed that the FbGH30 enzyme hydrolyzed the β-1,6-linked glucose side chains of laminarin (K_M: 3.1±0.2 mM and K_{cat}/K_M: 21124 $M^{-1}s^{-1}$) (FIG 2.2.4A), while it was inactive on the debranched substrate. The enzyme FbGH17A hydrolyzed both the debranched laminarin product of FbGH30 and the native laminarin, although with a markedly higher specific activity on the debranched product (K_M: 1.6±0.1 mM; K_{cat}/K_M: 36056 $M^{-1}s^{-1}$) than on laminarin itself (K_M: 4.3±0.1 mM; K_{cat}/K_M: 25744 $M^{-1}s^{-1}$) (FIG 2.2.4B). Preference for debranched laminarin was even more pronounced with FbGH17B, which only hydrolyzed debranched laminarin (K_M: 2.6±0.3 mM; K_{cat}/K_M: 30803 $M^{-1}s^{-1}$) (FIG 2.2.4C) and was inactive on the branched form.

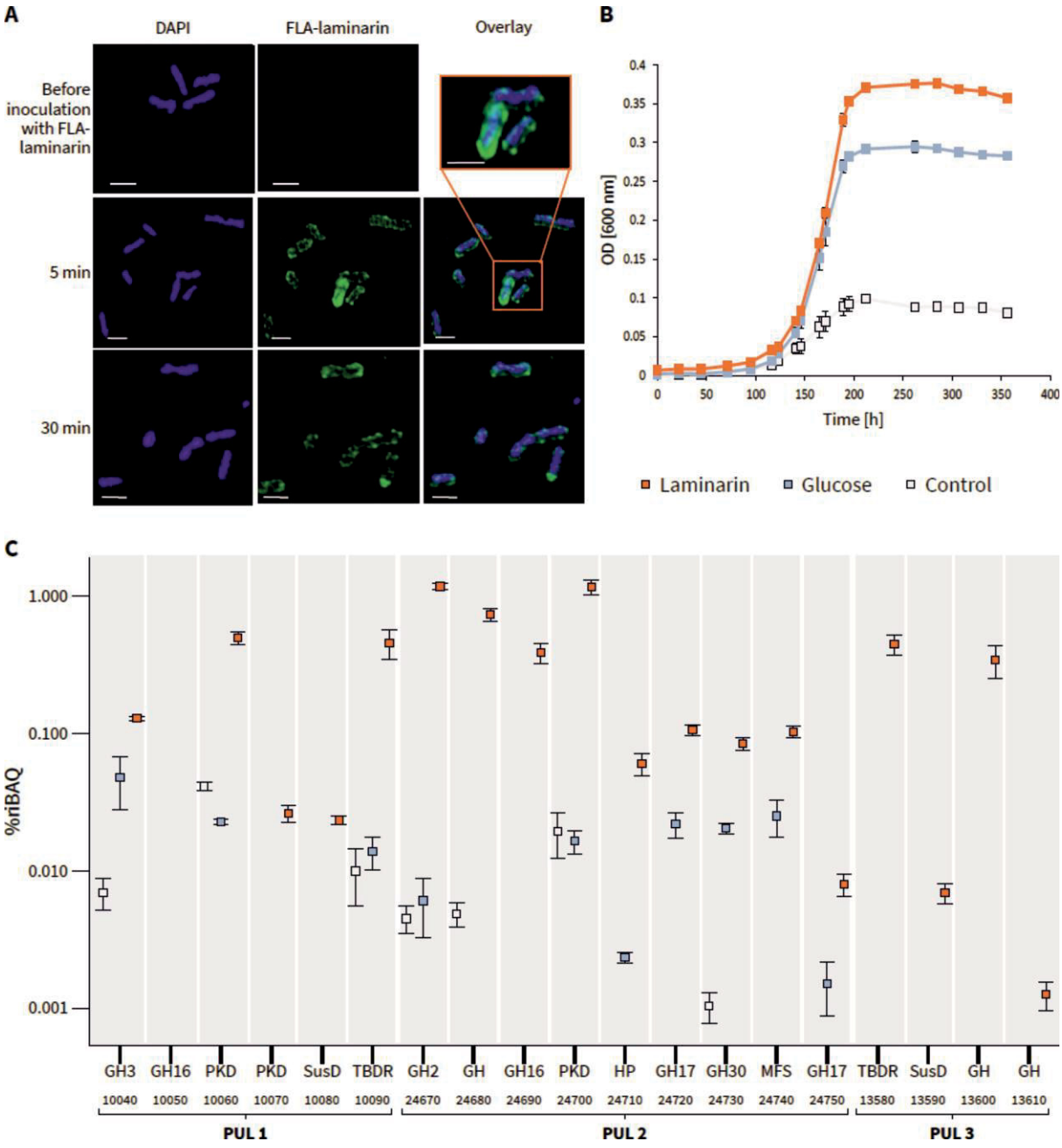


FIG 2.2.3 Laminarin utilization of *Formosa* B. **(A)** SR-SIM of *Formosa* B cells before inoculation with FLA-laminarin and 5 and 30 min after incubation with FLA-laminarin. Images show cell staining by DAPI (left, blue), FLA-laminarin (middle, green) and an overlay showing both FLA-laminarin staining and DAPI (right). Scale bar = 1 µm. **(B)** Growth curves at 12°C in modified HaHa_100V medium (Hahnke et al., 2015) with 2 g L^{-1} laminarin or 2 g L^{-1} glucose. The "control" culture contained only 0.1 g L^{-1} peptone, 0.1 g L^{-1} yeast extract and 0.1 g L^{-1} casamino acids but no additional carbon sources. **(C)** Expression profile and gene organization of the laminarin utilization PULs 1-3 in *Formosa* B. Relative protein abundances (in %riBAQ) of PUL-encoded proteins detected in the membrane protein fractions of cultures grown on laminarin (orange), glucose (blue) and chitin (control, grey) are shown (for riBAQ values see Supplementary TAB 2.2.6A and 7). Putative protein functions (e.g. GH3) and the respective locus tags (e.g. 10040) are indicated. The squares represent the mean values of the replicates for every protein and each substrate. The error bars refer to the standard error of the mean (SEM; standard deviation/√number of replicates). Proteins that could be detected in at least 2 out of 3 independent biological replicates are shown (for individual replicate numbers see Supplementary TAB 2.2.6A).

To elucidate how these enzymes work together in successive laminarin degradation, we used HPLC-PAD analyses. The data indicated an enzymatic functional cascade in three steps: The exo-

acting β-1,6-glucosidase FbGH30 removes the glucose side chains from laminarin. The endo-acting β-1,3-glucan hydrolase FbGH17A degrades the remaining debranched laminarin into oligosaccharides. The exo-acting β-1,3-glucosidase FbGH17b processes these oligosaccharides into glucose (FIG 2.2.4D). FbGH17b is part of a multi-modular protein, which is encoded by a gene that also codes for an N-terminal MFS transporter, suggesting that hydrolysis and product uptake might be coupled. The MFS transporter contains 12 transmembrane-spanning helices (as predicted by the phyre server), with the last C-terminal helix and the attached GH17 domain located in the cytoplasm (FIG 2.2.4E), where oligosaccharides transported through the MFS are thus likely to be cleaved into glucose. Blast analysis revealed that this fusion is common among marine *Flavobacteriia* suggesting that such multi-modular transporter-associated enzyme may be a conserved mechanism for boosting laminarin utilization.

In order to examine the molecular basis of substrate specificity the X-ray crystal structure of GH17A was solved (FIG 2.2.4F, see supplementary information). Compared to the monomeric GH17 structures, FbGH17A has significant insertions and is larger. The structure of GH17A allows for the deduction of the molecular basis of substrate specificity for the laminarinase. Based on the GH17 complexes obtained for *R. miehei* (Qin et al., 2015), a model was generated of a laminarin product involving 5 monomers bound to the catalytic groove, two on the aglycon side and three on the glycon side (FIG 2.2.4G). The reducing and non-reducing ends of the modelled glycan are free, suggesting the protein can act in the middle of the chain as expected for an endo-acting glycoside hydrolase (see supplementary information). This structural data supports the observation that GH17 activity on laminarin is bolstered by the action of the debranching enzyme GH30.

2.2.4.7 Laminarin stimulates the co-expression of selected peptidases and transporters

Formosa B encodes 69 peptidases in its relatively small genome (2.7 MB). Other laminarin-degrading marine Bacteroidetes like *Gramella forsetii* KT0803T (79 peptidases), *Polaribacter* sp. Hel1_85 (84 peptidases), *Jejuia pallidilutea* (58 peptidases), and *Flaviramulus ichthyoenteri* (63 peptidases) show a comparable number of peptidase genes although their genomes are around two-fold larger than that of *Formosa* B. Our proteome analysis of *Formosa* B revealed that 41 peptidases are expressed in the presence of laminarin (Supplementary TAB 2.2.9). Nine of these peptidases showed a significantly higher protein abundance on laminarin in the enriched membrane proteome, compared to glucose or the control culture or were exclusively expressed with laminarin (FIG 2.2.5A and Supplementary TAB 2.2.7). Eight of these *Formosa* B peptidases were also detected in the metaproteome of the 2009 spring bloom (FIG 2.2.2B and Supplementary TAB 2.2.5A). In addition, a putative peptide ABC transporter ATP-binding protein (FORMB_10920) and a putative oligopeptide permease ABC transporter protein (OppC; FORMB_20460) were detected, which showed a significantly higher abundance under laminarin

conditions (Supplementary TAB 2.2.7). This indicates a coupling of the peptide metabolism with laminarin utilization in *Formosa* B.

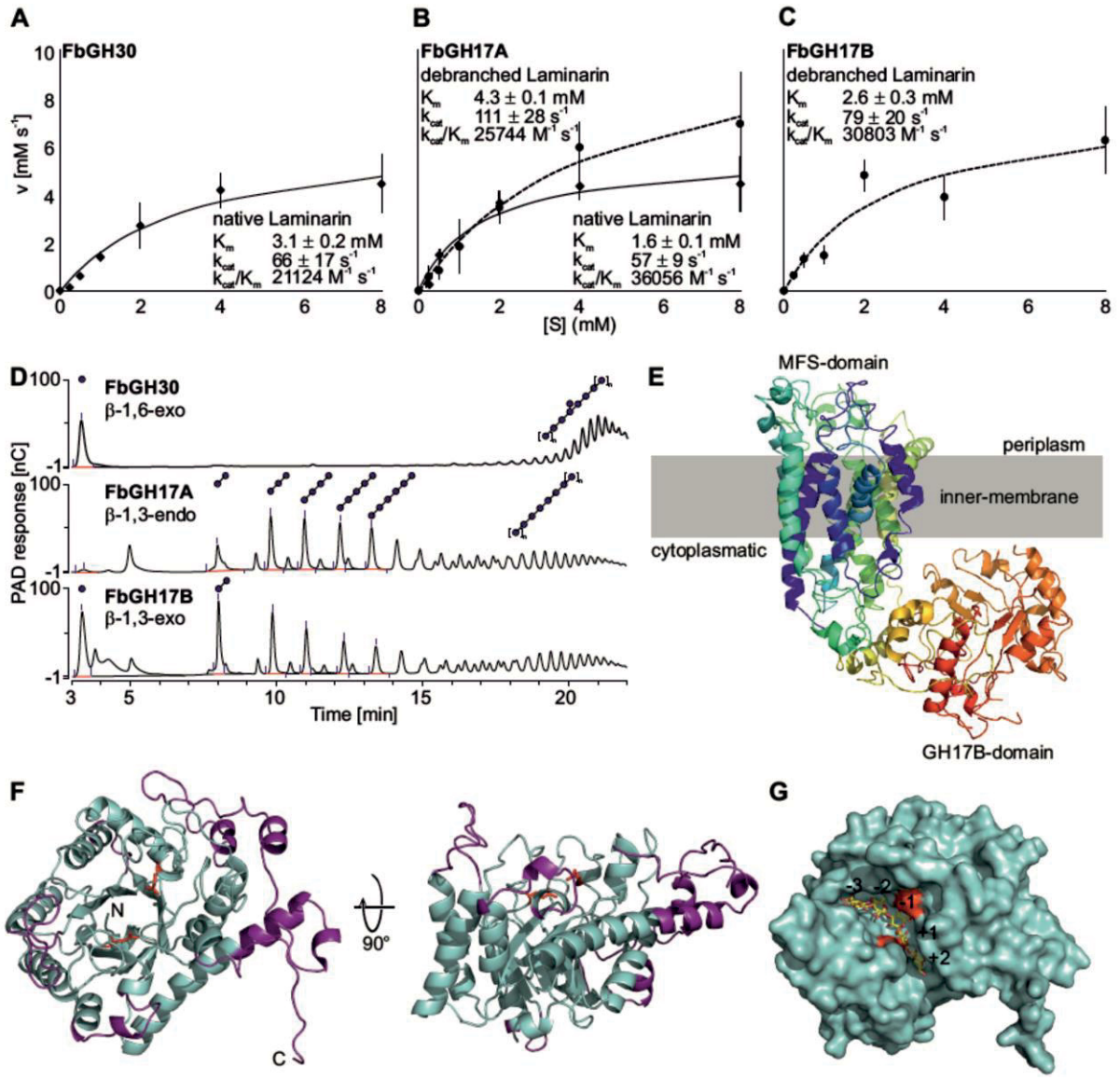


FIG 2.2.4 Biochemical characterization of different laminarinases from *Formosa* B. Michaelis-Menten-Kinetic **(A)** of FbGH30 on native laminarin, **(B)** of FbGH17A on native and debranched laminarin and **(C)** of FbGH17B on debranched laminarin. **(D)** Visualization of all three enzymatic activities was done using HPAEC-PAD. FbGH30 hydrolyzed native laminarin. After this debranching reaction, the laminarin was purified to remove glucose for the following steps. This debranched laminarin was used in the FbGH17A reaction. FbGH17B hydrolyzed the products of the previous FbGH17A reaction without any further purification in between. **(E)** 3D structure model of both the MFS-domain and the associated FbGH17B-domain and its arrangement within the inner-membrane. The modelling was performed using Phyre2. **(F)** The overall structure of FbGH17A is displayed in cyan with the additions colored purple. Highlighted in red are the catalytic residues and the N- and C-termini are labeled. **(G)** A surface view of FbGH17A with a modelled substrate complex shown as sticks in yellow from a GH17 transferase of *Rhizomucor miehei* with laminaritriose and laminaribiose in the -3 to -1 and +1 to +2 subsites, respectively.

An exceptionally high expression with glucose and laminarin was visible for a putative porin (FORMB_11920, 10% riBAQ; FIG 2.2.5B and C), an outer membrane protein, which was not

detectable in the control cultivations with peptone (Supplementary TAB 2.2.7). This *Formosa* B protein was also detected in the metaproteome analyses of the spring bloom 2009 (FIG 2.2.2B, Supplementary TAB 2.2.5A). The porin-encoding gene is located in an operon with a putative ammonium transporter and clusters with several genes involved in nitrogen metabolism, including two putative glutamate synthase genes and an additional supposed ammonium transporter (FIG 2.2.5C). All nitrogen metabolism-related genes in the direct vicinity of the porin-encoding gene were specifically induced by glucose and laminarin in peptone-containing cultures in comparison to the peptone-only control culture without these carbon sources (FIG 2.2.5C).

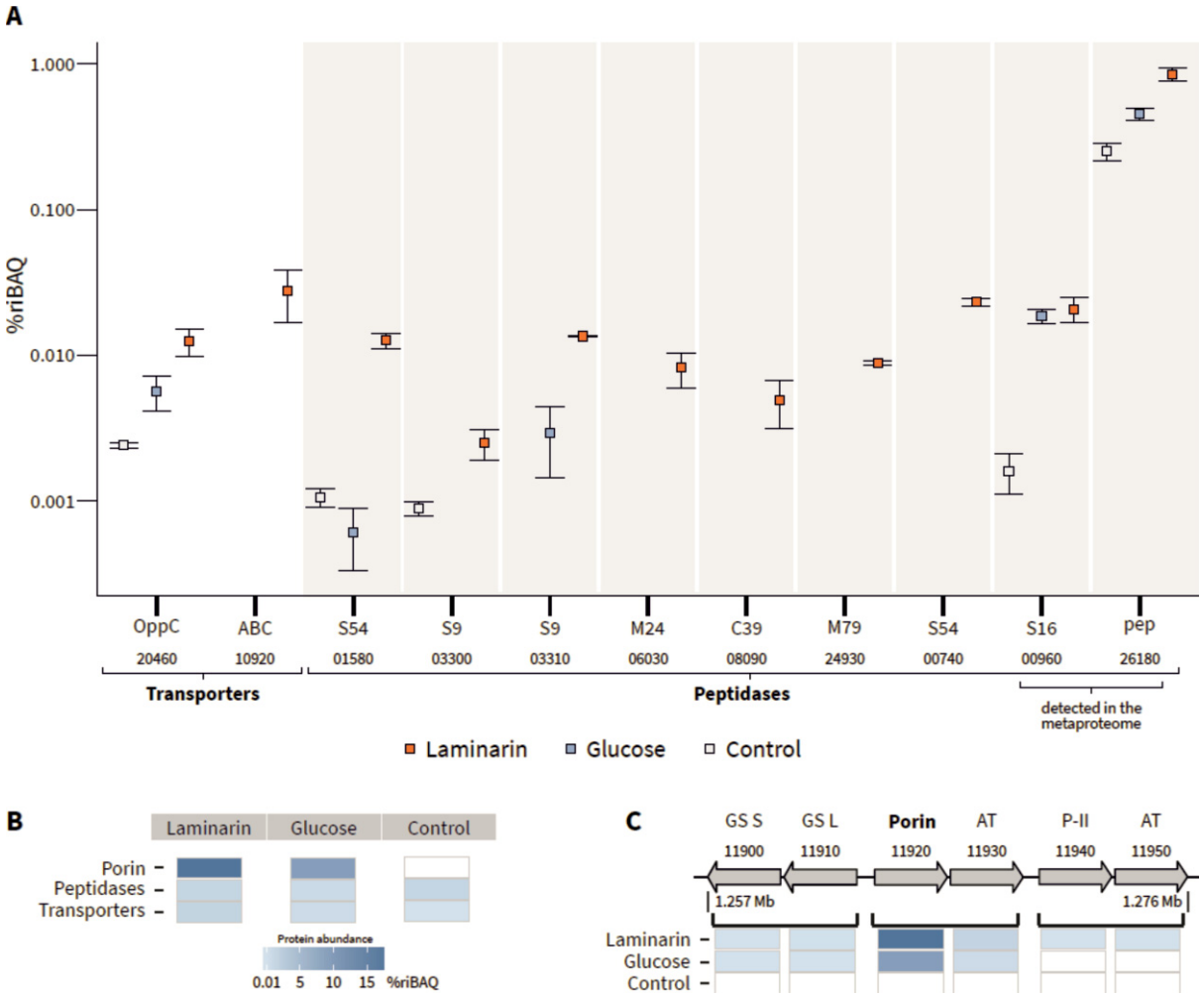


FIG 2.2.5 Differential expression of peptidases and putative peptide transporter proteins of *Formosa* B. **(A)** The two laminarin-induced peptidases shown on the right (00960, 26180) were also detected in the metaproteome of a spring bloom in 2009 (Supplementary TAB 2.2.5A). Relative protein abundances are depicted as % riBAQ values. The putative peptidase families and the respective locus tags are indicated. The squares represent the mean values of the replicates for every protein and each substrate. The error bars refer to the standard error of the mean (SEM; standard deviation/√number of replicates). Proteins that could be detected in at least 2 out of 3 independent biological replicates are shown (for individual replicate numbers see Supplementary TAB 2.2.6A). **(B)** Comparison of total protein abundances (in %riBAQ) of peptidases (Supplementary TAB 2.2.9), nitrogen-associated transporters and the porin in *Formosa* B under the three investigated substrate conditions (Supplementary TAB 2.2.7). **(C)** Genomic structure of the porin-encoding cluster and the expression patterns of the corresponding genes. Brackets indicate putative operons. Protein functions and the respective locus tags are indicated. GS S: glutamate synthase subunit S, GS L: glutamate synthase subunit L, AT: ammonium transporter, P-II: nitrogen regulatory protein P-II.

2.2.5 Discussion

This study provides detailed insights in the adaptations which make *Formosa* strains successful competitors in the early breakdown of organic matter during diatom blooms. Combining comparative in vitro and in situ proteogenomics with biochemical enzyme characterization reveals that the key to this process is the sensing and utilization of laminarin. Our data indicate that this polysaccharide is used both as a major source of energy and as a signal molecule inducing transporters and digestive enzymes to use also other compounds released from the lysis of diatom cells.

The two environmentally relevant *Formosa* strains examined in this study feature streamlined genomes which are significantly smaller as those of many other marine *Flavobacteriia*. With a lower number of total proteins to synthesize, *Formosa* A and B can dedicate a higher relative proportion of their genomic and proteomic resources to the digestion of laminarin. Their CAZyme repertoire is strongly reduced compared to versatile polysaccharide degraders such as *F. agariphila* (Mann et al., 2013) and *Zobellia galactanivorans* (Barbeyron et al., 2016), which were isolated from macroalgae. It is, however, similar to another member of North Sea spring bacterioplankton, *Polaribacter* sp. Hel1_33_49 (Xing et al., 2015). In contrast to macroalgae-associated laminarin-degrading bacteria, such as *Z. galactanivorans* (Groisillier et al., 2015), neither of the *Formosa* strains possesses a mannitol dehydrogenase, which indicates a specialization of *Formosa* A and B to chrysolaminarin. This type of laminarin lacks mannitol residues and is preferentially produced by diatoms.

We found a laminarin-specific proteome in *Formosa* B, which is not induced by the sugar monomer of this polysaccharide, glucose. A similar laminarin-specific control of gene expression was suggested for the marine flavobacterium *G. forsetii* (Kabisch et al., 2014). Interestingly, this induction in *Formosa* B includes not only the proteins required for laminarin uptake and utilization, but also peptidases and transporters for amino acid utilization. The *Formosa* cells, upon sensing of laminarin, thus appear to react in two ways: Firstly, they enhance the expression of outer membrane proteins to degrade and rapidly transport the energy molecule laminarin into their periplasm, utilizing the selfish polysaccharide uptake mechanism recently demonstrated for marine *Flavobacteriia* (Reintjes et al., 2017). Secondly, the expression of amino acid- and nitrogen metabolism-related proteins is increased to boost the recycling of nitrogen building blocks, which are required for rapid growth of *Formosa* bacteria and become available simultaneously with laminarin upon algal lysis.

Formosa strain B possesses an extended repertoire of laminarin-specific enzymes and transporters, which is larger than that of other laminarin-degrading bacteria such as *Polaribacter* sp. Hel1_33_49 (Xing et al., 2015) or *G. forsetii* (Kabisch et al., 2014). Our subproteome and bioinformatic analyses indicate that many laminarin-degrading enzymes of *Formosa* B are surface

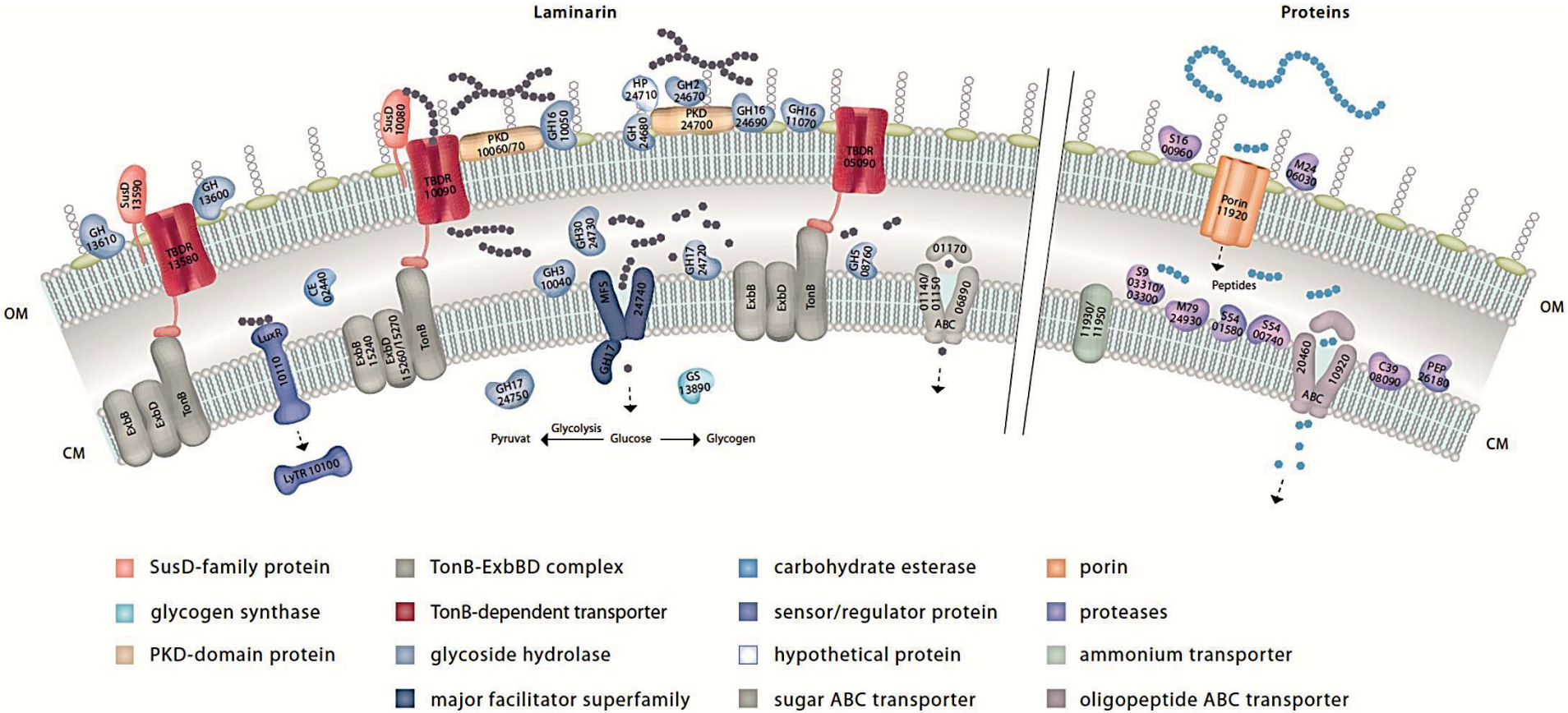


FIG 2.2.6 A tentative model of laminarin utilization pathways and of the proposed laminarin induced peptide uptake in *Formosa* B. Protein localizations were predicted *in silico* according to (Romine et al., 2011) and were deduced by subproteome analyses (Supplementary TAB2.2.7). Additional biochemical experiments are required to ascertain this model.

tethered or localized in the periplasmic space and in the cytoplasmic membrane, respectively. The different TBDR, laminarinases, transporters and additional enzymes combine complementary activities into an efficient laminarin disassembly line for degradation and uptake (FIG 2.2.6). The biochemical experiments presented here support the annotation of the conserved cluster of genes as encoding for a laminarin utilization pathway. Here, two enzymes that are likely residents of the periplasm are shown to work together towards the complete degradation of laminarin in a highly specific manner. The X-ray crystal structure of GH17A reveals the possible molecular determinants of substrate specificity and the propensity of the enzyme to be more active on unbranched laminarin.

The multi-modular protein FORMB_24740 (FbGH17b) combines a glycoside hydrolase (GH17) with a membrane spanning transport protein and may represent an adaptive mechanism for laminarin utilization. The integration of the transport and hydrolysis processes into a single protein could facilitate improved consumption of the sugars by increasing the activity of the cytosolic GH17. Increased activity would also reduce the necessary enzyme copy number, and thereby resource consumption for synthesis of this protein. To our knowledge, such a transporter-CAZyme-fusion has not been described for other bacteria as yet, but its conservation in nature suggests that this could provide a significant benefit.

The exceptionally strong accumulation of a putative porin in glucose- and laminarin-controlled cultures and the co-induction of the 12 surrounding genes of this porin-encoding genomic cluster, all of which play a role in nitrogen metabolism, could indicate a function of this transporter protein in the uptake of peptides as nitrogen and amino acid source. The highly abundant endo-acting proteases might degrade proteins released by lysed microalgae into peptides, which are then imported through the porin into the periplasm (FIG 2.2.6). An efficient capture of these peptides with a highly abundant porin system might be especially useful in the highly diffusive marine environment. With laminarin and glucose as easily metabolizable carbon sources, such a strategy could be crucial for a balanced carbon and nitrogen diet.

Our results show an extraordinary degree of specialization for the *Formosa* strains A and B, which could enable these marine Bacteroidetes to successfully compete for laminarin against a multitude of other laminarin-degrading microbes in bloom situations (Bennke et al., 2016; Xing et al., 2015; Cardman et al., 2014; Kabisch et al., 2014; Alderkamp, van Rijssel, et al., 2007). However, as indicated for *Formosa* B, fast growth on beta-glucans such as laminarin requires a balanced diet that also includes nitrogen sources like peptides. The induction of several cell wall-associated peptidases and peptide-specific transporters in *Formosa* B upon growth on laminarin suggests that these bacteria pursue a complex uptake strategy, which encompasses both sugars and nitrogen compounds, and may make marine *Flavobacteriia* so successful in this diffusion-open environment.

2.2.6 Acknowledgements

The authors are grateful to Jana Matulla for technical assistance and Sebastian Grund for mass spectrometry analyses. The Federal Ministry of Education and Research (BMBF) funded parts of this work within the "Microbial Interactions in Marine Systems" project (MIMAS project 03F0480A). The work was also financially supported by the DFG in the framework of the research unit FOR2406 "Proteogenomics of Marine Polysaccharide Utilization" (POMPU) by grants of R. Amann (AM 73/9-1), H. Teeling (TE 813/2-1), B. Fuchs (FU 627/2-1), D. Becher (BE 3869/4-1), J.-H. Hehemann (HE 7217/2-1) and T. Schweder (SCHW 595/10-1). F. Unfried was supported by scholarships from the Institute of Marine Biotechnology e.V. and the Ph.D. graduate program of the International Max Planck Research School of Marine Microbiology (MarMic).

2.2.7 Supplementary information

2.2.7.1 Supplementary methods

2.2.7.1.1 Bioinformatic analyses

Sequence alignments between genome and metagenome sequences were performed using Bl2seq (BLASTn, E-value 1e-5) (Camacho et al., 2009). Sequence data and the BLAST comparison files were drawn with the R package genoPlotR (Guy et al., 2010) version 0.8.4 and edited in Inkscape version 0.91. BLAST results were automatically edited, so that short hits contained in longer hits and hits with a bitscore below 100 were removed.

For the estimation of the in situ abundance of the *Formosa* strains A and B during algal spring blooms quality filtered reads from 44 metagenomes (Supplementary TAB 2.2.1) were mapped on the respective genomes. Reads were quality filtered and adapters trimmed using bbduk.sh (v. bbmap-35.14) from the BBtools package (Bushnell, 2016) with the following settings: ktrim=r, k=28, mink=12, hdist=1, tbo=t, tpe=t, qtrim=rl, trimq=20, minlength=100. Reads were then mapped on the genomes using bbmap.sh (BBtools package) with the following settings, using a nucleotide identity threshold of ≥95%: minid=0.95, idfilter=0.95. Sequence alignment map files (SAM) were further processed to binary format (BAM) using samtools (v. 1.2) (Li et al., 2009). Unmapped (-F 4) and low quality mapped reads (-q <10), and reads that were detected as PCR duplicates (Validation_Stringency=Lenient) using picard tools (v. 1.119) (http://broadinstitute.github.io/picard/) were removed. The number of mapped reads and coverage information were subsequently calculated using pileup.sh (BBtools package). For detection of closely related Formosa species in 4 representative metagenomes (2009/04/07, 2010/04/30, 2011/05/16, 2012/05/03), reads were mapped to the *Formosa* strain B genome using bbmap.sh with nucleotide identity threshold of ≥70% (settings: minid=0.76, idfilter=0.70). Mapped reads were directly extracted from SAM files and are shown as the fraction of reads mapped to the *Formosa* B genome compared to the total number reads from the metagenome [%].

The amino acid sequence of the catalytic domains from putative glycoside hydrolases of family GH16 from *Formosa* sp. Hel1_33_131 were blasted (BlastP) against all characterized GH16 enzymes from the CAZy database. Hits with an identity of ≥25% and a query coverage of ≥80% were selected for a MUSCLE alignment (Edgar, 2004). Afterwards the alignment was used to construct a maximum likelihood tree using bootstrap values. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 6 (Tamura et al., 2013).

2.2.7.1.2 Determination of laminarin uptake

For the determination of laminarin uptake by *Formosa* B, cells were grown in HaHa medium (Hahnke et al., 2015) with laminarin (2 g L^{-1}) and harvested during exponential growth. The cells were then inoculated (1:10) in minimal medium and 35 µmol fluorescently labelled (FLA) laminarin was added (Arnosti, 2003). The cells were sampled for visualization before the inoculation and after 5, 10 and 30 min. Cells were fixed using formaldehyde (1%) for 1 h at room temperature and subsequently filtered onto polycarbonate (0.2 µm pore size) filters. The cells were counterstained using DAPI and visualized using super-resolution structured illumination microscopy (SR-SIM) as described in detail by (Reintjes et al., 2017).

2.2.7.1.3 Subproteome fractionation

To extract the soluble intracellular proteome and the enriched membrane-associated proteome, cell pellets of *Formosa* strain B were resuspended in TE buffer (10 mM Tris-HCl pH 7.5, 10 mM EDTA pH 8.0) with Roche 'cOmplete Protease Inhibitor'. Cells were disrupted by sonication (3 x 30 s) and cell debris was removed by centrifugation (8.000 × g, 4°C, 10 min). Soluble intracellular proteins and membrane-associated proteins were separated via ultracentrifugation (100.000 × g, 4°C, 60 min).3 Purification of the membrane-bound proteome was performed as described by (Kabisch et al., 2014). To enrich the soluble extracellular proteome, the culture supernatants were precipitated with 10% TCA (trichloroacetic acid) overnight at 4°C (Antelmann et al., 2001), and precipitated proteins were extracted by centrifugation (9.500 × g, 4°C, 60 min) and washed with ice-cold ethanol (99.9%) before dissolving them in 8 M urea/2 M thiourea.

From each subproteome fraction and triplicate, 25 µg of protein were loaded on a 10% 1D-SDS polyacrylamide gel and separated according to their molecular weight for 75 min at 150 V. After fixing with ethyl acetate and Coomassie G-250 staining (Candiano et al., 2004), the proteins were in-gel digested for 16 h using trypsin as described by (Heinz et al., 2012). Peptides were eluted in an ultrasonic bath for 15 min and subsequently desalted using ZipTip columns (Millipore, Billerica, MA, USA) according to the manufacturer's guidelines.

2.2.7.1.4 Cloning procedures

The gene encoding the putative laminarinase FbGH17A (locus tag FORMB_24720) was amplified from genomic DNA of *Formosa* sp. Hel1_33_131 (GenBank accession number GenBank:

CP017260.1)(Hahnke et al., 2015). The following primers were designed to amplify the coding regions without the signal peptide-encoding sequences using Gibson assembly (Gibson et al., 2009): FbGH17Aforward: CAGCGGCCTGGTGCCGCGCGGCAGCCATAAAAATAAGCAATCAAATACG, FbGH17Areverse: CTCAGTGGTGGTGGTGGTGGTGCTCGAGTTAATGGTTCTTTGTGATTTC. They include ~30 bp overhangs complementary with the pET28a+ vector, for Gibson assembly (Gibson et al., 2009), which results in an N-terminal hexahistidine tag in the recombinant protein. The PCR followed standard PCR amplification protocols and the Gibson cloning kit (NEB) was used according to the manufacturer's instructions. Colonies of positive clones were identified by colony PCR using T7. The vector from a positive clone was purified from 2-ml Luria-Bertani (LB) broth cultures using the QIAprep Spin Miniprep kit (Qiagen). For protein expression, the vector was used to transform Escherichia coli BL21(DE3) (NEB). The FaGH17B gene (locus tag FORMB_24740) was synthesized and cloned into a pET28a+ vector by the company GenScript. Unlike the two other proteins, the vector was transformed in E. coli Rosetta (DE3) (Merck-Millipore). The cells were stored at -80°C until further use.

2.2.7.1.5 Enzyme overexpression, refolding and purification

Protein expression was conducted in ZYP autoinduction medium (Studier, 2005), which was inoculated with an overnight preculture at 37°C in LB medium supplemented with 100 µg mL-1 kanamycin. Cultures were grown for 3 days at 20°C in ZYP medium with 100 µg mL^{-1} kanamycin, with rotation at 150 rpm. Cells were harvested by centrifugation at 4,500 × g for 25 min at 4°C, and the pellet was stored at -20°C until further use. After thawing, the cells were suspended in 15 mL sucrose solution (20 mM Tris-HCl, 25% [w/v] sucrose). 30 mg of lysozyme were added and the cells were incubated for 10 min at room temperature and constant stirring. Afterwards 30 mL of deoxycholate solution (20 mM Tris-HCl pH 7.5, 1% [w/v] deoxycholate, 100 mM NaCl, 1% Triton X-100) and 0.2 mL of 1 M $MgSO_4$ were added. The highly viscous solution was liquefied by adding 10 mg of DNase (Sigma) at room temperature. The resulting lysate was centrifuged at 16,000 × g for 45 min at 4°C.

The supernatant of FbGH17A and FbGH30 was loaded onto a 5 mL HiTrap IMAC HP column (GE Healthcare), which was charged with one column volume (CV) of 500 mM $NiSO_4$ and equilibrated with 5 CVs of IMAC buffer A (20 mM Tris-HCl pH 8, 500 mM NaCl, 20 mM imidazole). After sample injection, the column was washed with IMAC buffer A (15 CVs), and the protein was eluted with 5 CVs of a linear gradient from IMAC buffer A to IMAC buffer B (20 mM Tris-HCl pH 8, 500 mM NaCl, 500 mM imidazole), at a flow rate of 5 mL min^{-1}, at 20°C. The 1 mL fractions were analyzed by SDS-PAGE. Fractions corresponding to a band at the expected size were pooled and concentrated in an ultrafiltration stirred cell (Amicon, Merck-Millipore), on a polyether-sulfone membrane with a 10-kDa cutoff value (Merck- Millipore). An aliquot of 1 to 2 mL was loaded onto a 120 mL HiPrep 16/60 Sephacryl S-200 HR column that had been equilibrated previously with

3 CVs of SEC buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl) at 20°C. The protein was eluted with 1 CV of SEC buffer, and the fractions were analyzed by SDS-PAGE. Fractions with a single band at the expected size and at an elution time corresponding to monomeric or dimeric protein were combined and concentrated. The concentration and hydrodynamic radius were determined using a BioSpectrometer (Eppendorf) and dynamic light scattering (DLS).

Since FbGH17B was expressed insoluble, it had to be refolded in vitro (adapted from (Qi et al., 2015)). After bacterial cell lysis, the resulting pellet contained the denatured protein. The pellet was resuspended in 30 mL washing buffer (20 mM Tris-HCl pH 8, 300 mM NaCl, 1 mM EDTA, 1% [v/v] Triton X-100, 1 M Urea), followed by centrifugation at 16,000 × g for 15 min and 4°C. This washing step was repeated three times. The pellet was resuspended in 10 ml refolding buffer (20 mM Tris-HCl pH 8, 1 M Urea) and stored overnight at -20°C. After thawing, the solution was centrifuged as described before. To remove the urea, the solution was dialyzed overnight against 1 L SEC buffer and using a Spectra/Por 1 dialysis membrane with a 6-8 kDa cutoff (Spectrum Labs). The reservoir was kept at 4°C and stirred permanently. To remove the precipitate resulting from the dialysis, the protein solution was centrifuged again. Afterwards, it was ready to be purified via IMAC in the same way as the other proteins. The 1 mL fractions were analyzed by SDS-PAGE but in contrast to the other proteins, they were pure enough to omit the SEC purification. Fractions corresponding to a band at the expected size were pooled and concentrated. In order to remove the imidazole, the protein was additionally applied on a 5 mL HiTtrap desalting column (GE Healthcare). The procedure was executed according to the manufacturer's instructions and the column was equilibrated with SEC buffer.

2.2.7.1.6 Hydrolysis of laminarin

Laminarin from Laminaria digitata (0.1% [w/v]; Sigma) was hydrolyzed over the course of 60 min at 37°C with 100 nM purified enzyme (~5 µg mL^{-1} of FbGH30, FbGH17A, or FbGH17B) in 50 mM MOPS buffer at pH 7. Aliquots of 100 µl were taken at 0 s, 5 min, 10 min, 20 min, 40 min and 60 min. Each reaction was stopped by boiling the sample for 5 min at 100°C.

2.2.7.1.7 Debranching and purification of laminarin

Laminarin was debranched with FbGH30. 100 mg laminarin was hydrolyzed overnight under the conditions mentioned above. The reaction was stopped by boiling the sample for 5 min at 100°C. Precipitated protein was removed by filtration through 0.2 µm Costar Spin-X Filters (Corning). Afterwards, the debranched high molecular laminarin was separated from glucose by size exclusion chromatography using a HiTrap Desalting column (GE Healthcare) according to the manufacturer's instructions. The column was equilibrated and the sample was eluted with Milli-Q water. The water was evaporated overnight in a vacuum concentrator at 45°C and constant rotation (Eppendorf).

2.2.7.1.8 Michaelis-Menten-Kinetics

The kinetic parameters of the investigated enzymes acting on native and debranched laminarin from *L. digitata* (Sigma) were determined using 100 nmol of the enzyme in a reaction mixture of 600 µL over 10 min at 37°C and in MOPS buffer at pH 7. Seven different laminarin concentrations were measured in triplicates: 0% (w/v), 0.05%, 0.1%, 0.2%, 0.4%, 0.8%, 1.6%. 100 µL aliquots were taken every 2 min and the reaction was stopped by boiling the sample for 5 min at 100°C. The amount of released reducing ends was measured by the PAHBAH reducing sugar assay (Moretti and Thorson, 2008). One mL of a freshly prepared 9:1 mixture of reagent A (0.3 M 4-hydroxybenzhydrazide, 0.6 M HCl) and reagent B (48 mM trisodium citrate, 10 mM $CaCl_2$, 0.5 M NaOH) was added to each aliquot and the mixture was heated for 5 min at 100°C. Absorbance was determined at 410 nm using a BioSpectrometer (Eppendorf).

2.2.7.1.9 High Performance Anion Exchange Chromatography with Pulsed Amperometric Detection (HPAEC PAD)

An HPAEC-PAD system was applied for qualitative product analysis. An ICS-5000+ (Dionex) with electrochemical detection on the gold working electrode and a pH reference electrode (Ag/AgCl) was used. The detector wave form was: E1=100 mV (T=0.4 s), E2=-2000 mV (T=0.42 s), E3=600 mV (T=0.43 s) and E4=-100 mV (T=0.5 s). A 25 µL sample loop was used on a Dionex CarboPac PA100 analytical column (2×250 mm) coupled with a Dionex CarboPac PA100 guard column (2×50mm). Chromatography and detection were conducted at 25°C. A mixed sugar standard solution consisting of 1µg mL^{-1} Glucose (Sigma), laminaribiose, laminaritriose, laminaritetraose, laminaripentaose and laminarihexaose (all from Megazyme) was used as reference. Eluent 1 consisted 1 of 0.15 M NaOH (HPLC grade, VWR) and eluent 2 of 0.15 M NaOH and 1 M sodium acetate (HPLC grade, Sigma). All eluents were dissolved in Milli-Q water, degassed with helium for at least 10 min and kept pressurized after connecting the bottles to the system. Eluent 2 was filtered through a 0.2 µm nylon filter membrane. Separation was achieved by a linear gradient course of two mobile eluents (from 100% eluent 1 to 50% eluent 1 and 50% eluent 2) over a duration of 19.5 min, followed by an increase of the eluent 2 concentration to 70% within 6 s and a linear gradient to 100% eluent 2 over 114 s immediately afterwards. Finally, the concentration of eluent 1 was increased back to 100% over 30 s. The entire sequence was conducted at a flow rate of 0.25 mL min^{-1}. The column was equilibrated with eluent 1 for 3 min between each sample.

2.2.7.1.10 Phylogenetic analysis

All available Formosa-related 16S rRNA sequences from SILVA SSURef v.128 database (Quast et al., 2012) were loaded into ARB v.6.1 (Ludwig et al., 2004) and consistently aligned using SILVA Incremental Aligner (SINA) (Pruesse et al., 2012). Phylogenetic tree reconstructions were done with the ARB internal programs (i) for maximum likelihood method RAxML v.8 (Stamatakis, 2014) using the 'GTRGAMMA' substitution model and 'thorough tree search' option activated,

(ii) the ARB neighbor joining program using the Jukes-Cantor correction and (iii) the 'ARB_PARSIMONY' maximum parsimony method for global and local optimizations. All treeing algorithms were run on sequences filtered with and without a 30% and 50% positional conservation filter for all *Flavobacteriia*. A consensus tree was built from these trees following the recommendation by (Peplies et al., 2008). Therefore, bootstrap values cannot be given, and branches indicating ≤1% distance should be regarded as uncertain.

2.2.7.2 Supplementary results

2.2.7.2.1 Metaproteomic identification of Formosa-specific enzymes and transporters during microalgal blooms

For the metaproteome analysis, we generated combined in silico databases of all potential peptides using the recorded metagenome data from the spring blooms in 2009 (Teeling et al., 2012) and 2010 (Teeling et al., 2016), respectively, as well as the predicted protein sequences of the *Formosa* A (Hel3-A1_48) and B (Hel1_33_131) genomes. The combined search of the measured peptide spectra from the spring bloom samples taken on April 7th 2009 against this database yielded several marker proteins of *Formosa* B's PUL 1, including the glycoside hydrolase GH3 (FORMB_10040), a PKD domain protein (FORMB_10070), the TBDR protein (FORMB_10080), and the SusD-like protein (FORMB_10090) (FIG 2.2.2B and Supplementary TAB 2.2.5A). In addition, we detected the PUL 2-encoded Formosa B proteins glycoside hydrolase GH2 (FORMB_24670), two putative GH16 (FORMB_24680, FORMB_24690), a GH30 (FORMB_24730), a PKD domain protein (FORMB_24700), a hypothetical protein (FORMB_24710), and the MFS transporter protein (FORMB_24740). Furthermore, also the TBDR protein (FORMB_13580) of *Formosa* B's PUL 3 could be identified in the environmental samples (FIG 2.2.2B and Supplementary TAB 2.2.5A). Altogether, this analysis provided evidence that a significant proportion of *Formosa* strain B's putative laminarin PUL-encoded proteins were expressed in situ during the spring bloom in 2009. Furthermore, the metaproteome analysis also revealed the three *Formosa* B marker proteins of PUL 1, FORMB_10060, FORMB_10080, and FORMB_10090, in the environmental samples of 2010 (Supplementary TAB 2.2.5B).

Additional *Formosa* B proteins that were detected in the environmental metaproteome samples of 2009 are involved in the central catabolism of the laminarin-specific monosaccharide glucose (FIG 2.2.2B and Supplementary TAB 2.2.5A). Nearly all glycolysis-related enzymes were detected. This includes glucose-6-phosphate isomerase (EC 5.3.1.9, FORMB_15410), fructose-bisphosphate aldolase (EC 4.1.2.13, FORMB_06360), triosephosphate isomerase (EC 5.3.1.1, FORMB_18610), glyceraldehyde-3-phosphate dehydrogenase (GAPDH, EC 1.2.1.12, FORMB_25050), glyceraldehyde-3-phosphate dehydrogenase (NAD-dependent GAPDH, EC 1.2.1.12, FORMB_22380), 2,3-bisphosphoglycerate independent phosphoglycerate mutase (EC 5.4.2.11, FORMB_06010), and pyruvate kinase (EC 2.7.1.40, FORMB_11840). In addition, the pyruvate dehydrogenase E1 component beta subunit (EC 1.2.4.1, FORMB_09930), the anaplerotic

enzyme phosphoenolpyruvate (PEP) carboxylase (EC 4.1.1.31, FORMB_11960) and a putative glycogen synthase (EC 2.4.1.21, FORMB_13890) of *Formosa* B could be identified in the metaproteome samples of the spring bloom 2009 (FIG 2.2.2B and Supplementary TAB 2.2.5A). These data demonstrate that the *Formosa* B strain substantially contributed to laminarin degradation and turnover during a diatom-driven phytoplankton bloom.

Furthermore, we also identified glycolytic marker proteins of *Formosa* strain A in the metaproteome samples (Supplementary TAB 2.2.5A-B). These include, for example, the fructose-bisphosphate aldolase class II (EC 4.1.2.13, FORMA_07040), and the pyruvate dehydrogenase E1 component beta subunit (EC 1.2.4.1, FORMA_09500), during the spring bloom 2009, and the glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12, FORMA_16240) in 2010 environmental samples.

2.2.7.2.2 Laminarin-induced proteins of *Formosa* B

With laminarin as growth substrate, the TonB-dependent-receptor protein FORMB_10090 belonged to the most strongly induced proteins within the laminarin PULs (FIG 2.2.3B). Moreover, this TBDR protein was one of the most abundant proteins of the entire membrane-enriched proteome in this study (0.8% riBAQ; Supplementary TAB 2.2.6A) and was also particularly abundant in the bacterioplankton metaproteome during the spring bloom in April 2009 (FIG 2.2.2B and Supplementary TAB 2.2.5A). Furthermore, the TBDR protein FORMB_10090 was likewise identified together with the PUL 1 - specific proteins FORMB_10080 (SusD-like protein) and FORMB_10060 (PKD domain protein) in the metaproteome of the spring bloom in May 2010 (FIG 2.2.2B and Supplementary TAB 2.2.5B). This highlights the importance of this polysaccharide uptake protein in glycan utilization. In the control experiments with glucose or peptone, this receptor protein (FORMB_10090) was also detectable, albeit at significantly lower levels. Such a weak constitutive expression pattern was also observed for other PUL-encoded proteins and might be explained by a sensory function of these proteins (Kabisch et al., 2014; Thomas et al., 2013; Hehemann, Kelly, et al., 2012). The constitutive, low basal expression of selected proteins, which are involved in initial polysaccharide degradation and uptake steps, would allow bacteria to scan their environment for potential substrates without having to maintain the entire machinery for polysaccharide turnover.

A phylogenetic analysis of 9 putative laminarinases of *Formosa* B indicated a clear affiliation to the GH16 family for three of the enzymes (Supplementary FIG 2.2.7). Two of the respective genes are located in the laminarin-inducible PULs 1 and 2 (FIG 2.2.3C). The laminarin-induced protein machinery of *Formosa* B includes additional proteins, which are not organized in PULs, such as the putative laminarinase (FORMB_24690) of the GH16 family (Supplementary FIG 2.2.7A and Supplementary TAB 2.2.7). This enzyme belonged to the most highly expressed proteins in the proteome during growth with laminarin, but was not detectable with glucose or in the control culture (Supplementary FIG 2.2.7B). We furthermore detected a GH5 protein (FORMB_08760),

which is encoded in an operon with three other proteins, all of which were also upregulated in the presence of laminarin (Supplementary TAB 2.2.7). A putative polysaccharide deacetylase (FORMB_02440), which possesses a NodB-like domain of the carbohydrate esterase family 4 (CE4), was also found to be induced by laminarin (Supplementary TAB 2.2.7). The second domain of this multi-modular enzyme seems to be a glycoside hydrolase of the GHL13 family. Elevated protein abundance under laminarin conditions was also observed for an ABC-transporter protein (FORMB_12400), which is encoded in an operon with a putative aspartate-semialdehyde dehydrogenase. It is interesting to note that the glycolytic enzymes glyceraldehyde-3-phosphate dehydrogenase (FORMB_25050) and the pyruvate kinase (FORMB_11840) were also found to be induced in *Formosa* B cultures with laminarin, which indicates an upregulation of this part of the glycolytic pathway under laminarin conditions. Furthermore, the increased level of the phosphoenolpyruvate carboxylase (FORMB_11960) during the growth on laminarin could ensure the replenishment of oxaloacetate in the citric acid cycle to support the metabolic flux through this pathway and to thus boost anabolic and energy generation processes.

Beside the PUL-encoded TBDRs we identified FORMB_05090 as an additional laminarin-induced TBDR protein (Supplementary FIG 2.2.9 and Supplementary TAB 2.2.7). We detected further TBDR proteins, the expression of which was upregulated by both laminarin and glucose (Supplementary FIG 2.2.9). This includes FORMB_02790 und FORMB_02850, which are located in an operon with two hypothetical proteins in a proposed mannan PUL (Supplementary FIG 2.2.2). The glucose and laminarin-induced TBDR protein FORMB_07380 is located in an operon with a SusD-like protein of unknown function. The gene encoding the upregulated TBDR protein FORMB_15500 is located in an operon with a gene encoding a hypothetical membrane protein, which contains a putative galactose-binding domain. These TBDRs were also detected in the metaproteome samples of the spring bloom in 2009. It is interesting to note that the highly abundant TBDR protein FORMB_15420, which could also be identified in the metaproteome samples (FIG 2.2.2B), clusters genomically together with the gene of the glucose-6-phosphate isomerase (FORMB_15410) and a hypothetical protein of unknown function.

2.2.7.2.3 Subproteome analyses

The bioinformatic prediction of the putative cellular localization of laminarin-induced proteins of *Formosa* B as suggested in FIG 2.2.6 was based on the strategy recently proposed by Romine (Romine, 2011)(Supplementary TAB 2.2.7). The computational analysis of specific domains and putative protein sorting signals was combined with the subproteome analyses of the enriched soluble intracellular, membrane-associated and extracellular protein fractions to improve confidence in our protein location predictions. A PorSS signal sequence of the novel type IX protein secretion system (Sato et al., 2010)(Sato et al., 2010) was predicted for the laminarinase GH16 (FORMB_11070) as well as for a number of hypothetical proteins, suggesting an extracellular localization of these enzymes (Supplementary TAB 2.2.7). These analyses also

suggested a surface-tethered localization of the laminarinase GH16 FORMB_10050, the hypothetical protein FORMB_24710, the GHs FORMB_13600 and FORMB_24680 or the PKD-domain proteins of the laminarin specific PULs. The multi-modular MFS-GH17 fusion protein FORMB_24740 with 12 membrane-spanning domains was enriched in the membrane proteome fraction.

It needs to be kept in mind, though, that our subproteome enrichment procedure does not provide a clean, exclusive separation of intracellular, membrane-associated and secreted proteins from each other, but can only give indications on the actual protein localization. The proposed model of the laminarin and protein utilization machineries in *Formosa* B illustrated in FIG 2.2.6 is, therefore, a working model and requires further in-depth analyses in the future in order to prove specific protein interactions and their subcellular organization.

2.2.7.2.4 Structural analysis of GH17A

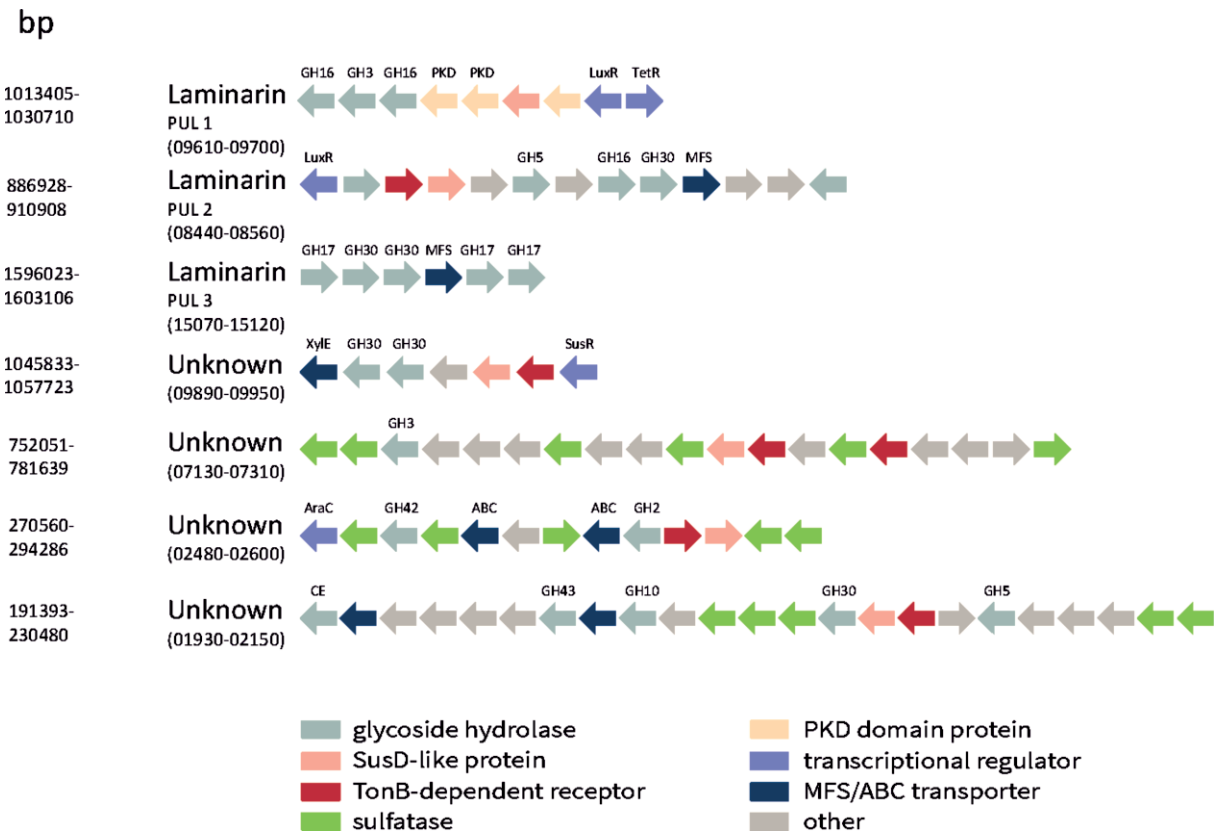
In order to examine the molecular basis of substrate specificity the X-ray crystal structure of GH17A was solved to a resolution of 2.6 Å by molecular replacement using a GH17 from the fungus *Rhizomucor miehei* (Qin et al., 2015)(Supplementary FIG 2.2.10). The structure was modelled without gaps from residues 38 to 430, the C-terminus (Supplementary TAB 2.2.8). Within the crystal lattice, the protein is found as a dimer of trimers where a C-terminal loop blocks the active site of the adjacent monomer (Supplementary TAB 2.2.8). The structure consists of a modified $(\alpha/\beta)_8$ fold whereas other characterized members of GH17 are smaller and unmodified (Varghese et al., 1994). Compared to the monomeric GH17 structures, FbGH17A has significant insertions and is larger (FIG 2.2.4F). The closest structurally characterized GH17 homolog from *R. miehei* (Qin et al., 2015) shares an overall Cα RMSD of 1.85 Å and a structurally aligned sequence identity of 23%. The insertions of FbGH17A are found in two main places. First, the loop β3-α3 is longer and is found adjacent to the active site. Second, the helix α6 is extended and the C-terminal loop wraps around this extension to reach into the adjacent active site (FIG 2.2.4F). The latter insertions appear to be involved in oligomerization as opposed to catalysis as they are far away from the active site.

The structure of GH17A allows for the deduction of the molecular basis of substrate specificity for the laminarinase. Based on the GH17 complexes obtained for *R. miehei* (Qin et al., 2015), a model was generated of a laminarin product involving 5 monomers bound to the catalytic groove, two on the aglycon side and three on the glycon side (FIG 2.2.4G). The reducing and non-reducing ends of the modelled glycan are free, suggesting the protein can act in the middle of the chain as expected for an endo-acting glycoside hydrolase. Furthermore, given the conformation of the modeled glycan, 6-O-β- glucose branching would be possible only at subsite +1 and anything further away (+3 or –4). In other words, within the native polysaccharide the enzyme would need a stretch of at least three free β-1,3-glucose moieties to act. This structural data

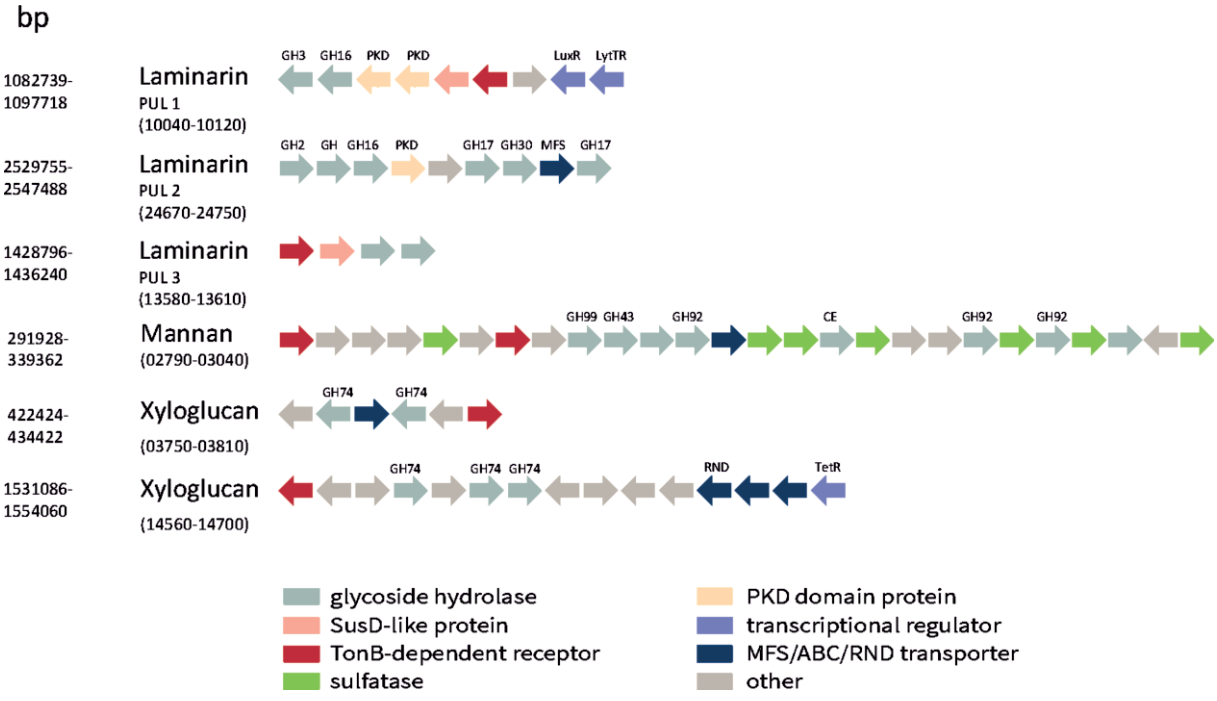
supports the observation that GH17 activity on laminarin is bolstered by the action of the debranching enzyme GH30.

2.2.7.3 Supplementary figures

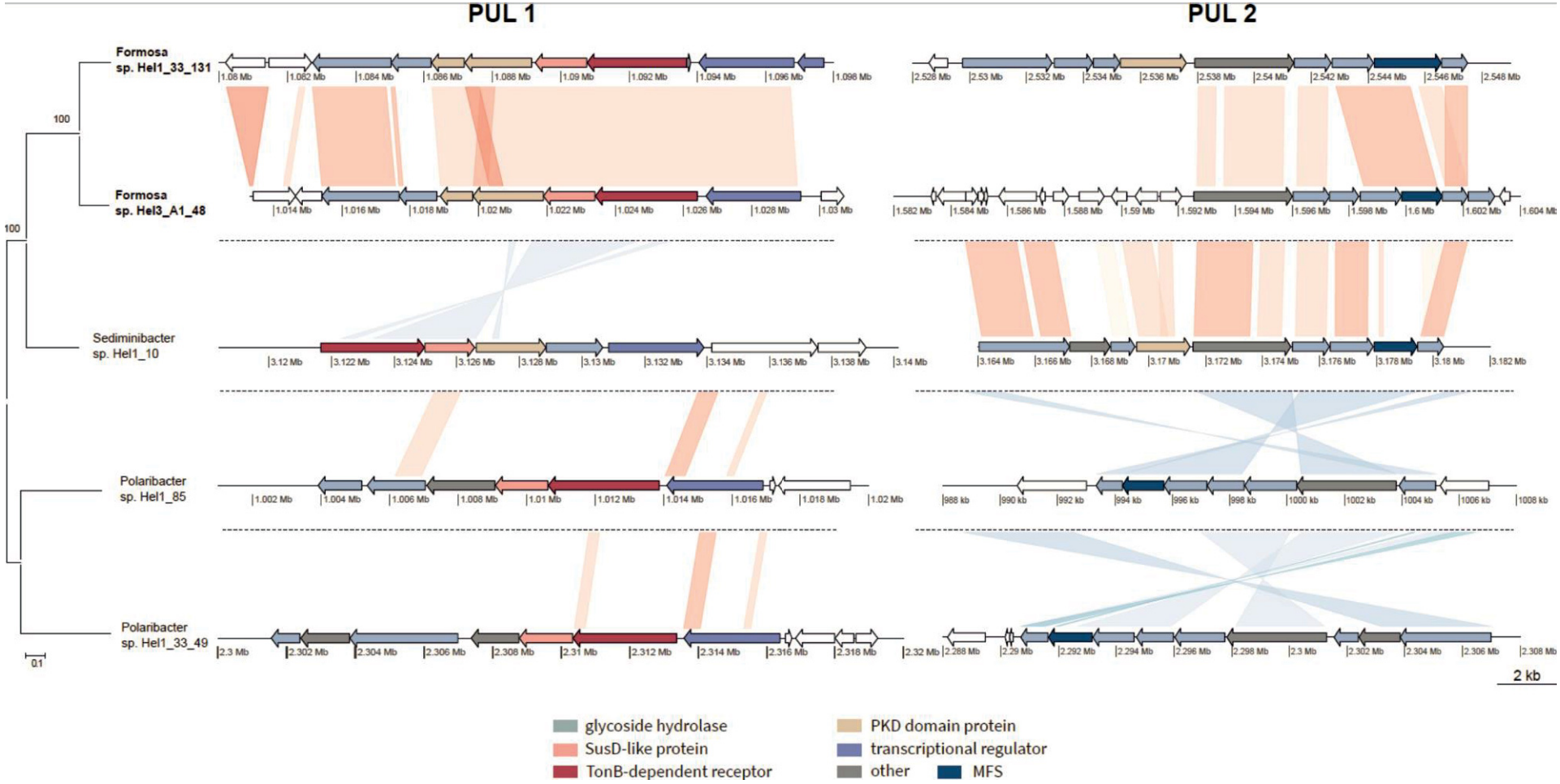
Supplementary FIG 2.2.1 Phylogenetic consensus tree of the *Formosa* clade. Accession numbers are given at the end of each twig. The two new Formosa strains are marked in red. 16S rRNA gene clone sequences retrieved from Helgoland Roads are shown in green. The numbers in brackets indicate the number of grouped sequences. The pink box indicates the specificity of the 16S rRNA-targeted probe FORM181B for Hel1_33_131 (*Formosa* strain B) and closely related clones.



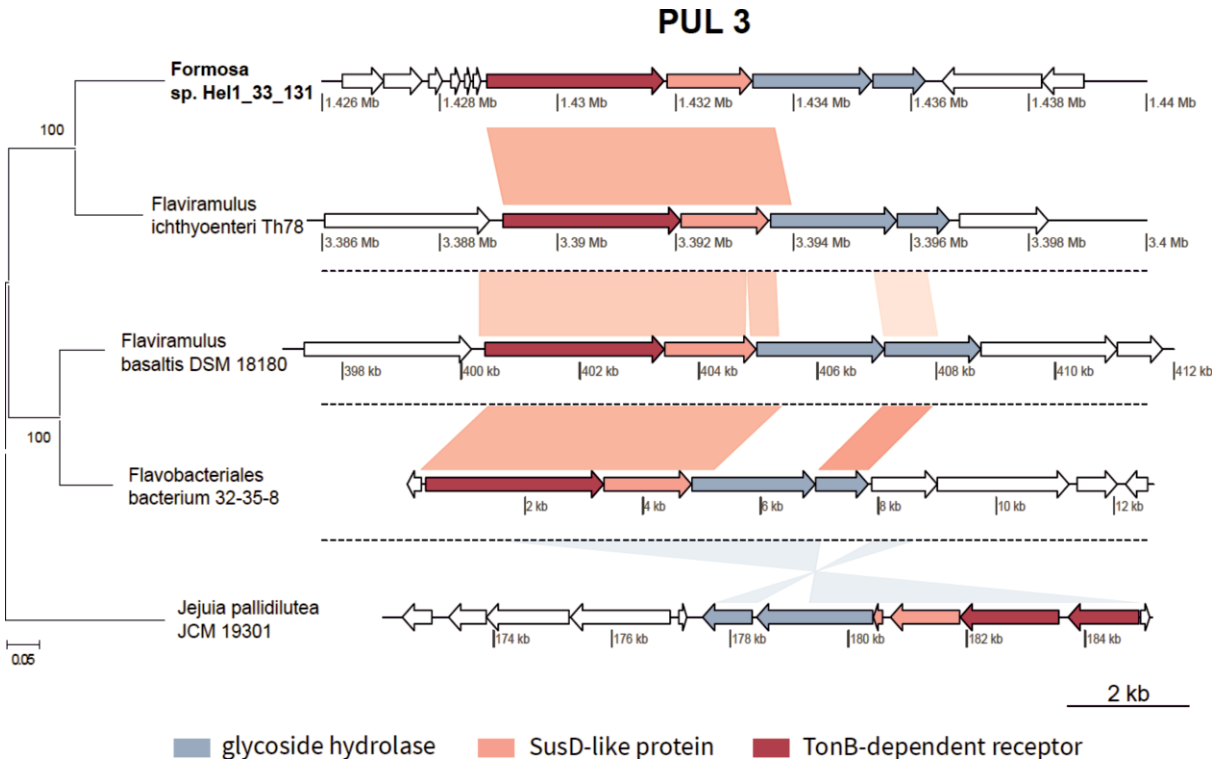
Supplementary FIG 2.2.2 PUL organization in Formosa A



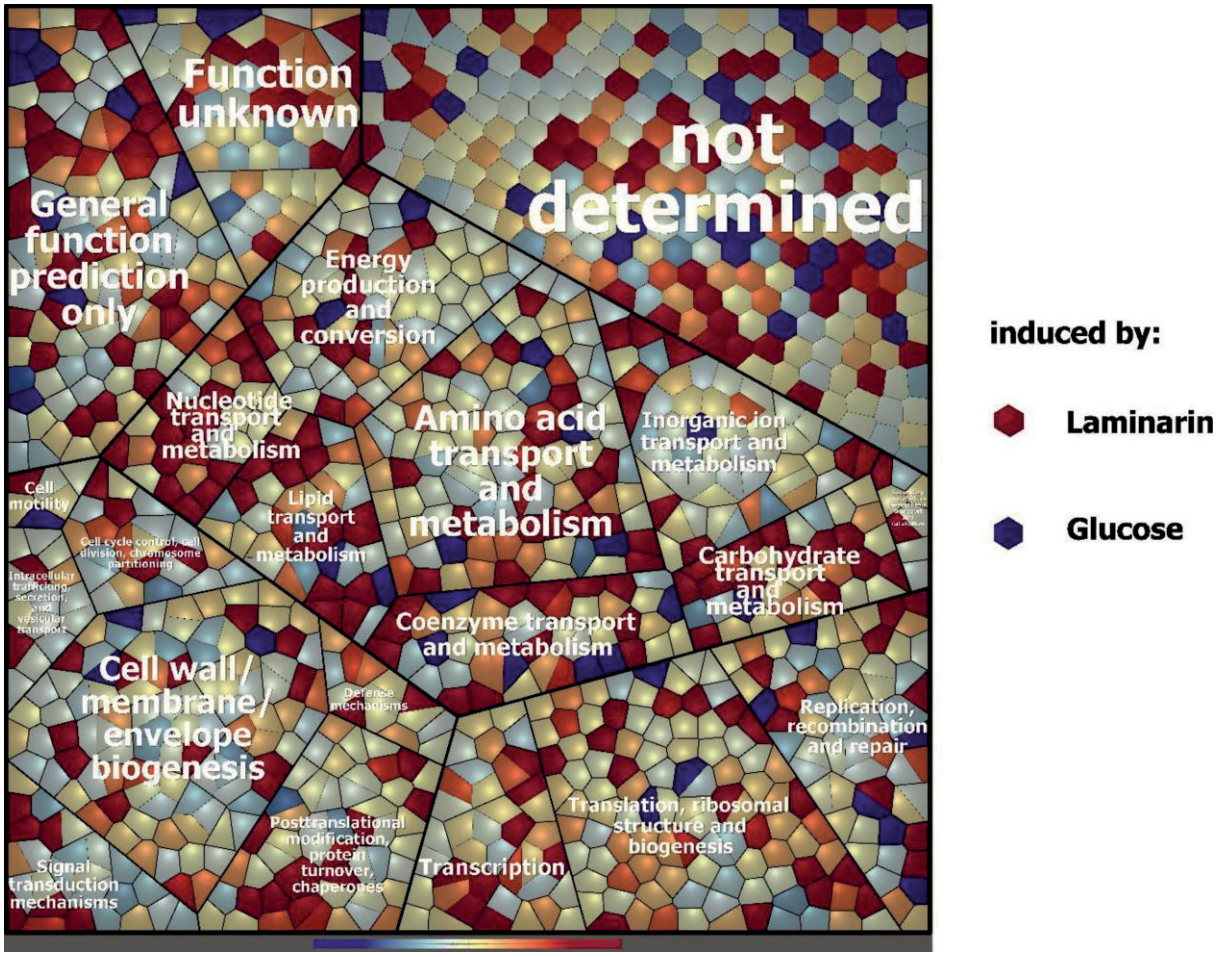
Supplementary FIG 2.2.3 PUL organization in Formosa B



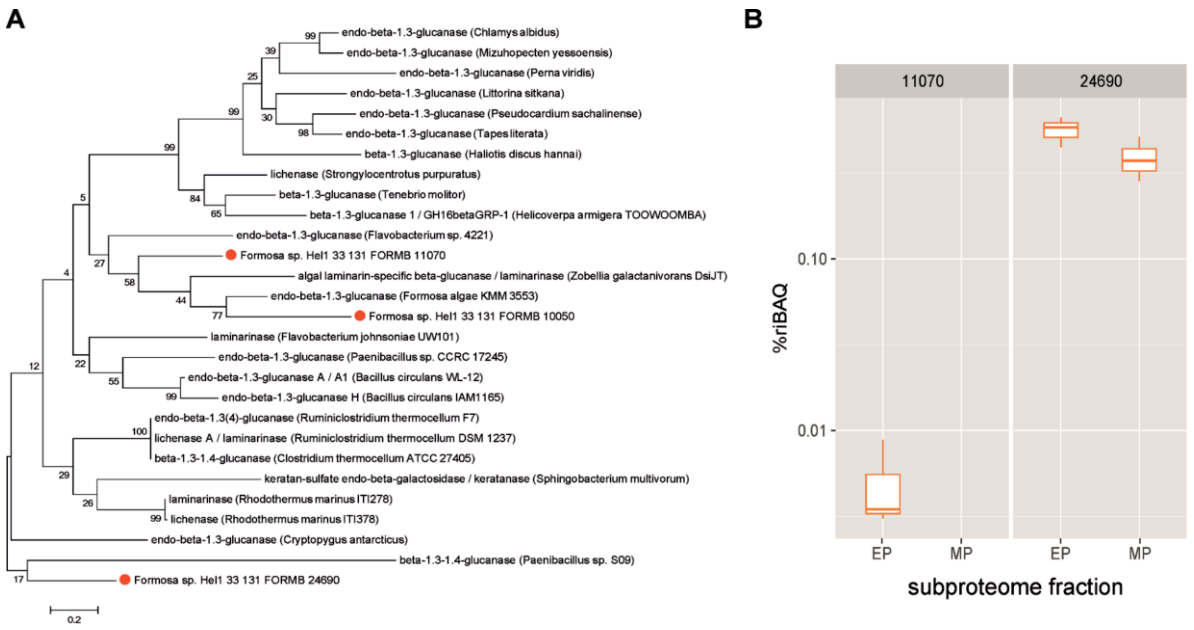
Supplementary FIG 2.2.4 The laminarin-specific PULs 1 and 2 in *Formosa* B and A. Gene organization of these PULs is conserved among other marine Flavobacteriaceae as exemplified. The PUL structures and their syntenies to the *Formosa* B PULs are arranged according to the position of their PUL-specific TBDR proteins in the phylogenetic tree depicted on the left of this figure. The phylogenetic analysis was done by the Maximum Likelihood method based on the JTT matrix-based model (Jones et al., 1992). The tree with the highest log likelihood (-8848.5785) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 5 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 954 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).Amino acid sequences of the following genes were used to generate this tree: FORMB_10090 (*Formosa* sp. Hel1_33_131), FORMA_09680 (*Formosa* sp. Hel3_A1_48), WP_007807822.1 (*Sediminibacter* sp. Hel1_10), PHEL85_0899 (*Polaribacter* sp. Hel1_85), PHEL49_2085 (*Polaribacter* sp. Hel1_33_49).



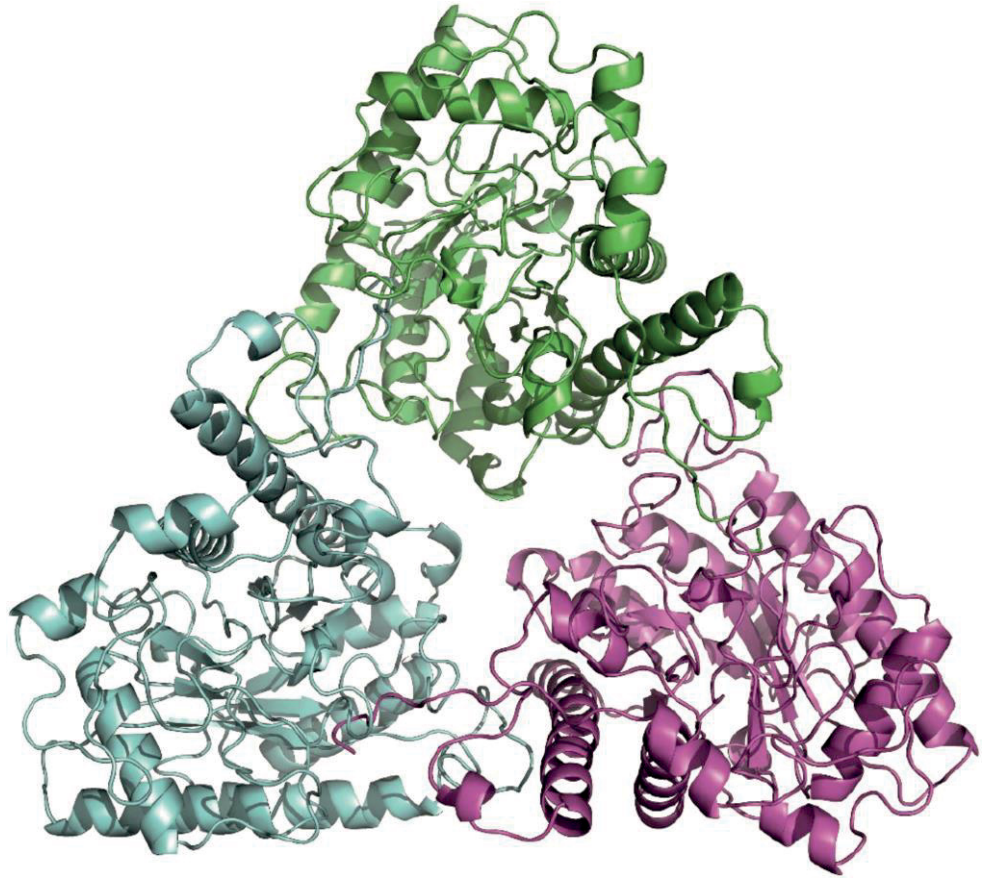
Supplementary FIG 2.2.5 *Formosa* B PUL 3. The third laminarin-specific PUL of *Formosa* B (PUL 3) shows high syntenies to other Flavobacteriaceae, which do not possess PUL 1 and 2. The PUL structures and their syntenies to the *Formosa* B PUL 3 are arranged according to the position of their PUL-specific TBDR proteins in the phylogenetic tree depicted on the left of this figure. The phylogenetic analysis was done by the Maximum Likelihood method based on the JTT matrix-based model (Jones et al., 1992). The tree with the highest log likelihood (-3670.6338) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 5 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 549 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013). Amino acid sequences of the following genes were used to generate this tree: FORMB_10090 (*Formosa* sp. Hel1_33_131), RG22_RS14790 (*Flaviramulus ichthyoenteri* Th78), SAMN05428642_102383 (*Flaviramulus basaltis* DSM 18180), OYX28435.1 (*Flavobacteriales bacterium* 32-35-8), JCM19301_3833 (*Jejuia pallidilutea* JCM 19301).



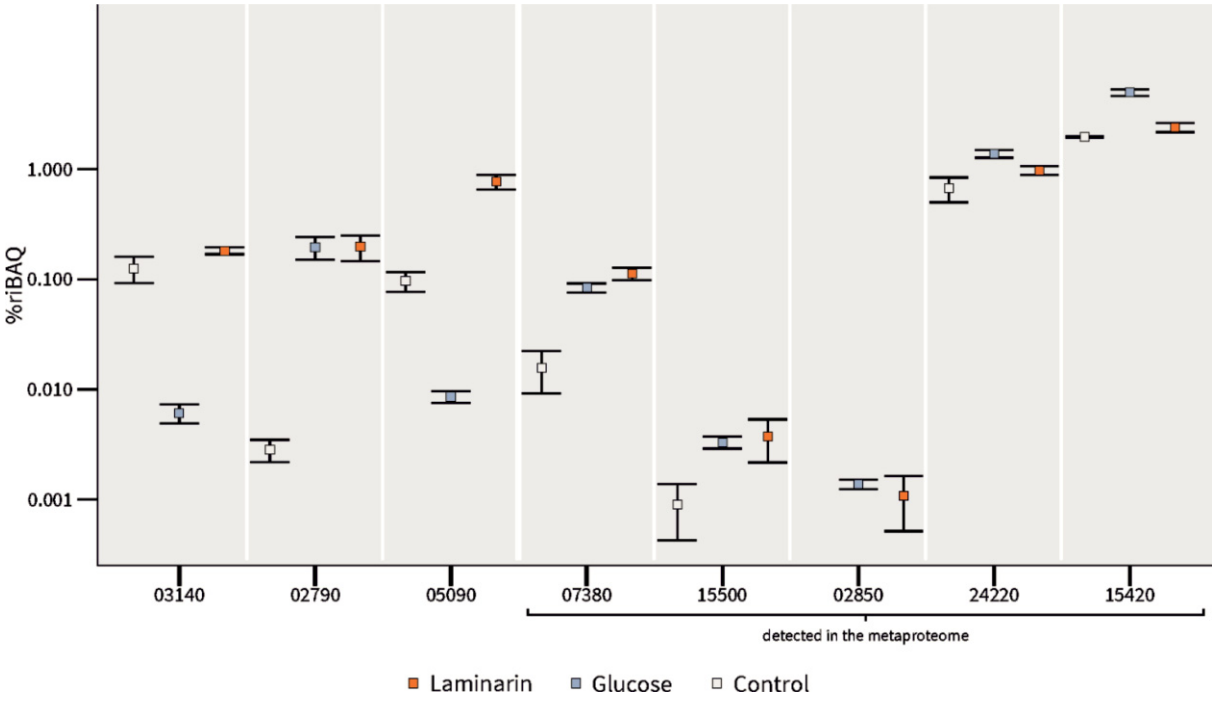
Supplementary FIG 2.2.6 Differential protein expression in *Formosa* B. Voronoi tree map comparing *Formosa* B protein expression patterns during growth on laminarin vs. growth on glucose. The color code reflects the calculated log2 ratios of protein abundances under these two substrate conditions. Mean NSAF values (normalized spectral abundance factors) from three biological replicates of the enriched membrane proteome were used for the calculation of expression ratios (fold change values). Functional protein categories were classified with Prophane (Schneider et al., 2011) and the tree map was created using Paver (www.decodon.com/paver.html).



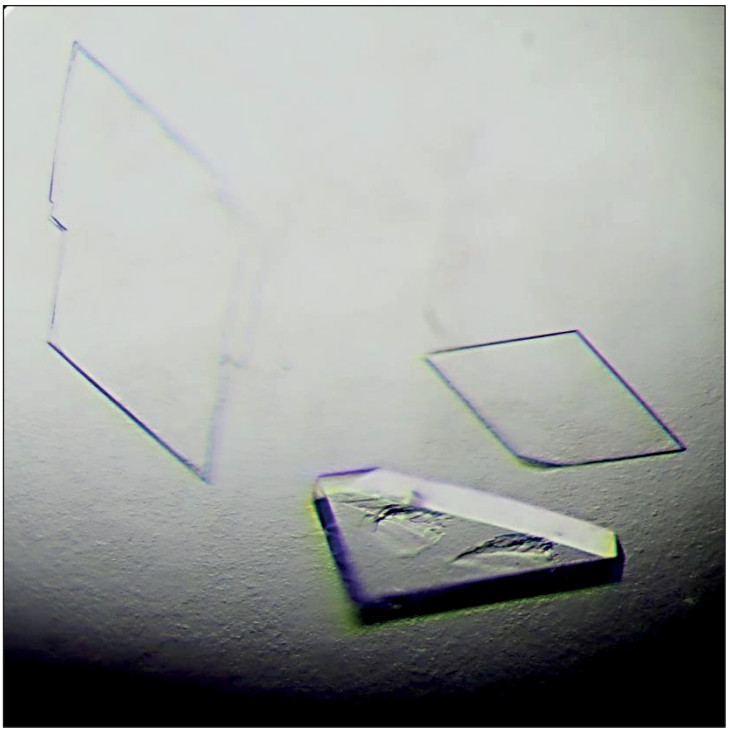
Supplementary FIG 2.2.7 GH16 enzymes in *Formosa* B. **(A)** Maximum likelihood phylogenetic tree of the three GH16 enzymes FORMB_11070, FORMB_24690 and FORMB_10050. The alignment was prepared with MUSCLE (Edgar, 2004) and the tree with MEGA6 (Tamura et al., 2013). Numbers given on the branches are bootstrap proportions as a percentage of 1000 replicates. The scale bar represents the number of amino acid substitutions per site. **(B)** Abundance of the putative *Formosa* B laminarinases GH16 FORMB_11070 and FORMB_24690 in the presence of laminarin. None of these proteins could be detected during growth on glucose. The third GH16, FORMB_10050, was not at all detected in our *Formosa* B proteome analysis under either of the conditions. EP: enriched extracellular proteome, MP: membrane-associated proteome.



Supplementary FIG 2.2.8 Trimer of FbGH17A found in the asymmetric unit. The assembly is shown with each of the three monomers in cyan, green and purple.



Supplementary FIG 2.2.9 Abundance of non-PUL-encoded TBDRs the expression of which were laminarin- and/or glucose-induced. These proteins were detected in the membrane-enriched proteome of *Formosa* B. The five proteins shown on the right were also detected in the metaproteome of a spring bloom in 2009. Relative protein abundances are depicted as % riBAQ values, i.e., as % of all proteins in the same sample. The respective locus tags of the proteins are indicated. The squares represent the mean values of the replicates for every protein and each substrate. The error bars refer to the standard error of the mean (SEM; standard deviation/√number of replicates). Proteins that could be detected in at least 2 out of 3 independent biological replicates are shown.



Supplementary FIG 2.2.10 FbGH17A crystals growing in a condition containing 0.03 M $MgCl_2$, 0.1 M MOPS (pH 7), 9% PEG8K and 15% ethylene glycol.

2.2.7.4 Supplementary tables

Supplementary TAB 2.2.1 Percentage of metagenomic reads mapped to the *Formosa* strain A and B genomes at different nucleotide identities. Quality filtered reads from 44 metagenomes from phytoplankton blooms of four consecutive years were mapped to the respective genomes. Reads were quality filtered and adapters trimmed using bbduk.sh (v. bbmap-35.14) from the BBtools package (Bushnell, 2016). Reads recruited at ≤93% nucleotide identity represent other Formosa spp. that were abundant during the bloom events. For better readability, the table is also attached on the DVD.

				Formosa A		*Formosa* B		
Date	Sequencing platform	Total number of reads	No. of quality filtered reads	Reads recruited by the genome (%)	Genome coverage (%)	Reads recruited by the genome (%)	Genome coverage (%)	Source of metagenome
20090211	454 FLX Ti	2,203,446	1,604,288	0.003	0.6	0.058	10.3	ENA: ERP001227
20090331	454 FLX Ti	1,666,535	1,140,320	0.007	0.7	0.120	14.9	ENA: ERP001227
20090407	454 FLX Ti	2,109,239	1,537,921	0.018	1.4	2.937	91.4	ENA: ERP001227
20090414	454 FLX Ti	4,588,441	3,236,386	0.007	1.2	0.516	76.6	ENA: ERP001227
20090616	454 FLX Ti	1,120,072	817,516	0.005	0.4	0.003	0.2	ENA: ERP001227
20090901	454 FLX Ti	2,714,430	1,728,646	0.280	52.6	0.002	0.2	ENA: ERP001227
20100303	HiSeq2000	285,509,546	215,734,422	0.002	2.5	0.001	1.1	JGI CSP COGITO met01
20100330	HiSeq2500	87,007,870	81,716,370	0.019	2.7	0.010	4.0	JGI CSP COGITO mtgs 100330
20100408	HiSeq2000	330,603,192	232,384,184	0.036	4.7	0.027	60.7	JGI CSP COGITO met02
20100413	HiSeq2500	85,204,698	79,754,314	0.014	2.3	0.034	63.3	JGI CSP COGITO mtgs 100413
20100420	HiSeq2500	83,499,192	78,919,256	0.025	3.2	0.285	98.1	JGI CSP COGITO mtgs 100420
20100423	HiSeq2500	79,594,912	75,009,358	0.025	3.6	0.937	99.6	JGI CSP COGITO mtgs 100423
20100430	HiSeq2500	72,546,152	68,614,694	0.026	4.2	0.996	99.5	JGI CSP COGITO mtgs 100430
20100504	HiSeq2000	274,336,674	224,993,942	0.033	6.3	0.367	99.4	JGI CSP COGITO met03
20100511	HiSeq2500	75,567,206	71,051,770	0.018	5.2	0.043	28.1	JGI CSP COGITO mtgs 100511
20100518	HiSeq2000	309,971,624	199,665,578	0.039	6.6	0.108	47.3	JGI CSP COGITO met04
20110321	HiSeq2500	84,894,030	79,587,610	0.003	2.8	0.003	0.9	JGI CSP COGITO mtgs 110321
20110324	HiSeq2000	304,742,176	244,800,088	0.003	3.1	0.003	1.2	JGI CSP COGITO met05
20110328	HiSeq2500	75,242,542	69,406,062	0.003	2.0	0.003	0.8	JGI CSP COGITO mtgs 110328
20110331	HiSeq2500	74,047,178	69,292,462	0.004	2.8	0.004	0.8	JGI CSP COGITO mtgs 110331
20110404	HiSeq2500	79,429,002	73,995,840	0.005	2.3	0.004	0.8	JGI CSP COGITO mtgs 110404
20110407	HiSeq2500	88,393,908	82,789,928	0.006	2.5	0.005	1.0	JGI CSP COGITO mtgs 110407
20110414	HiSeq2500	76,260,848	71,086,194	0.004	1.9	0.003	0.8	JGI CSP COGITO mtgs 110414
20110421	HiSeq2500	81,789,256	76,032,310	0.004	2.4	0.002	1.5	JGI CSP COGITO mtgs 110421
20110426	HiSeq2500	84,364,922	78,545,118	0.006	3.8	0.004	4.6	JGI CSP COGITO mtgs 110426
20110428	HiSeq2000	295,426,382	235,544,646	0.009	11.0	0.011	44.5	JGI CSP COGITO met06
20110506	HiSeq2500	80,085,580	74,970,646	0.019	12.5	0.114	85.9	JGI CSP COGITO mtgs 110506
20110509	HiSeq2500	90,174,260	84,434,414	0.022	10.2	0.134	90.5	JGI CSP COGITO mtgs 110509
20110512	HiSeq2500	81,683,140	75,835,970	0.017	8.9	0.127	88.5	JGI CSP COGITO mtgs 110512
20110516	HiSeq2500	88,557,186	83,034,602	0.024	8.7	0.982	97.0	JGI CSP COGITO mtgs 110516
20110519	HiSeq2500	100,457,324	94,028,844	0.026	9.2	0.673	96.7	JGI CSP COGITO mtgs 110519
20110523	HiSeq2500	227,207,716	207,300,790	0.021	11.4	0.296	96.2	JGI CSP COGITO mtgs 110523
20110526	HiSeq2000	324,219,736	250,091,456	0.015	38.1	0.039	67.5	JGI CSP COGITO met07
20110530	HiSeq2500	83,974,950	77,770,314	0.009	8.3	0.016	12.8	JGI CSP COGITO mtgs 110530
20120308	HiSeq2000	311,415,662	207,556,978	0.002	7.8	0.001	1.1	JGI CSP COGITO met08
20120405	HiSeq2500	173,626,414	161,992,212	0.003	4.6	0.002	1.2	JGI CSP COGITO mtgs 120405
20120412	HiSeq2500	175,790,324	164,053,740	0.002	2.9	0.001	1.2	JGI CSP COGITO mtgs 120412
20120416	HiSeq2000	258,751,220	210,093,962	0.003	2.6	0.002	1.5	JGI CSP COGITO met09
20120426	HiSeq2500	110,390,598	96,979,192	0.008	3.6	0.048	63.9	JGI CSP COGITO mtgs 120426
20120503	HiSeq2500	105,224,208	94,343,038	0.010	5.0	0.214	94.1	JGI CSP COGITO mtgs 120503
20120510	HiSeq2000	272,267,348	222,320,174	0.005	6.4	0.072	91.4	JGI CSP COGITO met10
20120524	HiSeq2500	113,234,474	101,681,308	0.021	47.0	0.038	47.2	JGI CSP COGITO mtgs 120524
20120531	HiSeq2500	102,492,998	91,212,794	0.011	27.9	0.009	10.9	JGI CSP COGITO mtgs 120531
20120607	HiSeq2500	189,120,228	169,486,964	0.017	57.8	0.022	13.6	JGI CSP COGITO mtgs 120607

Supplementary TAB 2.2.2 Summary of the phylogeny-guided CAZyme annotations for *Formosa* strain A and B in comparison to the macro algae-associated strain *Formosa agariphila* KMM 3901T. For better readability, the table is also attached on the DVD.

Glycoside Hydrolase Family	1	2	3	5	10	13	15	16	17	18	20	23	25	26	27	28	29	30	31	33	36	39	43	50	63	64	65	73	78	82	88	89	92	95	97	105	109	110	113	114	116	117	120	127	129	NC	total
Formosa agariphila KMM 3901^T (3561 genes)	1	13	7			5	1	9	2		1	2	1	2		2	5	2	3			3	2			1	1	3		5		1	1	1	1		1				6				2		84
Formosa sp. A (1848 genes)		3	3	2	2			3	3			2						5					1	1			1																			2	28
Formosa sp. B (2532 genes)			2	1				3	2	2		2						1					1				1					4						2								21

GlycosylTransferase Family	2	4	5	9	19	20	28	30	50	51	NC	total
Formosa agariphila KMM 3901^T (3561 genes)	33	21	2	1	1	1	1	1	1	2	1	65
Formosa sp. A (1848 genes)	14	7	1	1	1		1	1		1		27
Formosa sp. B (2532 genes)	21	12	1	2	1		1	1		2		41

Polysaccharide Lyase Family	1	6	7	8	9	12	14	17	NC	total
Formosa agariphila KMM 3901^T (3561 genes)		3	4	1				1	1	10
Formosa sp. A (1848 genes)					1					1
Formosa sp. B (2532 genes)						1				1

Carbohydrate Esterase Family	1	4	6	7	9	10	11	14	15	NC	total
Formosa agariphila KMM 3901^T (3561 genes)	7	1	1			2	1	1			13
Formosa sp. A (1848 genes)	3					2	1	1			7
Formosa sp. B (2532 genes)			1			1	1	1			4

Carbohydrate-Binding Module Family	4	5	6	9	13	16	32	35	42	47	48	50	51	57	total
Formosa agariphila KMM 3901^T (3561 genes)			1	1			2				2	4			10
Formosa sp. A (1848 genes)	1		1	1								3			6
Formosa sp. B (2532 genes)												2			2

Formylglycine-dependent sulfatase family (S1)	7	8	11	14	15	16	17	19	20	22	23	24	25	27	28	29	30	O	NC	total
Formosa agariphila KMM 3901^T (3561 genes)	3	4	2		4	1	2	3		1			2	1	3				2	28
Formosa sp. A (1848 genes)	7	2	3		1	1		1												15
Formosa sp. B (2532 genes)			1		1			1										1	1	5

Supplementary TAB 2.2.3 Identity and bitscore values of the sequence comparison depicted in FIG 2.2.3 of the synteny analysis with the laminarin PULs of *Formosa* sp. strain Hel1_33_131 (*Formosa* B) and partial PUL sequences in the metagenomes from the spring bloom at Helgoland in 2009 (Teeling et al., 2012).

PUL 1

Contig	Identity	Bitscore	E-value
62762	99.683	3386	0
33580	99.395	7038	0
60478	98.98	690	0
60478	100	710	0
59666	100	9822	0
24070	99.852	4845	0

PUL 2

Contig	Identity	Bitscore	E-value
60332	99.66	2626	0
43208	99.856	1243	0
11922	100	4046	0
29536	99.933	2695	0
62924	100	315	8.97E-85
44510	99.826	3099	0
61980	100	630	8.93E-180
39376	99.832	1068	0
63976	100	504	8.95E-142
63217	100	646	0
63659	88.187	926	0
42380	99.876	5820	0
58299	99.581	1278	0
37199	100	471	5.29E-132
32517	99.823	13151	0

PUL 3

Contig	Identity	Bitscore	e-value
60995	99.754	13109	0
46749	97.71	668	0
08399	99.917	8700	0
16086	99.886	11068	0

Supplementary TAB 2.2.4 Laminarin-specific, PUL-encoded genes of *Formosa* B as detected in the metagenomes of four consecutive annual spring blooms from 2009 until 2012. Green: consecutive contigs; red: overlapping contigs; (): large gap (>300 nt); nd: not detectable. For better readability, the table is also attached on the DVD.

Locus_tag	protein_ID	annotation	07.04.2009	14.04.2009	04.05.2010	18.05.2010	26.05.2011	10.05.2012
PULs								
FORMB_10040	AOR28054	glycoside hydrolase, GH3 family	2	3	5	7	1	3()
FORMB_10050	AOR28055	glycoside hydrolase, GH16 family	1	1	2	1()	nd	1
FORMB_10060	AOR28056	PKD domain protein	1	2	1	nd	nd	1
FORMB_10070	AOR28057	PKD domain protein	2	3	3	1()	nd	2 (1)
FORMB_10080	AOR28058	SusD-like protein	1	2()	2	nd	nd	1
FORMB_10090	AOR28059	TonB-dependent receptor	2	4()	2	1()	nd	2()
FORMB_24670	AOR29484	glycoside hydrolase, GH2 family	3 (1)	6()	5	3	3	2()
FORMB_24680	AOR29485	glycoside hydrolase	2	5	2	2	1	2()
FORMB_24690	AOR29486	glycoside hydrolase, GH16 family	1	1()	1	2	1	nd
FORMB_24700	AOR29487	PKD domain protein	4	3()	2	3()	2()	3 (1)
FORMB_24710	AOR29488	hypothetical protein	6 (1)	7()	6	5	3	3()
FORMB_24720	AOR29489	glycoside hydrolase, GH17 family	1	1	2	3	1	nd
FORMB_24730	AOR29490	glycoside hydrolase, GH30 family	4 (1)	2()	4	3	2	1()
FORMB_24740	AOR29491	sugar transporter, MFS family	1	2()	5	2	2	2()
FORMB_24750	AOR29492	glycoside hydrolase, GH17 family	1	1	3	nd	nd	1()
FORMB_13580	AOR28402	TonB-dependent receptor	3	2()	4	5	nd	nd
FORMB_13590	AOR28403	SusD-like protein	1	2()	1()	4	nd	nd
FORMB_13600	AOR28404	beta-glucanase precursor	1	2	3	3()	nd	nd
FORMB_13610	AOR28405	beta-glucanase precursor	2	1	3	3	nd	nd

green: consecutive contigs
red: overlapping contigs
(): large gap (>300 nt)
nd: not detectable

Supplementary TAB 2.2.5 Relative abundance of *Formosa* A and B proteins in metaproteome samples taken in 2009 **(A)** and 2010 **(B)**. The table is too big and thus was the attached via DVD.

Supplementary TAB 2.2.6 Overview of *Formosa* B subproteome analyses performed in this study during growth on laminarin, glucose and chitin. The table is too big and thus was the attached via DVD.

Supplementary TAB 2.2.7 List of laminarin-induced *Formosa* B proteins and their predicted subcellular localization. The table is too big and thus was the attached via DVD.

Supplementary TAB 2.2.8 Crystallographic data and refinement statistics for FbGH17A.

Data collection	
X-ray source	DESY P11
Wavelength (Å)	1.0332
Space group	P2
Unit cell a, b, c (Å)	93.34 149.14 107.22
Unit cell α, β, γ (˚)	90 103.35 90
Resolution range, (Å)	104.32-2.60 (2.65-2.60)
R_{merge}	0.20 (0.766)
Completeness (%)	99.7 (99.9)
Redundancy	5.2 (5.4)
$\langle I/\sigma(I)\rangle$	6.9 (2.6)
No. of Reflections	453483 (23840)
No. Unique	87396 (4421)
Mosaicity	0.08
Refinement	
R_{work}/R_{free} (%)	20.5/23.9
No. Of Atoms	19699
Protein	18825
Calcium	2
Water	872
***B* factors**	
Overall	26.10
Protein	26.29
Calcium	41.43
Water	19.33
R.m.s. deviations	
Bond Lengths (Å)	0.013
Bond Angles (°)	1.668
Ramachandran statistics (%)	
Favored	98.3
Allowed	1.7
Outliers	0.0
PDB accession code	6FCG

Supplementary TAB 2.2.9 List of all peptidases of *Formosa* B detected in the proteome analyses. For better readability, the table is also attached on the DVD.

Peptidase family	locus tag	IP Glucose	IP Laminarin	IP Control	MP Glucose	MP Laminarin	MP Control	EP Glucose	EP Laminarin	EP Control
M1	FORMB_00840	0.001895	0.000124	0.000392	0.006115	0.005038	0.005132	0.001198	0.002791	0.002466
M1	FORMB_12550	0.000000	0.000000	0.000254	0.000000	0.000000	0.009440	0.009123	0.007364	0.005440
M1	FORMB_19820	0.060000	0.023290	0.740274	0.001019	0.001344	0.191432	0.299688	0.206446	0.191639
M1	FORMB_23640	0.000000	0.000000	0.001296	0.000000	0.000000	0.000000	0.001175	0.000000	0.000000
M3	FORMB_07630	0.000000	0.037299	0.450314	0.004533	0.013008	0.000047	0.005420	0.014382	0.113852
M3	FORMB_23530	0.190187	0.070842	0.019589	0.019831	0.027399	0.020216	0.046193	0.053483	0.058041
M10	FORMB_07880	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
M12B	FORMB_06690	0.023836	0.009927	0.021841	0.001029	0.002018	0.009261	0.038629	0.014942	0.007163
M13	FORMB_20170	0.025715	0.015324	0.005557	0.003155	0.007090	0.005914	0.020684	0.104199	0.037838
M14	FORMB_00380	0.001417	0.009965	0.009779	0.000185	0.000000	0.000000	0.002287	0.007172	0.001125
M16	FORMB_09380	0.024029	0.012135	0.437715	0.000000	0.000000	0.076576	0.157829	0.126061	0.082532
M16	FORMB_09390	0.011990	0.006973	0.347985	0.001250	0.000142	0.090606	0.166395	0.183824	0.126700
M20	FORMB_01710	0.000364	0.025698	0.061879	0.000751	0.007164	0.110541	0.043062	0.063127	0.460133
M20	FORMB_03530	0.000000	0.003325	0.026608	0.000000	0.000000	0.000000	0.002081	0.011624	0.009389
M22	FORMB_10870	0.000699	0.042688	0.043894	0.000000	0.000000	0.000135	0.000000	0.001535	0.001758
M23	FORMB_02290	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
M23	FORMB_07730	0.003124	0.000000	0.000000	0.005366	0.000342	0.000000	0.000000	0.012556	0.002393
M23	FORMB_12790	0.002116	0.000000	0.000942	0.000652	0.002044	0.001772	0.000000	0.000000	0.001261
M23	FORMB_13360	0.004189	0.002184	0.000000	0.004116	0.003654	0.001818	0.008287	0.012719	0.027815
M23	FORMB_15400	0.000888	0.000000	0.000000	0.000427	0.000000	0.002479	0.001262	0.000775	0.003295
M23	FORMB_25680	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
M24	FORMB_06030	0.001264	0.003658	0.003313	0.000000	0.005475	0.000000	0.000000	0.032316	0.000000
M24	FORMB_09420	0.004178	0.020680	0.001288	0.002105	0.004982	0.000079	0.019708	0.036304	0.003495
M28	FORMB_01090	0.002754	0.001337	0.501765	0.000000	0.000000	0.010933	0.018971	0.006823	0.007448
M28	FORMB_08950	0.000772	0.000000	0.003628	0.000000	0.000000	0.000000	0.020411	0.002906	0.000650
M28	FORMB_18170	0.001831	0.000000	0.000000	0.001466	0.002101	0.000000	0.000000	0.004216	0.000000
M28	FORMB_20990	0.002185	0.000296	0.044592	0.000000	0.000000	0.000296	0.007759	0.005149	0.001269
M41	FORMB_19270	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.002479	0.000000
M42	FORMB_08210	0.000000	0.000000	0.010675	0.000000	0.000000	0.033882	0.000000	0.000000	0.097556
M42	FORMB_26670	0.020477	0.034736	0.025754	0.001453	0.000983	0.000770	0.017551	0.045225	0.026447
M48	FORMB_00980	0.002212	0.000000	0.000518	0.000814	0.000624	0.006460	0.008621	0.006911	0.008748
M49	FORMB_20750	0.182549	0.097687	0.313915	0.074634	0.116938	0.356023	0.083252	0.175352	0.229734
M50B	FORMB_14030	0.028720	0.014679	0.002713	0.025072	0.036764	0.032413	0.001244	0.021962	0.007299
M79	FORMB_24930	0.000000	0.000000	0.000000	0.000000	0.005918	0.000000	0.000000	0.000000	0.000000
S1C	FORMB_04130	0.000000	0.000154	0.021363	0.000000	0.000170	0.006325	0.000000	0.001109	0.050987
S8A	FORMB_01970	0.002341	0.000409	0.000175	0.003984	0.003150	0.003346	0.002872	0.002936	0.000000
S8	FORMB_02630	0.000000	0.000000	0.003535	0.000000	0.000000	0.002899	0.000000	0.000000	0.000000
S8	FORMB_07470	0.000000	0.000000	0.002888	0.000000	0.000000	0.009629	0.000000	0.000000	0.041947
S8	FORMB_12560	0.000000	0.000000	0.000067	0.000000	0.000000	0.007305	0.009562	0.015751	0.010679
S9A	FORMB_23080	0.018069	0.012131	0.032669	0.005135	0.004581	0.002745	0.009176	0.022147	0.001447
S9	FORMB_03300	0.002129	0.005300	0.000000	0.000280	0.002494	0.000886	0.002819	0.018391	0.007672
S9	FORMB_03310	0.002656	0.001440	0.000000	0.002954	0.013533	0.000000	0.002383	0.008025	0.000000
S9	FORMB_20360	0.052699	0.007312	0.100110	0.000000	0.000000	0.003119	0.008830	0.019036	0.012832
S9	FORMB_21490	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
S10	FORMB_23430	0.020579	0.011494	0.026383	0.000000	0.000000	0.005048	0.056238	0.032718	0.003755
S14	FORMB_25320	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
S16	FORMB_00960	0.002129	0.002219	0.008293	0.012407	0.020804	0.001607	0.008324	0.018205	0.040229
S33	FORMB_26700	0.057257	0.032444	0.009696	0.037749	0.045906	0.025996	0.015100	0.032552	0.009058
S41A	FORMB_17660	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
S41A	FORMB_24010	0.034290	0.015790	0.011595	0.145815	0.081913	0.140121	0.046267	0.072711	0.038406
S41	FORMB_12530	0.000828	0.000000	0.000000	0.000000	0.000000	0.000173	0.000000	0.000312	0.000000
S41	FORMB_14380	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.001172	0.001499	0.000472
S41	FORMB_20450	0.000000	0.000000	0.000694	0.000000	0.000000	0.000000	0.001617	0.004824	0.002706
S49	FORMB_18980	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
S54	FORMB_00740	0.000000	0.000000	0.000000	0.002669	0.015472	0.000000	0.000000	0.000000	0.000000
S54	FORMB_01580	0.000000	0.000000	0.000000	0.000407	0.012675	0.001060	0.000000	0.000000	0.000698
S54	FORMB_01590	0.000571	0.000000	0.000000	0.000719	0.000230	0.000000	0.000000	0.000000	0.000000
S54	FORMB_07490	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
S66	FORMB_19940	0.003428	0.008838	0.022231	0.000000	0.000000	0.000087	0.000000	0.004298	0.000000
C25	FORMB_18120	0.002882	0.000682	0.009450	0.001296	0.001493	0.005611	0.004593	0.006506	0.011717
C39	FORMB_06070	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
C39	FORMB_08090	0.000000	0.000000	0.000000	0.000052	0.004949	0.000000	0.000000	0.000000	0.000000
C40	FORMB_21950	0.000000	0.028812	0.119435	0.000000	0.000000	0.000000	0.000000	0.000000	0.002380
A8	FORMB_10010	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
A8	FORMB_15320	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
A28	FORMB_15200	0.037837	0.010792	0.003797	0.000000	0.000000	0.000000	0.000000	0.001781	0.000000
unknown	FORMB_05990	0.002664	0.029636	0.096803	0.000000	0.000000	0.000000	0.000000	0.002499	0.011705
unknown	FORMB_06930	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
unknown	FORMB_01540	0.008102	0.003231	0.000000	0.026711	0.028858	0.011940	0.004415	0.005201	0.007449
unknown	FORMB_15100	0.015455	0.000831	0.000885	0.004623	0.003178	0.001277	0.000000	0.005746	0.000056
unknown	FORMB_26180	0.528378	0.411974	0.067840	0.454888	0.852356	0.251292	0.356181	0.359978	0.380939
sum (%riBAQ)		1.389685	1.016335	3.614388	0.853664	1.334792	1.446690	1.510380	1.798866	2.150617

2.3 Manuscript III: Laminarin quantification in microalgae with enzymes from marine microbes

Stefan Becker[1,2] and Jan-Hendrik Hehemann[1,2]

[1] MARUM – Center for Marine Environmental Sciences at the University of Bremen, Germany
[2] Max-Planck-Institute for Marine Microbiology, Bremen, Germany

In print at: Bio-Protocol

Contributions to the manuscript (% of SB's contribution to the total workload):
SB (35%) and JHH designed the study. SB (100%) performed experimental work and data analysis. SB (100%) prepared figures and tables. SB (80%) and JHH wrote the manuscript.

2.3.1 Abstract

The marine beta-glucan laminarin is an abundant storage polysaccharide in microalgae. High production rates and rapid digestion by heterotrophic bacteria turn laminarin into an ideal carbon and energy source, and it is therefore a key player in the marine carbon cycle. As a main storage glucan laminarin also plays a central role in the energy metabolism of the microalgae (Painter, 1983; Myklestad, 1974; Percival and Ross, 1951). We take advantage of enzymes that digest laminarin selectively and can thereby quantify only this polysaccharide in environmental samples. These enzymes hydrolyze laminarin into glucose and oligosaccharides, which are measured with a standard reducing sugar assay to obtain the laminarin concentration. Prior to this assay, the three enzymes need to be produced via heterologous expression and purification. The assay can be used to monitor laminarin concentrations in environmental microalgae, which were concentrated from seawater by filtering, or in samples derived from algal lab cultures.

2.3.2 Background

Marine polysaccharides play an important role in the marine carbon cycle and are a major part of the physiology of phytoplankton, but are severely understudied. For decades, the agro-food industry has been using ready-to-use kits based on enzymatic assays to analyze a wide range of different polysaccharides in their processes (Whitaker, 1974). These fast, robust and specific enzyme based methods assess polysaccharides originating from land-based plants, i.e., starch, as they are widely used in food, feed and other industrial applications (Brunt et al., 1998). However, similar assays for marine polysaccharides are still lacking. Inspired by the idea of using enzymes for polysaccharide quantification in algae, we developed an enzyme-based method to quantify the ecologically relevant beta-glucan laminarin, also known as chrysolaminarin, in diatoms and other microalgae.

The three glycoside hydrolases (GH) for this application are from *Formosa* spp. and they were characterized as follows: FbGH30 is an exo-acting β-1,6-glucanase of the GH30 family, specifically hydrolysing the β-1,6-linked glucose monomer branches attached to the laminarin backbone; and FaGH17A and FbGH17A are two endo-acting β-1,3-glucanases of the GH family 17, which acts specifically on the β-1,3-linked laminarin backbone (Becker et al., 2017; Unfried et al. unpublished).

This method enables the quantification of laminarin in crude substrate mixtures, without the need for purification of the laminarin. This enzymatic method is fast, does not require sophisticated instruments, the enzymes are stereospecific and they selectively cleave laminarin into glucose and oligosaccharides, which can be quantified with a common reducing sugar assay. The method can be easily applied in fieldwork. The assay itself comprises only the three steps of extraction, hydrolysis and the reducing sugar assay (FIG 2.3.1). It can be done within only a few hours. The limit of detection (LOD) of the assay is at 1.5 µg/ml. The three enzymes need to be

produced only once and can be stored for years. After their production and purification, one has enough material to analyze thousands of samples. We decided to include the plasmid transformation and recombinant enzyme production part into the protocol, since we consider these steps feasible to be done by marine labs with less experience in biotechnology.

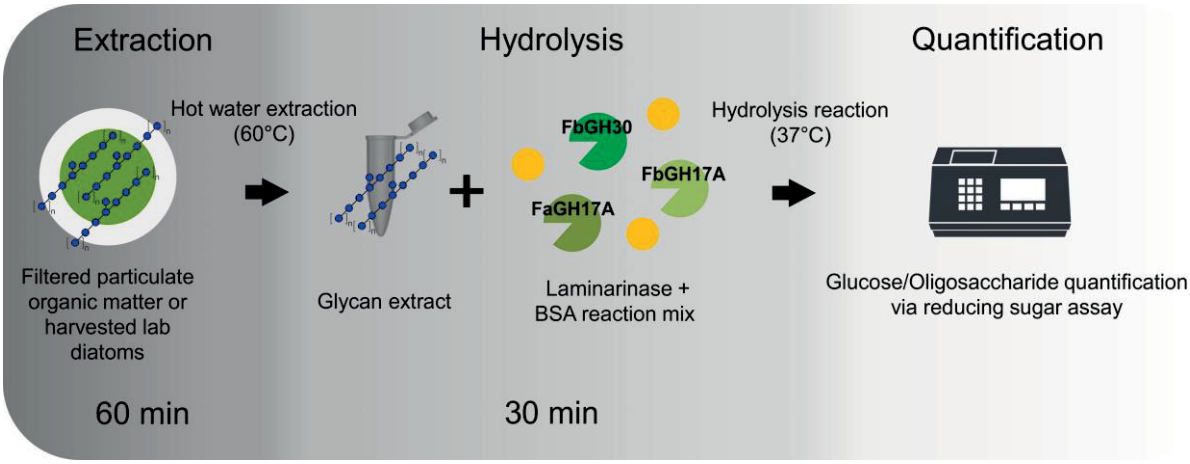


FIG 2.3.1 Schematic protocol after the production of the enzymes. A brief outline of the three main steps and their approximate duration.

2.3.3 Materials and reagents

1. 2.0 ml reaction tubes (SARSTEDT, catalog number: 72.691)
2. Sterile toothpick
3. Aluminium foil
4. 5 ml HiTrap IMAC HP column (GE Healthcare, catalog number: 17-0920-05)
5. Polyethersulfone ultrafiltration membrane with 10-kDa cutoff value (Merck, catalog number: PBGC04310)
6. 5 ml HiTrap Desalting column (GE Healthcare, catalog number: 29-0486-84)
7. Whatman GF/F filters (GE Healthcare, catalog number: 1825-047), size is dependent on used filtration setup
8. 50 ml centrifuge tubes (SARSTEDT, catalog number: 62.547.254)
9. Half-micro-cuvettes (BRAND, catalog number: 759015)
10. Petri dishes (SARSTEDT, catalog number: 82.1473)
    Note: Alternatively a Pierce Protein Concentrator PES, 10K MWCO, 5-20 ml (Thermo Fisher Scientific, Thermo ScientificTM, catalog number: 88527) (see Note 2)
11. BL21 (DE3) competent E. coli (New England Biolabs, catalog number: C2527I)
12. SOC-medium (is usually enclosed in BL21 (DE3) competent E. coli product)
13. FaGH17A plasmid (Addgene, catalog number: 86462) (see Note 1)
14. FbGH30 plasmid (Addgene, catalog number: 86463) (see Note 1)
15. FbGH17A plasmid (Addgene, catalog number: 100911) (see Note 1)

16. IPTG (Isopropyl β-D-thiogalactopyranoside from AppliChem, catalog number: A4773)
17. Lysozyme (from chicken egg white) (Sigma-Aldrich, catalog number: L6876)
18. Sodium chloride (NaCl) (Carl Roth, catalog number: 3957.2)
19. Tryptone (Sigma-Aldrich, catalog number: 95039)
20. Yeast extract (Sigma-Aldrich, catalog number: 09182)
21. Agar (Sigma-Aldrich, catalog number: A7002)
22. Kanamycin sulfate (Sigma-Aldrich, catalog number: 60615)
23. Ammonium sulfate ($(NH_4)_2SO_4$) (Sigma-Aldrich, catalog number: A4418)
24. Monopotassium phosphate (KH_2PO_4) (Sigma-Aldrich, catalog number: 9791)
25. Disodium phosphate (Na_2HPO_4) (Sigma-Aldrich, catalog number: 71642)
26. Glucose (Sigma-Aldrich, catalog number: G8270)
27. Lactose (Sigma-Aldrich, catalog number: 17814)
28. Glycerol (Sigma-Aldrich, catalog number: G6279)
29. Magnesium sulfate ($MgSO_4$) (Sigma-Aldrich, catalog number: 208094)
30. Sucrose (Sigma-Aldrich, catalog number: 84100)
31. Tris(hydroxymethyl)aminomethane (Tris) (Sigma-Aldrich, catalog number: RDD008)
32. Sodium deoxycholate (Sigma-Aldrich, catalog number: D6750)
33. Triton X-100 (Sigma-Aldrich, catalog number: X100)
34. DNase I (from bovine pancreas) (Sigma-Aldrich, catalog number: 69182)
35. Nickel(II) sulfate hexahydrate ($NiSO_4 \cdot 6H_2O$) (Sigma-Aldrich, catalog number: 227676)
36. Imidazole (Sigma-Aldrich, catalog number: I5513)
37. Dithiothreitol (DTT) (Sigma-Aldrich, catalog number: D0632)
38. 3-(N-Morpholino)propanesulfonic acid (MOPS) (Sigma-Aldrich, catalog number: 69947)
39. Laminarin from Laminaria digitata (Sigma-Aldrich, catalog number: L9634)
40. Bovine serum albumin (BSA), lyophilized (Sigma-Aldrich, catalog number: A2153)
41. 4-Hydroxybenzhydrazide (Sigma-Aldrich, catalog number: H9882)
42. HCl (37%) (Sigma-Aldrich, catalog number: 320331)
43. Trisodium citrate dihydrate (Sigma-Aldrich, catalog number: W302600)
44. Sodium hydroxide (NaOH) (VWR, catalog number: 28244.295)
45. Calcium chloride ($CaCl_2$)
46. LB-Kana-Agar-Plates (see Recipes)
47. LB-Kana-Medium (see Recipes)
48. 20x NPS stock (see Recipes)
49. 50x 5052 stock (see Recipes)
50. $MgSO_4$ stock (see Recipes)
51. ZY-Medium (see Recipes)
52. ZYP-5052-Rich-Autoinduction-Medium (see Recipes)

53. Sucrose stock (see Recipes)
54. Deoxycholate stock (see Recipes)
55. DNase I stock (see Recipes)
56. $NiSO_4$ stock (see Recipes)
57. IMAC buffer A (see Recipes)
58. IMAC buffer B (see Recipes)
59. SEC buffer (see Recipes)
60. SEC/DTT buffer (see Recipes)
61. MOPS buffer (see Recipes)
62. BSA stock solution (see Recipes)
63. PAHBAH reagent A (see Recipes)
64. PAHBAH reagent B (see Recipes)
65. PAHBAH working reagent (see Recipes)

2.3.4 Equipment

1. Bunsen burner or sterile hood
2. 2 L Erlenmeyer flasks (or normal 2 L glass bottles)
3. Rotating shaker for Erlenmeyer flasks (temperature control is recommended)
4. Vortex mixer
5. Microcentrifuge for 1.5 or 2.0 ml reaction tubes with cooling function, i.e., Centrifuge 5418 R (Eppendorf, model: 5418 R, catalog number: 5401000013)
6. Centrifuge for 15 and 50 ml tubes with cooling function, i.e., Centrifuge 5804 R (Eppendorf, model: 5804 R, catalog number: 5805000017)
7. Standard membrane or peristaltic pump and respective filtration setup, i.e., ME 1 (Vacuubrand, model: ME 1) and Nalgene filter holder receiver (Thermo Fisher Scientific, Thermo ScientificTM, catalog number: 300-4100)
8. Standard fast protein liquid chromatography system (FPLC), i.e., ÄKTA start (GE Healthcare, model: ÄKTA start, catalog number: 29-0220-94)
   Note: Alternatively a peristaltic pump with a flow rate of $\leq$ 5 ml/min.
9. Spectrometer, i.e., BioSpectrometer basic (Eppendorf, model: BioSpectrometer® basic, catalog number: 6135000009)
10. Heated water bath, i.e., Thermolab (GFL, catalog number: 1070)
11. Two heat blocks, i.e., BioShake iQ (Analytik Jena, catalog number: 848-1808-0506)
12. Autoclave (is strongly recommended but not essential, since one always uses the antibiotic kanamycin for the bacterial growth)
13. Ultrafiltration stirred cell, i.e., Amicon (Merck, catalog number: UFSC05001)

Note: Alternatively a Pierce Protein Concentrator PES, 10K MWCO, 5-20 ml (Thermo Fisher Scientific, Thermo ScientificTM, catalog number: 88527) (see Note 2).

2.3.5 Procedure

A Transformation:

1. Transform each Addgene plasmid containing one enzyme into separate BL21 (DE3) competent E. coli according to the manufacturer's instructions and by using LB-Kana-Agar-Plates (see Recipes). A common transformation protocol uses the following steps, which have to be carried out next to a Bunsen burner or in a sterile hood.
2. Add 1 µl containing ca. 100 ng of plasmid DNA to the cell mixture and carefully invert the tube once. Do not vortex.
3. Keep on ice for 5 min and do not mix.
4. Heat shock at exactly 42°C for exactly 10 s and do not mix.
5. Keep on ice for 5 min and do not mix.
6. Add 500 µl SOC-medium, which is included in the E. coli BL21 (DE3) shipment.
7. Incubate at 37°C for 60 min and at 500 rpm.
8. Spread 100 µl on an LB-Kana-Agar-Plate and incubate for at least 15 h at 37°C.
9. After picking a colony with a sterile toothpick and growing it for at least 8 h in a 3 ml LB-Kana-Medium (see Recipes) preculture at 37°C and under constant stirring at 150 rpm, the cells can be stored at -80°C after adding 25% (v/v) sterile glycerol (i.e., 200 µl of glycerol in 600 µl of preculture).

B Enzyme overexpression and purification

1. Use 1 ml of the preculture from step A9 to inoculate 1 L of ZYP-5052-Rich-Autoinduction-Medium (see Recipes) (Studier, 2005) in a 2 L Erlenmeyer flask to have enough headspace. Close the flask only with aluminum foil to allow gas exchange into the flask.

   Notes: a. Overexpression and purification are performed in almost the exact same way for all three enzymes (FbGH30, FaGH17A and FbGH17A). However, each purification has to be performed separately. It is not possible to transform all three enzymes into one expression system and purify all together. Nonetheless, after the overexpression, it is possible to purify all three enzymes within one working day.

   b. Overexpression can also be performed in IPTG-induced LB-Kana-Medium. The inoculated cultures need to grow for 8 h at 37°C before 1 ml 1 M IPTG is added. After induction, the cultures need to be incubated for 8 h more at 16°C. Subsequent harvesting and purification are performed as already described. This method is faster but yields lower protein amounts.

2. Grow cultures for about 72 h at 20°C [or room temperature (RT)] and with rotation at 150 rpm.
3. Harvest cells by centrifugation at 4,500 x g for 25 min at 4°C. Discard supernatant. The resulting pellet can be stored at -20°C (for years) until further use.
4. Resuspend the cell pellet in 15 ml sucrose stock solution (see Recipes).
5. Add 5 mg of lysozyme and incubate for 10 min at RT under constant magnetic stirring at 500 rpm.
6. Add 30 ml deoxycholate stock solution (see Recipes) and 0.2 ml $MgSO_4$ stock solution (see Recipes).
7. Add 1 ml DNaseI stock solution (see Recipes) to liquefy the viscous solution.
8. Centrifuge the lysate at 16,000 x g for 45 min at 4°C. Transfer the supernatant containing the proteins of interest into a new tube.
9. All steps during the following immobilized metal affinity chromatography (IMAC) are carried out with an FPLC at RT and a flow rate of 5 ml/min. Alternatively, the procedure can also be performed by using a peristaltic pump at the same or lower flow rates.

   Note: Steps B10-B14 can be performed manually using a peristaltic pump at 5 ml/min flow rate instead of an FPLC system. The column is then placed after the pump. The linear buffer gradient in step B14 can be achieved similarly by using a stepwise increase of the imidazole concentration over the course of the total elution volume of 5 CV (~5 min). Starting with IMAC buffer A containing 50 mM imidazole, the buffer needs to be changed every 30 s. At each step the imidazole concentrations increases by 50 mM (30 s each): 50, 100, 150, 200, 250, 300, 350, 400, 450, 500 mM imidazole (the usual concentration for IMAC buffer A and B are 25 and 500 mM Imidazole, respectively). To prevent air bubbles on the column, the flow needs to be paused every time the inlet is switching from one buffer to the other.
10. Charge a 5 ml HiTrap IMAC HP column with 1 column volume (1 CV = 1 ml) of $NiSO_4$ stock solution (see Recipes). Discard the flow that comes off the column.
11. Equilibrate with 5 CVs of IMAC buffer A (see Recipes). Discard the flow that comes off the column.
12. Inject the entire supernatant of the lysis onto the column. If the solution is too viscous, it needs to be diluted up to a volume of 50 ml using IMAC buffer A. Discard the flow that comes off the column.
13. Wash the column with 15 CVs IMAC buffer A. Discard the flow that comes off the column.
14. Elute the protein in a total number of 5 CVs of a continuous linear gradient from IMAC buffer A to IMAC buffer B (see Recipes). Collect the flow that comes off the column in thirty 1 ml fractions.

15. For FbGH17A only: Add 50 µl of the reducing agent DTT (0.5 mg/ml) to each fraction immediately after the elution to prevent oxidation of the protein.
16. Pool fraction 6 to 30 (if there is a UV detector, take only the samples corresponding to the major peak). Keep the protein on ice.
17. The IMAC column has to be cleaned after each purification according to the manufacturer's instruction.
18. Assemble an ultrafiltration stirred cell together with a respective 10 kDa membrane. (alternatively, a centrifugation concentrator can be used, see Note 2)
19. Concentrate the protein solution in the ultrafiltration chamber at 4°C or on ice down to a volume of ≤ 1.5 ml.
20. Centrifuge the protein solution for 20 min at 13,000 x g and 4°C, to remove precipitated, inactive protein.
21. Transfer supernatant into a fresh tube and discard the pellet. Keep the protein on ice.
22. The following use of a desalting column removes the imidazole. The procedure followed the manufacturer's instructions and can be performed by using an FPLC system or a syringe at RT and a flow rate of 5 ml/min.
23. For FbGH17A only: Use SEC/DTT buffer (see Recipes) instead of SEC buffer in steps B24-26.
24. Equilibrate a HiTrap Desalting column with 5 CVs of SEC buffer. Discard the flow that comes off the column.
25. Apply exactly 1.5 ml of the protein solution. Dilute the sample with SEC buffer if necessary. Do not exceed the sample volume of 1.5 ml. Discard the flow that comes off the column.
26. Elute and collect the protein with 2 ml of SEC buffer and keep the protein on ice afterwards.
27. The desalting column has to be cleaned after the purification according to the manufacturer's instruction.
28. Determine protein concentration by measuring the absorbance at 280 nm using a spectrometer and under consideration of the molecular weight and the extinction factor of each protein (FbGH30: MW = 54,700 Da, ε = 128,480 M^{-1} cm^{-1}; FaGH17A: MW = 44,800 Da, ε = 87,905 M^{-1} cm^{-1}; FbGH17A: MW = 46,600 Da, ε = 88,935 M^{-1} cm^{-1}):

    $$c = A_{280} \times ([\varepsilon / MW] \times b)$$

    where, c: protein concentration in mg/ml; A: absorbance at 280 nm; ε: extinction factor in M^{-1} cm^{-1}; MW: molecular weight in Da; b: light path length in cm.

    Note: The success of overexpression (steps B1-B3), lysis (steps B4-B8), IMAC (steps B9-B17), concentration (steps B18-B21) and desalting (steps B22-B27) can be verified by SDS-Polyacrylamide gel electrophoresis (He, 2011).
29. Dilute the proteins in MOPS buffer (see Recipes) to a stock concentration of 10 µM. Keep protein on ice.

30. For FbGH17A only: After dilution in MOPS buffer the enzyme must be used immediately, either for preparing aliquots (step B31) or for hydrolysis reactions (steps D1-D7).
31. Prepare ready-to-use 50 µl aliquots and freeze them at -20°C.
    Note: So far, there is no very long-term data on the shelf life of the enzymes at -20°C, but they can be used for at least two years without any decrease in activity.

C Sampling and extraction

1. Environmental samples, i.e., seawater, or lab cultures of microalgae are filtered onto Whatman GF/F filters at 0.1-0.5 bar by using a standard membrane or peristaltic pump and a respective filtration setup. The required volume needs to be determined by experiments. In case of sea surface water, one needs at least 5 L, for highly concentrated lab cultures the demand is much smaller (~50 ml, depending on the cell numbers). The volume must be logged for each sample. The resulting filter can be stored at -20°C until further use.
    Notes: a. Polycarbonate filters can be used as well. The diameter of the filters depends on the filtration setup.
    b. For concentrated lab cultures, centrifugation can be used instead of filtration as well. Diatom cultures need to be centrifuged at 4,500 x g for 20 min.
    c. The method can be applied to macroalgal samples as well (in preparation L. Scheschonk et al.).
2. Place filter into a 15 or 50 ml centrifuge tube.
3. Add 5-10 ml of MOPS buffer, just enough so that the entire filter is covered with liquid.
4. Vortex every sample for 10 s.
5. Extract for 60 min at 60°C in a heated water bath.
6. Squeeze and take out the filter.
7. Centrifuge for 15 min at 4,500 x g.
8. Transfer the supernatant to a new tube. This extract can be stored for several weeks at -20°C until further use.
    Note: There is no long-term data on the stability of laminarin samples.

D Hydrolysis

1. Each sample is split up into six subsamples: three are hydrolyzed by the enzyme mixture and three are not hydrolyzed.
2. Additionally, one has to prepare and treat (steps C3 and C5-C7) samples for creating a calibration curve based on commercial laminarin from Laminarin digitata (Sigma-Aldrich). For microalgal samples, the following concentrations are suitable: 0, 7.8 µg/ml, 15.6 µg/ml, 31.3 µg/ml, 62.5 µg/ml, 125 µg/ml, 250 µg/ml, 0.5 mg/ml, 1.0 mg/ml, and 2.0 mg/ml (serial dilution in MOPS buffer).

3. Thaw fresh aliquots of each enzyme stock solution and keep them on ice.
4. Each hydrolyzation reaction mix consists of: 100 µl sample extract (or calibration standard), 1 µl FbGH30, 1 µl FaGH17A, 1 µl FbGH17A and 1 µl BSA stock solution (see Recipes) (see Note 3).
5. Each non-hydrolyzation reaction mix consists of: 100 µl sample extract, 1 µl BSA stock solution and 3 µl of MOPS buffer (see Note 11).
6. Mix all samples by shaking the reaction tubes.
7. Incubate for 25 min at 37°C in a heat block.
9. Stop enzyme reaction by incubating the samples for 5 min at 99°C in a second heat block. This extract can be stored for several weeks at -20°C until further use.
   Note: There is no long-term data on the stability of laminarin samples.

E Reducing sugar assay

1. Add 1 ml of PAHBAH working reagent (see Recipes) (Lever, 1972) to 0.1 ml of the sample.
2. Heat each sample for exactly 5 min at 100 °C in a heat block.
   Note: The duration of boiling directly influences the absorbance and therefore the standard deviation of the measured triplicates. One should try to take the samples out of the heat block at the exact same speed that is used for putting them in. By doing this one can make sure that every sample is getting heated for the exact same duration.
3. Determine absorbance at 410 nm using half-micro-cuvettes and a spectrometer.
4. Use PAHBAH working reagent as blank and for the dilution of samples that are too concentrated.
   Note: The range of absorbance in which one can make reliable measurements depends on the spectrometer and must be checked in the manual in advance. The samples can be diluted in PAHBAH working reagent.

2.3.6 Data analysis

1. Each sample is measured in six subsamples. The value of the non-hydrolyzed triplicates needs to be subtracted from the value of the hydrolyzed samples.
2. The laminarin concentration in the respective sample can be determined by comparing this difference to the calibration curve.
3. The concentration, which was measured in the extract, needs to be converted into the concentration of the original sample, by taking the exact filtration volume or the cell number into account.
4. Based on the calibration below, the equation and the data for laminarin concentrations in samples of 20 L are as follows (FIG 2.3.2, TAB 2.3.1):

$$c_{Lam} = (((Mean_{hydrolyzed} - Mean_{non\text{-}hydrolyzed})/3.0852)\text{-}0.0076) \times (1000/20)$$

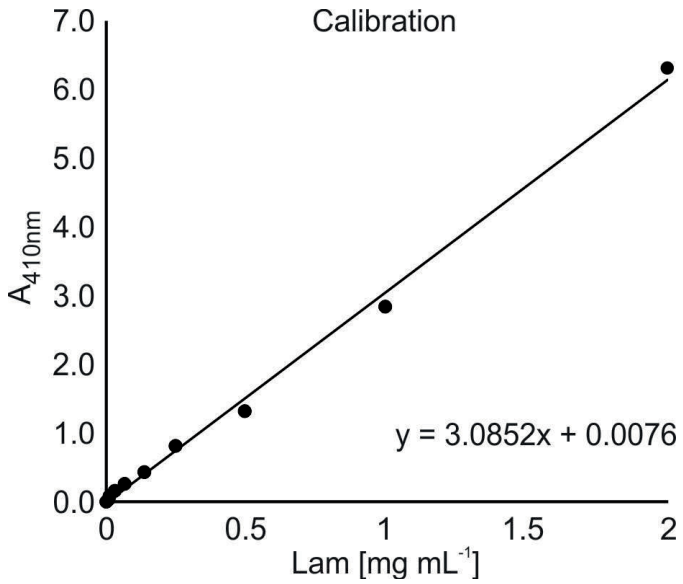


FIG 2.3.2 Exemplary calibration

TAB 2.3.1 Exemplary data

		Replicates			Mean	Sample blank	Laminarin [mg/L]
hydrolyzed samples	sample 1	0.956	0.979	0.964	0.966	0.259	5.69
	sample 2	0.570	0.615	0.597	0.594	0.127	2.61
	sample 3	0.836	0.851	0.851	0.846	0.222	4.83
non-hydrolyzed blanks	sample 1	0.694	0.706	0.722	0.708		
	sample 2	0.466	0.475	0.459	0.467		
	sample 3	0.623	0.629	0.620	0.624		

2.3.7 Notes

1. All three plasmid can be ordered by academics and non-profits via the Addgene plasmid repository (www.addgene.org). On the website, one can find protocols for the purification of the plasmid DNA if Addgene provides only bacterial stabs. The plasmids must get transformed into E. coli BL21 (DE3).
2. For protein concentration, a centrifugation-based system can be used alternatively. These tubes are not reusable and the recovery rate is quite low. However, for a single usage, it is cheaper compared to the ultrafiltration system. The concentration has to be conducted according to the manufacturer's instruction and at 4°C. The volume of the protein solution needs to be decreased to a volume of ≤ 1.5 ml.
3. It is possible and advisable to prepare master mixes containing all components that have to be added to each sample extract.

2.3.8 Recipes

Note : Milli-Q water was used to make up the following solutions unless otherwise indicated.

1. LB-Kana-Agar-Plates
   a. Dissolve:
      10 g NaCl
      10 g tryptone
      5 g yeast extract
      12.5 g agar
      Complete the volume to 1 L with non-sterile water
   b. Autoclave or boil the solution
   c. After letting the solution cool down to a temperature below 50°C, add 1 ml of a prepared kanamycin sulfate (50 mg/ml)
   d. After stirring the solution, the plates need to be poured next to a Bunsen burner or in a sterile hood
   e. When the plates are solid, they can be stored at 4°C for several weeks
2. LB-Kana-Medium
   a. Dissolve:
      10 g NaCl
      10 g tryptone
      5 g yeast extract
      Complete the volume to 1 L with non-sterile water in a 2 L Erlenmeyer flask or bottle
   b. Close the flask with aluminum foil but keep it air-penetrable
   c. It is recommended to autoclave the solution
   d. After letting the solution cool down to a temperature below 50°C, add 1 ml of a prepared kanamycin sulfate (50 mg/ml)
3. 20x NPS stock
   Dissolve:
   66 g $(NH_4)_2SO_4$
   136 g KH_2PO_4
   142 g Na_2HPO_4
   Complete the volume to 1 L with non-sterile water
   It is recommended to autoclave the solution before use
4. 50x 5052 stock
   Dissolve:
   25 g glucose
   100 g lactose
   250 g glycerol (weigh in a beaker) and complete the volume to 1 L with non-sterile water

It is recommended to autoclave the solution before use

5. $MgSO_4$ stock

   Dissolve 12 g $MgSO_4$ and complete the volume to 100 ml with non-sterile water

   It is recommended to autoclave the solution before use

6. ZY-Medium

   Dissolve 10 g tryptone and 5 g yeast extract in 928 ml H_2O

   It is recommended to autoclave the solution before use

7. ZYP-5052-Rich-Autoinduction-Medium

   Mix 928 ml ZY-Medium

   1 ml 1 M $MgSO_4$

   20 ml 50x 5052

   50 ml 20x NPS and 2 ml kanamycin sulfate (50 mg/ml) next to a Bunsen burner

8. Sucrose stock

   a. Dissolve 250 g sucrose and 6 g Tris
   b. Adjust to pH 7.5 and complete the volume to 1 L with non-sterile water
   c. It is recommended to autoclave the solution before use

9. Deoxycholate stock

   a. Dissolve:

      10 g sodium deoxycholate

      10 ml Triton X-100

      5.8 g NaCl

      2.4 g Tris

   b. Adjust to pH 7.5 and complete the volume to 1 L with non-sterile water
   c. It is recommended to autoclave the solution before use

10. DNase I stock

    Dissolve 50 mg DNase I in 35 ml IMAC buffer A and 15 ml glycerol

    The solution can be stored at -20°C

11. $NiSO_4$ stock

    Dissolve 3.9 g $NiSO_4$ and complete the volume to 50 ml

    Note: $NiSO_4$ is harmful. Protect yourself by using gloves, eye and mouth protection. Discard contaminated solutions in an appropriate heavy metal waste.

12. IMAC buffer A

    a. Dissolve 29.2 g NaCl, 2.4 g Tris and 1.7 g imidazole
    b. Adjust to pH 7.5 and complete the volume to 1 L with non-sterile water
    c. It is recommended to sterile filter the solution before use

13. IMAC buffer B
    a. Dissolve 29.2 g NaCl, 2.4 g Tris and 34 g imidazole
    b. Adjust to pH 7.5 and complete the volume to 1 L with non-sterile water
    c. It is recommended to sterile filter the solution before use
14. SEC buffer
    a. Dissolve 29.2 g NaCl and 2.4 g Tris
    b. Adjust to pH 7.5 and complete the volume to 1 L with non-sterile water
    c. It is recommended to sterile filter the solution before use
15. SEC/DTT buffer
    Dissolve 23 mg dithiothreitol in 50 ml SEC buffer
16. MOPS buffer
    a. Dissolve 5.2 g 3-(N-Morpholino)propanesulfonic acid
    b. Adjust to pH 7.0 and complete the volume to 500 ml
    c. It is recommended to sterile filter the solution before use
    d. Use a brown glass bottle and protect from light
17. BSA stock solution
    Dissolve 100 mg BSA in 1 ml of H_2O
    It is recommended to sterile filter the solution before use
18. PAHBAH reagent A
    a. Dissolve 10 g 4-hydroxybenzhydrazide in 60 ml H_2O
    b. Add 10 ml concentrated HCl (37%) and complete the volume to 200 ml by adding more H_2O
    Note: The solution can be stored at RT. Do not use if precipitate is present in the solution.
19. PAHBAH reagent B
    a. Dissolve 24.9 g trisodium citrate dehydrate in 500 ml H_2O
    b. Add 2.2 g $CaCl_2$ and mix
    c. Dissolve 40 g NaOH in a separate 1 L bottle
    d. Mix both solutions slowly and under constant stirring and complete the volume to 2 L with non-sterile water
e. The solution must become clear within several minutes
    Note: The solution can be stored at RT. Do not use if precipitate is present in the solution.
20. PAHBAH working reagent
    Prepare a fresh 9:1 mixture of PAHBAH reagent B and PAHBAH reagent A.
    Can be used for several hours, but needs to be kept on ice.

2.3.9 Acknowledgments

This protocol was adapted and modified from Becker et al., 2017. The research was supported by the Deutsche Forschungsgemeinschaft (grant HE 7217/1-1 to Jan-Hendrik Hehemann) and by the Max Planck Society.

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

2.4 Manuscript IV: Laminarin is the major marine sugar polymer and indicates bioenergy states of the surface ocean

Stefan Becker[1,2], Sarah Coffinet[1], Kai-Uwe Hinrichs[1], Karen Wiltshire[3], Morten Iversen[1,3] and Jan-Hendrik Hehemann[1,2]

[1] MARUM – Center for Marine Environmental Sciences at the University of Bremen, Germany
[2] Max-Planck-Institute for Marine Microbiology, Bremen, Germany
[3] Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany

In preparation.

Contributions to the manuscript (% of *SB's* contribution to the total workload):
SB (40%) and JHH designed the study. *SB* (20%) participated in environmental sampling. *SB* (90%) and SC conducted experimental work. *SB* (100%) performed data analysis and prepared figures and tables. *SB* (50%) and JHH wrote the manuscript. KW provided chlorophyll data. KUH and MI were involved in the discussion and editing of the manuscript.

2.4.1 One sentence summary

This work identifies the algal glucan laminarin as one of Earths most synthesized bioenergy molecules

2.4.2 Abstract

Marine algae sequester as much CO_2 into carbohydrates as terrestrial plants, yet the polysaccharide composition of the surface ocean remains unknown due to analytic challenges. Here we quantified algal laminarin using a biocatalytic assay, in which bacterial enzymes specifically cleave this polysaccharide into analyzable fragments. We measured laminarin in the Arctic, North-, Central- and South-Atlantic, coastal Pacific and in two North Sea time series and found that it accounts on average for 37±19% of the carbon in particulate organic matter. An observed correlation between laminarin and chlorophyll suggests that 18±9 gigatons of laminarin are produced annually. Marked variation of laminarin levels in different oceanic regions indicates a potentially strong influence of this molecule on the bioenergetic state of the surface ocean.

2.4.3 Report

The production rate of organic carbon is controlled by photosynthesis of microscopic algae in the sunlit ocean surface. Here, diatoms alone, which contribute about 40% of the marine primary production, are thought to convert more carbon dioxide into biomass than tropical forests (Field et al., 1998). Among the most abundant molecules synthesized by algae are glycans (carbohydrates composed of multiple, linked monosaccharides). However, our insights regarding the structure and diversity of glycans in nature fall short of those obtained for proteins by proteomics (Teeling et al., 2012) describing the enzymes and transporters used by heterotrophic microbes to degrade glycans and genomics (de Vargas et al., 2015) identifying the planktonic organisms that drive glycan production in the surface ocean. One reason is that the inherent structural complexity of glycans creates such vast molecular diversity whose elucidation poses a major analytical challenge for environmental glycomics (Gal et al., 2016; Senni et al., 2011; Biersmith and Benner, 1998; Biddanda and Benner, 1997; Painter, 1983).

Glycans are crucial for carbon export because they can aggregate and sink, which increases carbon sequestration (Engel et al., 2004). They were also found to accumulate in surface waters indicating their potential to store carbon in dissolved organic matter (Aluwihare et al., 1997). Finally, they are an important carbon and energy source for heterotrophs as judged from extensive expression of bacterial carbohydrate active enzymes and sugar transporters during algal blooms (Teeling et al., 2012). However, which glycans are foods or carbon sinks remains unconstrained. Carbohydrates build the most diverse and complex macromolecules in nature (Hofmann et al., 2015; Laine, 1994). This diversity of glycans is rooted in the sheer endless

possibilities of connecting different monomers (Supplementary FIG 2.4.1). Consequently, glycans can be branched whereas proteins and DNA are strictly linear. Proper identification and quantification of a glycan requires determination of its monomer types, the configuration of its glycosidic bonds, and location of the connectivity between bonds; this requirement explains why marine glycomics is still in its infancy. Most studies have focused on the composition of monomers (Engel et al., 2012), whereas only few studies have also constrained connectivity but none have elucidated configuration (Repeta and Aluwihare, 2006; Pakulski and Benner, 1994). Consequently we do not know the quantity, role and cycling of specific glycans in the ocean.

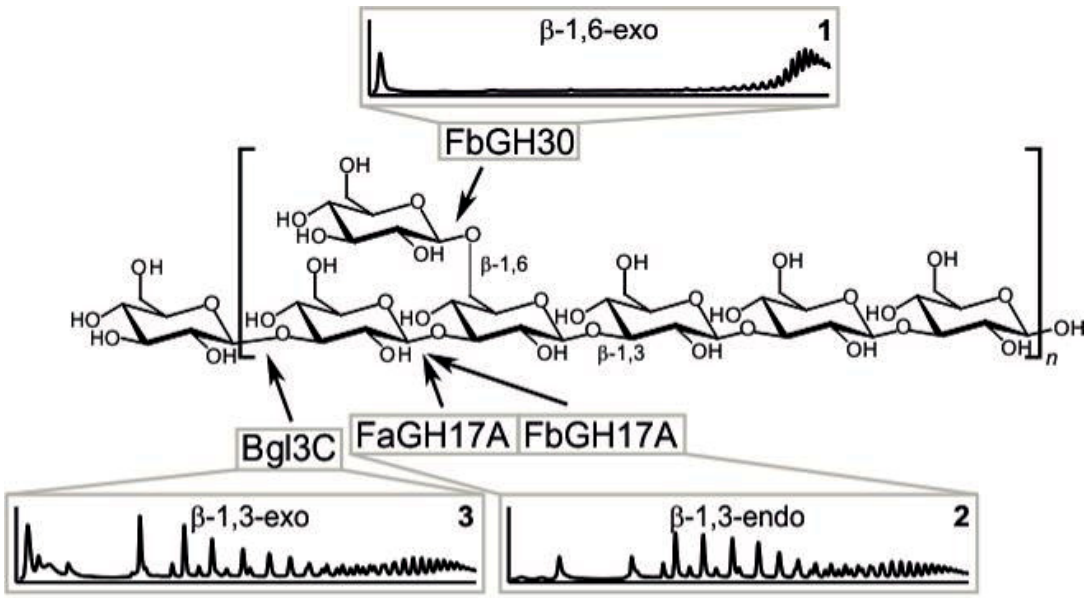


FIG 2.4.1 Enzymatic dissection of organic matter with specific enzymes enables quantification of environmental laminarin. Laminarin is a β-glucan composed of a β-1,3-glucose backbone with β-1,6 linked side chains. The four enzymes from different marine Bacteroidetes and terrestrial Gammaproteobacteria belong to the families GH3, GH17 and GH30. **(1)** The GH30 hydrolyzes the β-1,6-linked monomeric branches. **(2)** Both GH17 hydrolyze the β-1,3-linked backbone internally (endo) into smaller oligos. **(3)** The GH3 hydrolyzes the original backbone and smaller oligos externally (exo) into glucose monomers. *Chromatogram 2 and 3 are so far only dummies, since this experiment is still pending. However, it should look very similar.*

Microbes and their enzymes have evolved to overcome the diversity of glycans because they must adapt to accommodate the three-dimensional structure of the glycan for productive binding and catalysis (Labourel et al., 2014). To take advantage of the extraordinary enzymatic specificity offered by nature, we developed a biocatalytic assay based on enzymes that specifically recognize laminarin in marine organic matter (FIG 2.4.1) (Stefan Becker et al., 2017; CE Nelson et al., 2017). We chose to focus on laminarin because cultured algae (Størseth et al., 2005, 2006; McConville et al., 1986; Paulsen and Myklestad, 1978; Beattie et al., 1961) contain substantial amounts of this energy storage glycan (up to 80% per dry weight). Moreover, the most produced carbohydrate active enzymes by heterotrophic microbes in algal blooms are laminarinases hinting to the relevance of laminarin in the wild (Teeling et al., 2012). Laminarin is a branched polysaccharide of linear β-(1→3) linked glucose with an average degree of polymerization (DP) of ~20-30. In our

assay GH17 enzymes specifically cleave β-(1→3) linkages. Laminarin also contains β-(1→6) linked side chains consisting of one or more glucose monosaccharides (FIG 2.4.1), which are cleaved by the GH30 enzymes. The reaction products are glucose and oligosaccharides, which are easily measurable and proportional to the amount of laminarin in the sample (Stefan Becker et al., 2017).

We measured laminarin in organic matter obtained during six cruises and near the island Helgoland in the North Sea (Supplementary FIG 2.4.2). In short, particulate organic matter was concentrated by filtering seawater. Glycans were extracted with a buffer and digested with the enzyme mix. A reducing sugar assay and or high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) analysis revealed via the amount of degradation products the laminarin concentration by comparison to a non-digested (blank) control. The available data for each of these cruises is presented in Supplementary TAB 2.4.1; detailed analytical protocols are provided in the supporting text.

First, we determined to what extend laminarin correlates with algal growth. We monitored laminarin during two spring algal blooms the North Sea in 2016 and 2017 (FIG 2.4.2) where substantial nutrient input and strong mixing manifest in high algal productivity (Gattuso et al., 1998). We measured the highest laminarin concentrations of the entire dataset (~2.6 mg L^{-1}; mean 1.4±0.6 mg L^{-1}, 2016) and found that laminarin was present in smaller microalgae (between 10 and 3 µm) and not detectable below 3 µm. The laminarin increase during the bloom was driven by larger microalgae (log. regression: R^2=0.43; P<0.001, n=28) (Supplementary FIG 2.4.3A and FIG 2.4.4A). This was supported by multi-wavelength spectrofluorometric analysis, which resolved the contribution of diatoms, chlorophytes, cryptophytes and cyanobacteria (FIG 2.4.2A). This analysis revealed that laminarin increased with diatom growth (log. regression: R^2=0.46; P<0.001, n=28) (Supplementary FIG 2.4.4B) with a rate of ~0.06 mg d^{-1} L^{-1} in April 2017. In laboratory cultures of diatoms laminarin is released by algae due to active secretion during overflow metabolism and cell lysis accounting for up to 70% of the DOM (Myklestad, 2000). However, over the course of three months, weekly analyses of native DOM did not detect laminarin, likely because it was degraded too rapidly by bacteria. The most expressed carbohydrate active enzymes during algal blooms in the North Sea are laminarinases (Teeling et al., 2012) including the ones used in our assay (Stefan Becker et al., 2017). The high laminarinase activity was also found across the Atlantic (Arnosti et al., 2012) suggesting laminarin is widely enough produced to elicit basin-wide enzyme production by bacteria (Teeling et al., 2012; Arnosti, 2011).

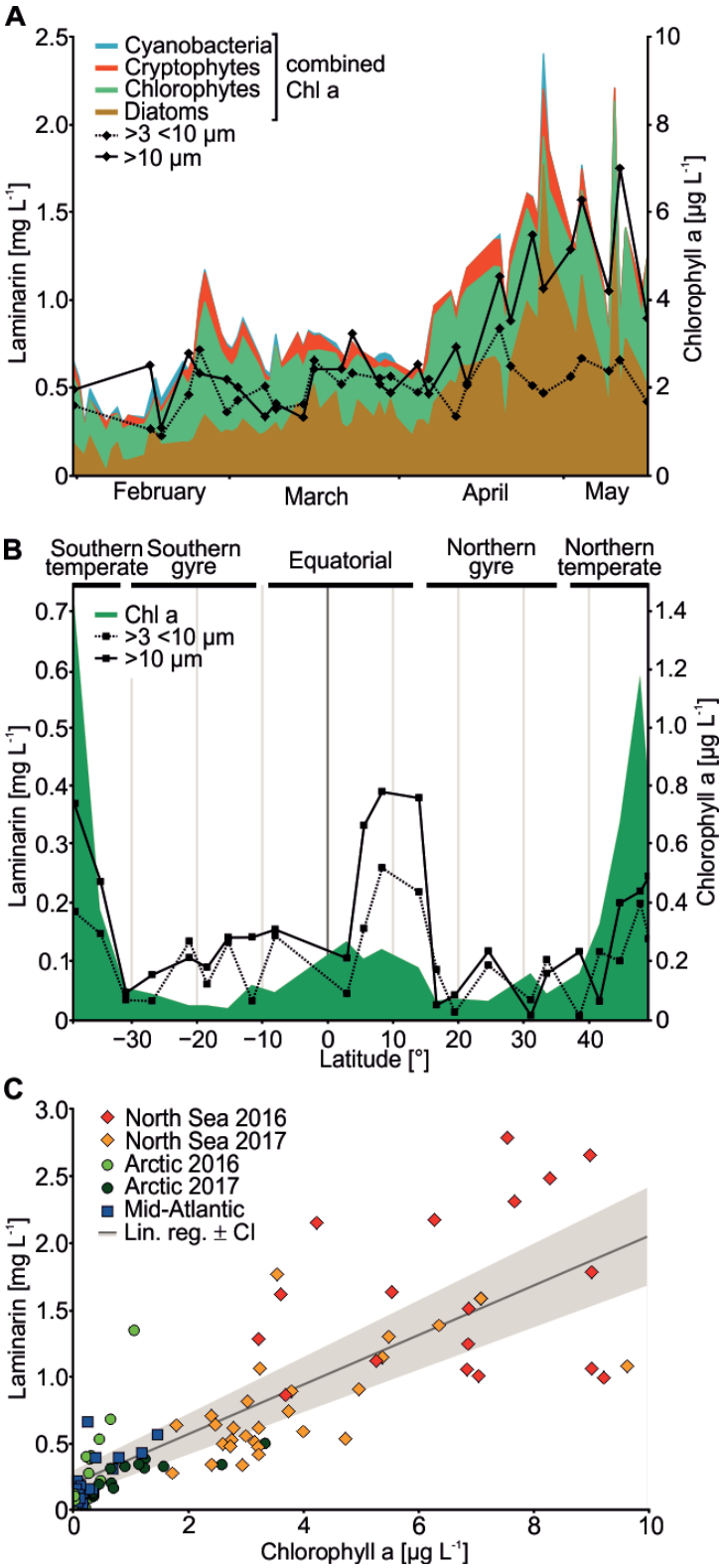


FIG 2.4.2 Comparison between laminarin and chlorophyll a in a spring bloom time series, across the Atlantic and in different oceanic regions. **(A)** Laminarin and chlorophyll a were determined during a phytoplankton spring bloom in the North Sea in 2017. Particulate organic matter was concentrated twice per week from 100 L seawater for over three month by sequential filtration through filters of decreasing size and laminarin was quantified by the biocatalytic assay. Multi-wavelength spectrofluorometric analysis of unfiltered seawater, resolved the contribution of four phytoplankton groups to the bloom. **(B)** Laminarin was measured along a meridional transect from the North to the South Atlantic. Particulate organic matter was concentrated from 20 L of seawater by sequential filtration through filters of decreasing size. **(C)** Comparison of laminarin and chlorophyll concentrations in all datasets where chlorophyll a was measured. Linear regression was applied to the laminarin to chlorophyll a relationship (R^2=0.66; P<0.001, n=101). The confidence intervall (CI) in grey was calculated at level 0.95.

To expand on the relationship between algal biomass and laminarin, using Chl-a as growth proxy (Field et al., 1998), we analyzed samples from the Arctic 2016, 2017 (Supplementary FIG 2.4.4C) as well as from a meridional Atlantic transect, including the northern temperate area, the northern gyre, the equatorial region, the southern gyre and the southern temperate area (FIG 2.4.2B). Chl-a ranged between 0.05 and 1.4 µg L^{-1}, its concentration was highest in temperate and lowest in the two oligotrophic gyres. We observed elevated laminarin values (~0.37 and ~0.21 mg L^{-1} on 10 and 3 µm size filters) in the northern equatorial upwelling region, reflecting higher productivity due to higher nutrients (Poulton et al., 2006; Robinson et al., 2006; Pérez et al., 2005). As seen for the North Sea, across the Atlantic transects the 10 µm POM contained more laminarin than the 3 µm POM. Different to the Arctic and the North Sea we observed a higher variation for the Lam:Chl-a ratio (FIG 2.4.4D). This discrepancy was driven by elevated Lam:Chl-a ratios near and at the equator, which may reflect the previously observed adaptive response of microalgae to increased irradiance of reducing the content of Chl-a (Falkowski and Chen, 2003). We compared the Arctic, North Sea and the Mid-Atlantic datasets and found that the Lam:Chl-a ratio correlated exponentially with water temperature (R^2=0.49; P<0.001, n=82) (Supplementary FIG 2.4.5). This suggests that the Lam:Chl-a ratio is sensitive to temperature and light. Combining Chl-a and laminarin measurements from all regions revealed a linear Lam:Chl-a correlation of R^2=0.66; P<0.001, n=101 (FIG 2.4.2C).

Laminarin and POC correlate (linear regression, R^2=0.81; p<0.001, n=88) (FIG 2.4.3A) but their ratio differs between regions revealing variable bioenergy levels in the surface ocean. North Sea and Peru upwelling regions formed a group of greater POC and laminarin values (>0.8 mg POC L^{-1} and >0.5 mg Lam L^{-1}), compared to the other regions (<1.1 mg POC L^{-1} and <1.3 mg Lam L^{-1}). Intriguingly the North Sea 2017 and Peru upwelling with the highest concentrations of POC and laminarin were characterized by the lowest LamC:POC ratios of 21±6% (SD) and significantly deviated from the overall mean of 37±19% (SD) (P<0.001 and P<0.05, Kruskal-Wallis-Test; N = 88) (FIG 2.4.3B). On the other hand the Arctic 2017 and the Canary upwelling gave LamC:POC ratios of around 58±13% (SD) exceeding the mean significantly (P<0.001 and P<0.05, Kruskal-Wallis-Test).

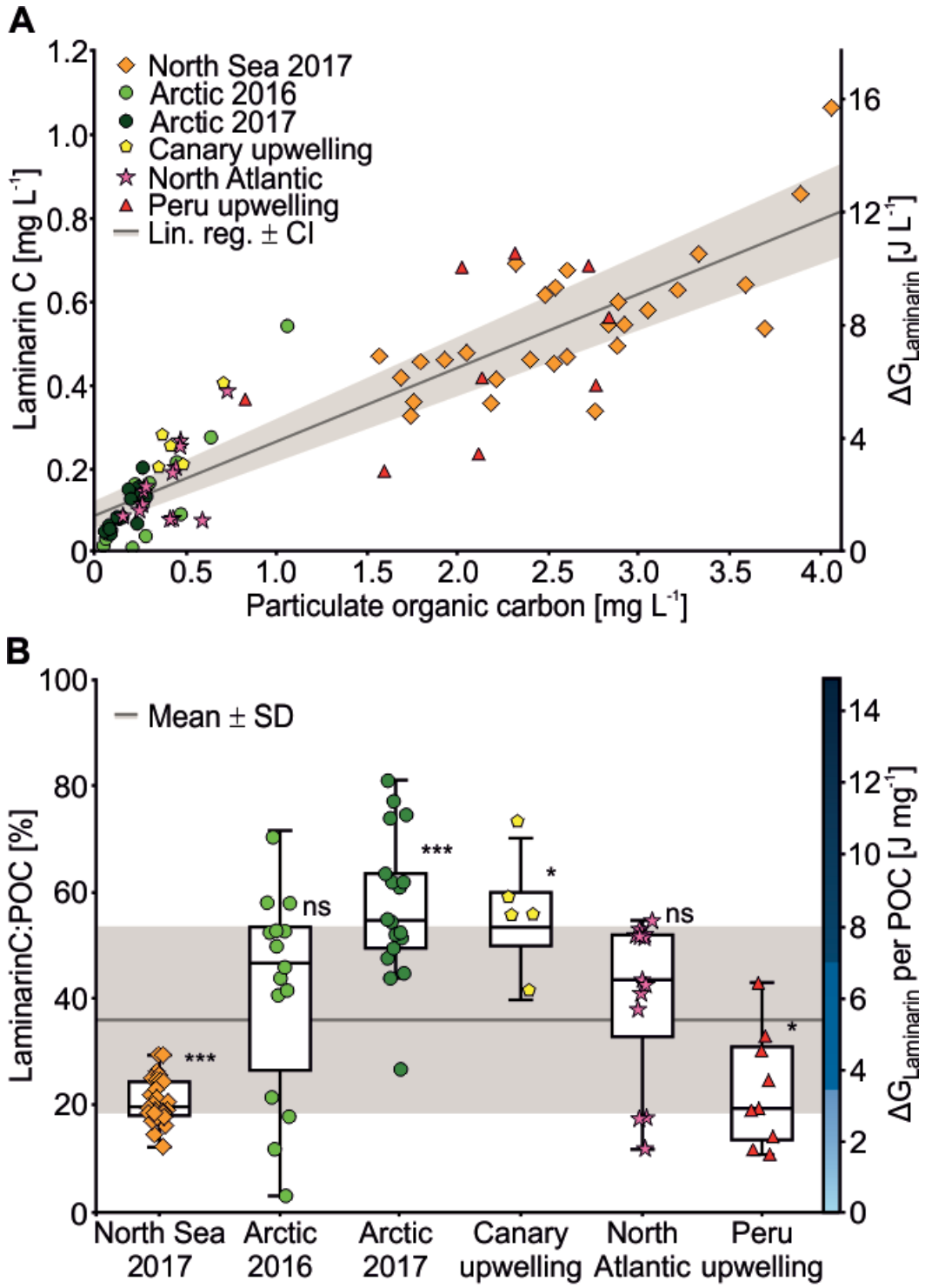


FIG 2.4.3 Correlation of laminarin to particulate organic carbon content and comparison of LaminarinC:POC ratios. Particulate organic matter was concentrated by filtering seawater. Laminarin was quantified by the biocatalytic assay and POC with an elemental analyzer using the same filters. Based on the max. theoretical ATP yield that would be gained through respiration of the glucose monomers with oxygen, the energy density of laminarin per volume was calculated. **(A)** The overview scatter plot comprises all regions where POC was measured. Linear regression was applied to the laminarin to POC relationship (R^2=0.81; P<0.001, n=88). The confidence interval (CI) in grey was calculated at level 0.95. **(B)** The boxplot depicts each individual dataset against the overall mean value of LaminarinC:POC. Significant deviations from the overall mean and its standard deviation (SD) in grey were tested (Kruskal-Wallis test; ****, P<0.0001; *, P<0.05; ns, not significant).

We further examined the quantitative relationships by establishing a mass balance of carbon in lipids, proteins, and laminarin in relation to POC. On a subset of 15 samples from the cruise to the Arctic in 2016, we determined protein and lipid masses and converted them into carbon equivalents (TAB 2.4.1). In addition to the 42±18% (SD) of carbon accounted for by laminarin, 44±14% (SD) consisted of protein carbon and 4±3% (SD) of lipid carbon. Together the laminarin, protein and lipid account for about 90% of the POC leaving room for additional polysaccharides (cell wall), RNA, DNA and minor molecules. Based on our measurements showing that laminarin carbon accounts on average for 37±19% (SD) of POC and the correlation with algal growth, we estimate that 18±9 gigatons of the 47.5 gigatons of the annual marine organic carbon production of phytoplankton is in the form of laminarin (Field et al., 1998).

TAB 2.4.1 Comparison of mass balances of lipid, protein, and laminarin or total carbohydrates in POC. Shown are mean values and standard deviations. In our study, the analysis of lipid and protein abundances was applied on the Arctic 2016 dataset.

	Total carbo-hydrates [%]	Laminarin [%]	Proteins [%]	Lipids [%]	Sum [%]
Parsons *et al.* 1961	23±11	-	39±13	8±5	70±29
Finkel *et al.* 2016	15±11	-	32±14	17±10	64±35
This study (Arctic 2016)	-	42±21	44±14	4±3	90±38

Presence of laminarin in particulate organic matter below the photic zone in the Arctic Ocean strongly suggests that it is part of sinking POM and potentially contributes to carbon export. In the Arctic we found the highest laminarin concentrations at the Chl-a maximum between 16 and 35 m (0.46±0.34 mg L^{-1} [SD]) (Supplementary FIG 2.4.6A), which confirms that active production by microalgae in the photic zone is the major source of laminarin (Alderkamp, van Rijssel, et al., 2007). We also measured laminarin in sinking particles in the North Atlantic (Porcupine Abyssal Plain) with a marine snow catcher and found that it was most abundant in the Chl-a maximum (~1.0 mg L^{-1} in the suspended fraction and ~0.5 mg L^{-1} in the settling particles) (Supplementary FIG 2.4.6D). The Lam:POC analysis (Supplementary FIG 2.4.5C) revealed in both Arctic years a ratio of 50±18% (SD), which did not significantly change with depth ($P>0.05$, Kruskal-Wallis-Test), suggesting that the laminarin status of the surface ocean might influence deeper waters. However, in the North Atlantic the Lam:POC ratio was greater at depth than at the surface supporting the observation that the ratio can be variable (Supplementary FIG 2.4.6E-F). The presence of laminarin in sinking particles (48±7% [SD] Lam:POC) indicates it contributes to carbon and energy export into the deeper ocean.

Laminarin can be enzymatically converted into glucose and further into the potential metabolic energy that would be gained through respiration with oxygen, yielding 32 ATP per molecule of glucose. The North Sea contained with a $\Delta G_{Laminarin}$ >15 J L^{-1} the highest potential laminarin-based metabolic energy (FIG 2.4.3A). However, normalized to POC the North Sea and

Peru belonged to the systems with relatively low laminarin energy levels, the North Atlantic was average and the Arctic 2017 and Canary upwelling had the greatest ratio (FIG 2.4.3B). This implies that the energy yield for grazers such as crustacean and fish would be different in each region. This is important because we show that laminarin can be easily digested and therefore rapidly provides energy to heterotrophic processes.

2.4.4 Acknowledgements

This research was supported by the Deutsche Forschungsgemeinschaft (grant HE 7217/1-1 to J.-H.H.) and by the Max Planck Society. We thank Bernhard Fuchs and Laura Bristow for providing samples from the cruises AMT22, DY077, POS508 and M138. We thank Antje Boetius for providing ship time for PS99.2 and PS107. We thank Rudolf Amann and Thomas Schweder for enabling our participation in the Helgoland spring bloom campaigns. Moreover, we thank Eva-Maria Nöthig and Wilken-Jon von Appen for providing us with chlorophyll/fluorescence data. Furthermore, we thank the crews of RV Polarstern, RV Meteor, RV Poseidon, RRS James Cook, RRS Discovery and the AWI Helgoland for helping with sample acquisition.

2.4.5 Supplementary information

2.4.5.1 Material and methods

2.4.5.1.1 Sites and sample collection

The entire dataset comprises filtered samples from 6 different oceanic regions (Supplementary FIG 2.4.1). They were taken during 8 separate cruises or campaigns (Supplementary TAB 2.4.1). The volume that went through every filter was logged. After sampling the filters were wrapped into pre-combusted aluminum foil and kept frozen at -20°C until further processing. Blank samples were taken by filtration of 10 L of MilliQ-H_2O.The North Sea was sampled during two sampling campaigns in spring 2016 (from March to June) and 2017 (from January to May) at the long term ecological research site (LTER) 'Helgoland Roads' at the island of Helgoland in the German Bight. The sea surface samples were taken in 50 L Nalgene bottle and the water was filtered sequentially through 10, 3 polycarbonate (PC) and 0.2 µm polyethersulfone filters. An additional 142 mm GF/F glass fiber filter sample at a pore size of 0.7 µm was taken during the 2017 campaign. Additionally, we concentrated 100 L of high molecular weight dissolved organic matter (HMW-DOM) using a 1 kDa membrane and tangential ultrafiltration (Sartorius). The Arctic was sampled during the two cruises PS99.2 in 2016 (from June to July) and PS107 in 2017 (from July to August) at 8 different stations of the LTER 'HAUSGARTEN' and by using GF/D filters at a pore size of ~3 µm. Different water depths were sampled *in situ* by deploying several Large Volume Water Transfer Systems (WTS-LV, McLane) at the same time. The pumps were attached to the cable of the rosette water sampler and filtered water over the duration of 60 min. The Mid-Atlantic was sampled during the Atlantic Meridional Transect 22 (AMT) cruise in 2012 (from October to November) at 23 stations. The rosette water sampler was

used to get sea surface water, which was sequentially filtered through 10 and 3 µm polycarbonate filters. The Canary upwelling system was sampled during the POS508 cruise in 2017 (from January to February) at 5 stations. Two sea surface samples were taken by the rosette water sampler and three were taken by using the ships inlet and all were filtered through GF/D filters. The Peru upwelling system was sampled during the M138 cruise in 2017 (from June to July) at 9 stations. The rosette water sampler and GF/D filters were used to sample sea surface water. The North Atlantic was sampled during the DY077 cruise in April 2017 at 3 stations in the area of the Porcupine Abyssal Plains. Sea surface water was sampled by using the ships inlet. Additionally a marine snow catcher at different depths was used to fractionate suspended and settling particle samples. Settling particles in the sampled water body were sedimented by letting the snow catcher stand upright for 9 to 31 h. Afterwards the suspended particles in the top section were slowly drained and later the remaining volume of the bottom section including the settling particles was recovered. Both fractions were filtered through GF/D filters.

2.4.5.1.2 HMW-DOM sample processing

The concentration of the 100 L sample resulted in a 0.5 L volume, which was frozen for later processing. The sample was then thawed and further concentrated to a volume of ~25 mL. Afterwards it was desalted using a 1 kDa membrane (Spectra/Por). The material was then frozen and freeze dried (Labogene). The procedure resulted in several milligrams of a crystalline powder, which could be again dissolved for further analyses.

2.4.5.1.3 Laminarin quantification

The sample preparation steps including extraction, enzymatic hydrolysis and the final measurement of hydrolysis products via an adapted protocol of the PAHBAH reducing sugar assay were performed according to Becker and Hehemann 2018 (unpublished Manuscript III).

2.4.5.1.4 Particulate organic carbon measurement

In case that a glass fiber filter was used for sampling, the same filter was used for both the laminarin and particulate organic carbon measurements. Prior to sampling, the filters were combusted for 4 h at 450°C to remove possible carbon contaminations. After punching out defined pieces of the filter in triplicates, the pieces were treated with concentrated HCl for 24 h in an excicator to remove inorganic carbon. The pieces were then dried for 24 h at 60°C. The dried filter pieces were packed into tin foil and later analyzed using an elemental analyzer (vario MICRO cube, Elementar Analysensysteme) and sulfanilamide for calibration.

2.4.5.1.5 Protein and lipid carbon quantification

Total protein and total lipid extractions were conducted on the same filters that were used for laminarin and POC measurements. After punching out defined pieces of the filter in triplicates, the pieces itself were used in the extraction according to an adapted protocol of Slocombe *et al.* (Slocombe et al., 2013). After adding 6% trichloroacetic acid to the filters and vigorous vortexing,

the samples were incubated at 95°C for 15 min. Precipitated protein and the filter material was centrifuged at 15.000 g for 20 min at 4°C and the supernatant was discarded. 1 ml of Lowry reagent was added to the pellet and the filter, followed by vigorous mixing. The rest of the protocol was performed according to the manufacturer's manual of the Total Protein Kit (Sigma-Aldrich), which is based on Peterson's modification of the original Lowry assay (Peterson, 1977; Lowry et al., 1951). Total lipids were extracted by acid hydrolysis with 1 M HCl in methanol as described by Becker *et al.* (Kevin W Becker et al., 2016). Filters were cut into an extraction vial containing the acid mix and combusted sand. The mixture was two times sonicated for 20 minutes and the hydrolysis was performed overnight (approx. 16 h) at 70°C. In the next morning, the mixture was sonicated two more times for 20 min and poured in a separatory funnel. The filters were then sonicated three times 20 min with 5:1 (v:v) dichloromethane:methanol and extracts were assembled in the separation funnel. Deionized water was added and the organic phase was drawn off after phase separation. The aqueous phase was then washed two times with dichloromethane and the organic phase three times with deionized water. The extract was dried under a stream of N2 and transferred to a pre-weighed 2 mL vial. The total lipid weight was determined by five successive measurements on a micro-scale balance (Mettler Toledo). The mean relative error was determined as 3.8%. A blank filter was extracted following the same procedure to check for potential non-lipid contamination. Total protein and lipid concentrations were converted to carbon equivalents assuming conversion factors of 0.50 and 0.75, respectively (Bourgeois et al., 2016; Satyanarayana and Chakrapani, 2013; Fabiano and Danovaro, 1994; Fichez, 1991; Rouwenhorst et al., 1991).

2.4.5.1.6 Supplementary data

The Helgoland Roads LTER time series provides chlorophyll a and temperature data on a weekday basis (Wiltshire et al., 2008). During both Arctic cruises, fluorescence measurements were conducted on the rosette water sampler that was deployed together with the *in situ* pumps (Tippenhauer et al., 2017) (Von Appen et al., unpublished). The measurements were calibrated with HPLC-based chlorophyll data (Nöthig et al., unpublished). All this data is accessible via the open database Pangaea (https://www.pangaea.de). The study uses also chlorophyll a and temperature data from Gavin Tilstone/Plymouth Marine Laboratory/Oceans 2025 project S01 Atlantic Meridional Transect, provided by the British Oceanographic Data Centre and funded by the Natural Environment Research Council (https://www.bodc.ac.uk).

2.4.5.1.7 Statistical analysis

Statistical analysis was carried out by using the R environment v.3.4.3 (R Core Team 2017, http://www.r-project.org). Linear and logarithmic models were fitted to the investigated variables. The relative quality of the models was estimated by the classical Akaike's information criterion (AIC, k=2) and the Bayesian information criterion (BIC). Significant differences in the

non-normal distributions of values among groups was tested by the Kruskal-Wallis-Test. The confidence interval for the mean of non-normal data was calculated via bootstrapping with 1000 replications (FK Wang, 2001). Multiple regression analysis with dummy variables was performed to investigate differences between varying groups of samples. p-values below 0.05 were considered statistically significant.

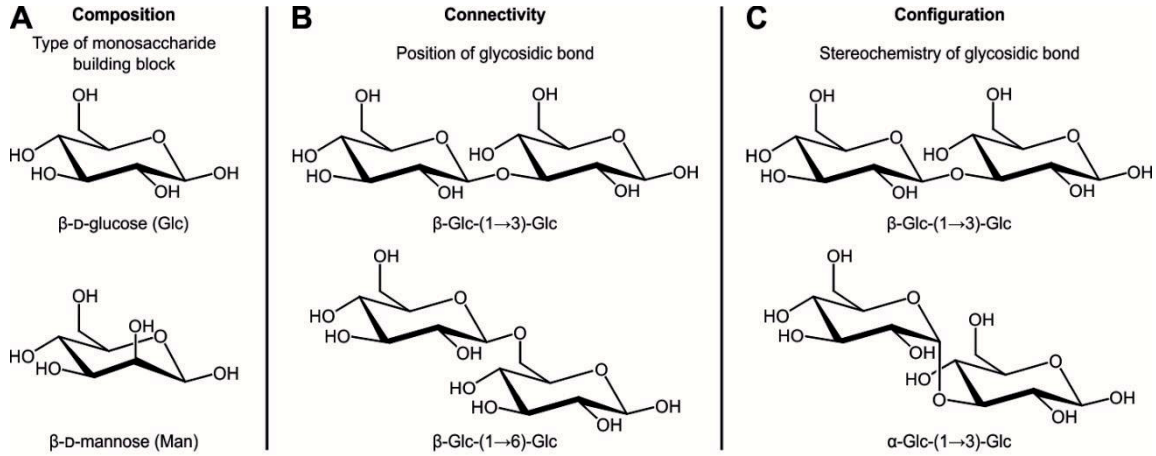
2.4.5.2 Supplementary results

In the North Sea 2017 dataset, we also tested 0.7 µm glass fiber filters using a single filtration step, which reproduced the laminarin signal of the 10 and 3 µm polycarbonate filters combined (Supplementary FIG 2.4.3B). This result was also illustrated by the slopes of the linear correlations (Supplementary FIG 2.4.3A). The laminarin content increases with a very similar slope on the glass fiber filter (GF/F), the 10 µm, and the combined 10 and 3 µm fraction on polycarbonate filter (PC), whereas the 3 µm fraction did not increase and barely correlates with chlorophyll a. This indicates not only that the laminarin increase during the bloom was driven by larger microalgae but also that different filters can be used for quantitative laminarin analysis. No laminarin was detected in the 0.2 µm fraction, which contains bacteria, and it was not present in particles or algae smaller than 3 µm suggesting laminarin is produced by larger microalgae.

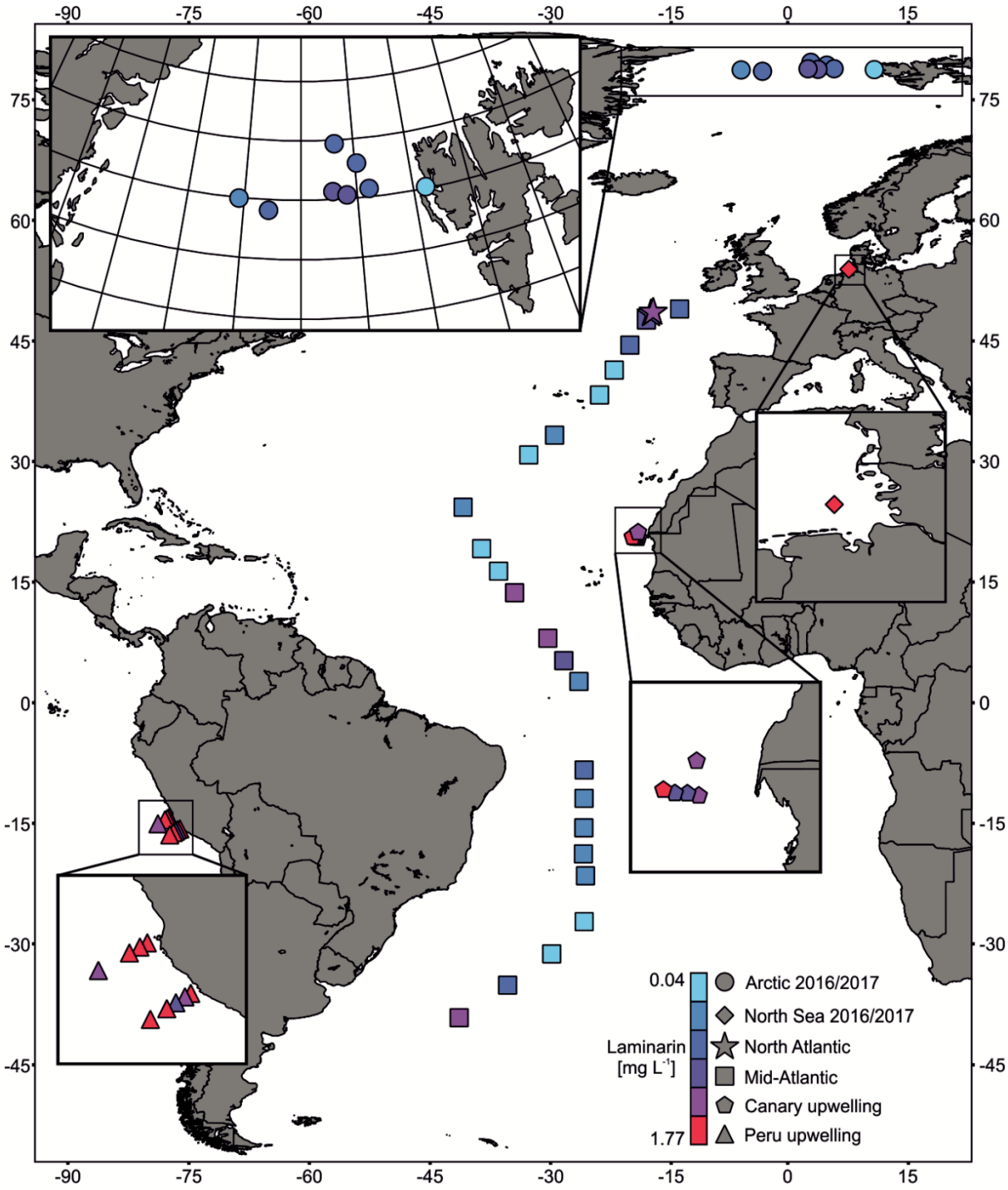
The sequentially filtrated 100 L were further concentrated on a 1 kDa membrane. In the resulting HMW-DOM of the three samples that were chosen from the bloom period (April 18, April 27 and May 4, 2017), we were not able to detect laminarin, meaning the hydrolyzed laminarin samples yielded the same low signal as the non-hydrolyzed blanks.

Both Arctic datasets exhibit a similar pattern of laminarin and POC distribution, with higher values at the depth of the chlorophyll a maximum between 16 and 35 m and lower values at the surface and the deeper water layer around 300 m (Supplementary FIG 2.4.6A-B). The cruise in 2016 took place earlier in the year (from June to July) and yielded on average higher values at the chlorophyll a maximum and lower values in the deep. Compared to this, the dataset from July to August 2017 shows a more even distribution over depth with lower values in the chlorophyll a maximum, but higher values at the surface and in the deep. This might indicate that at this later time in the year, the bloom was already in a more progressed stadium compared to 2016.

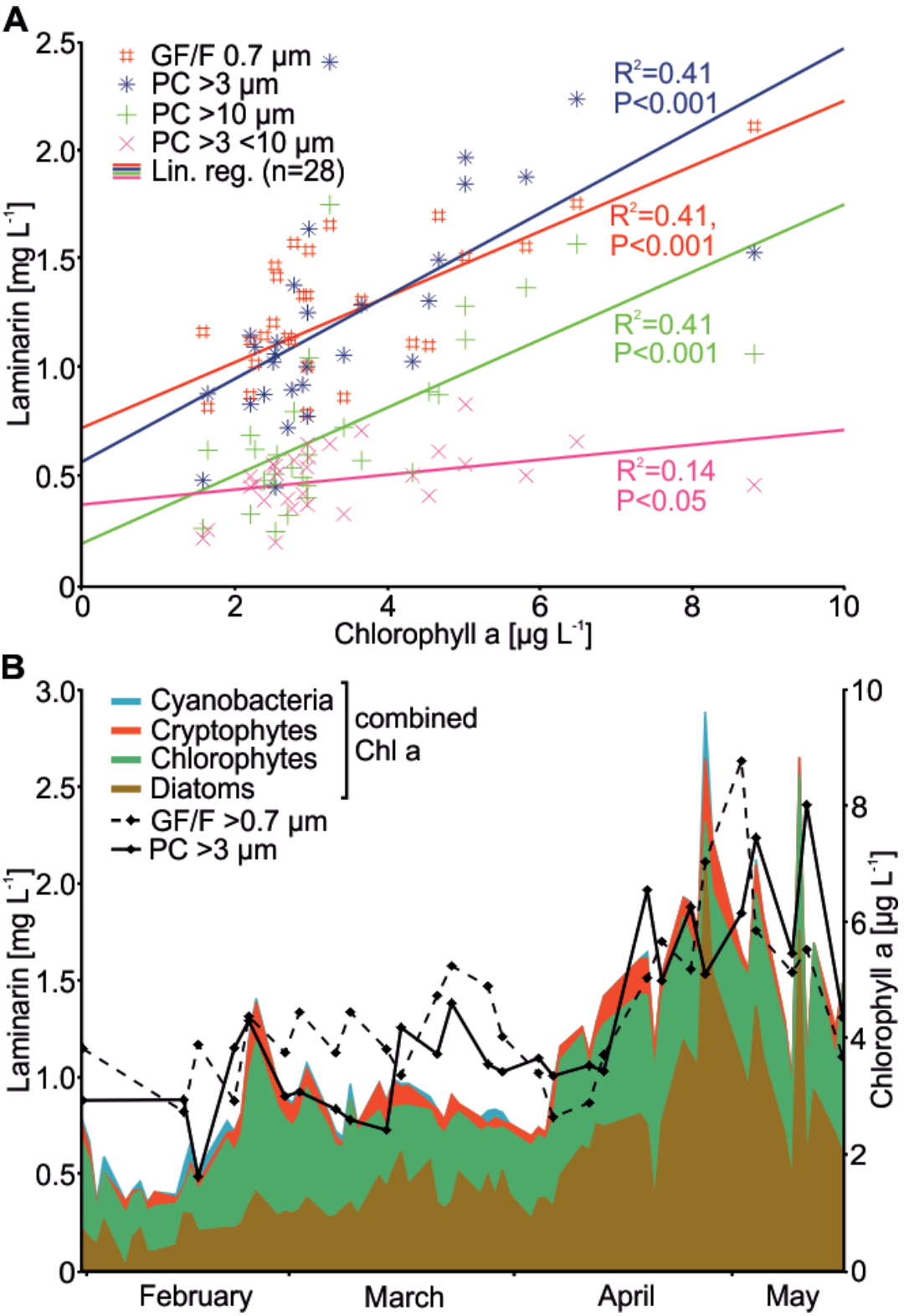
2.4.5.3 Supplementary figures



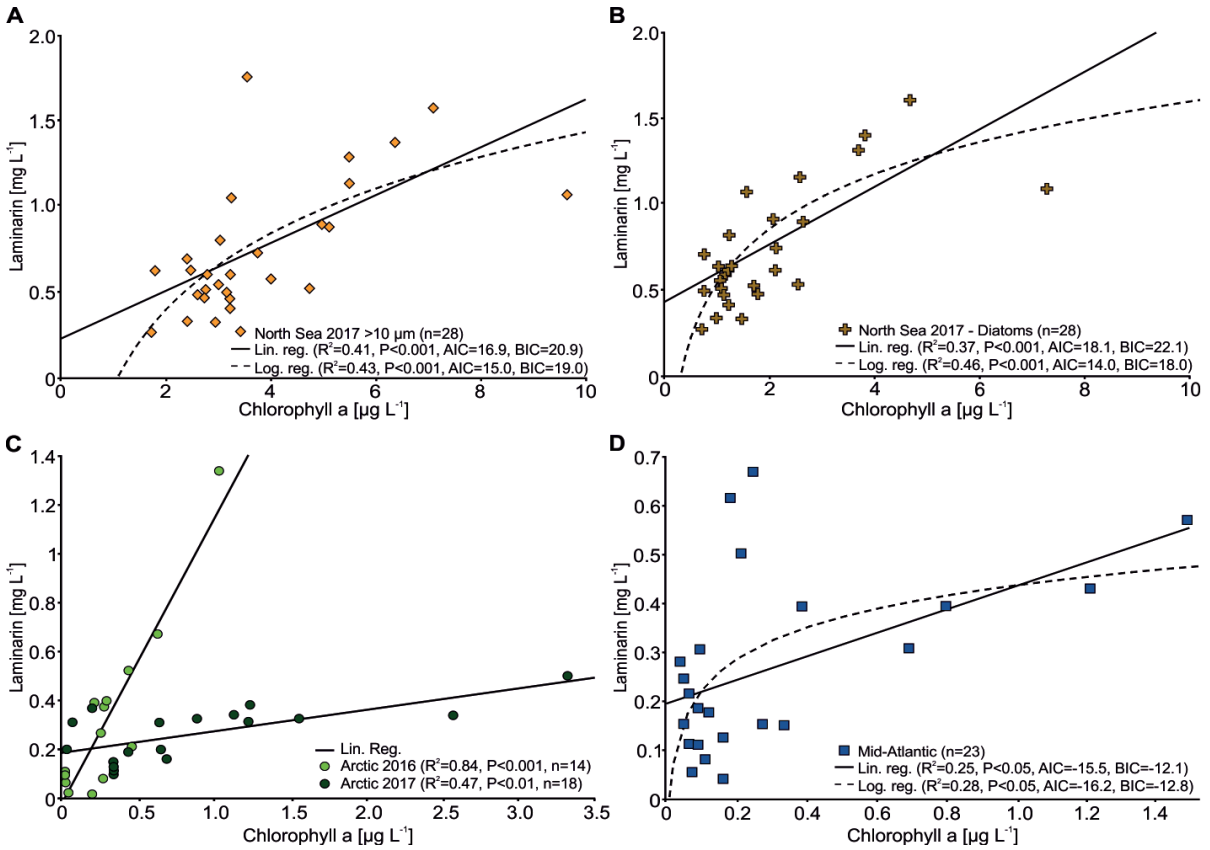
Supplementary FIG 2.4.1 Structural features of polysaccharides adapted from Hofmann et al. 2015. Carbohydrates build the most diverse and complex macromolecules in nature. **(A)** The composition of a glycan is the abundance of its monomers, which can differ in the stereochemical orientation of one of their carbon atoms as seen for mannose (Man) and glucose (Glc). **(B)** The connectivity of a glycan describes the fact that each monomer has multiple hydroxyl groups each of which can serve as the anchor point to connect via the glycosidic bond to another monomer. This enables glycans to be branched whereas proteins and DNA are linear. **(C)** The configuration of the glycoside bonds of a glycan can be in two ways namely α or β. Combining these bespoke structural possibilities results in high molecular diversity, but to identify and quantify a glycan in situ all three layers of information are required. For example by composition alone the glycans cellulose (glucose β1→4 linked), glycogen (α1→4, β1→6) or laminarin (β1→3, β1→6) cannot be distinguished because they are all made of glucose.



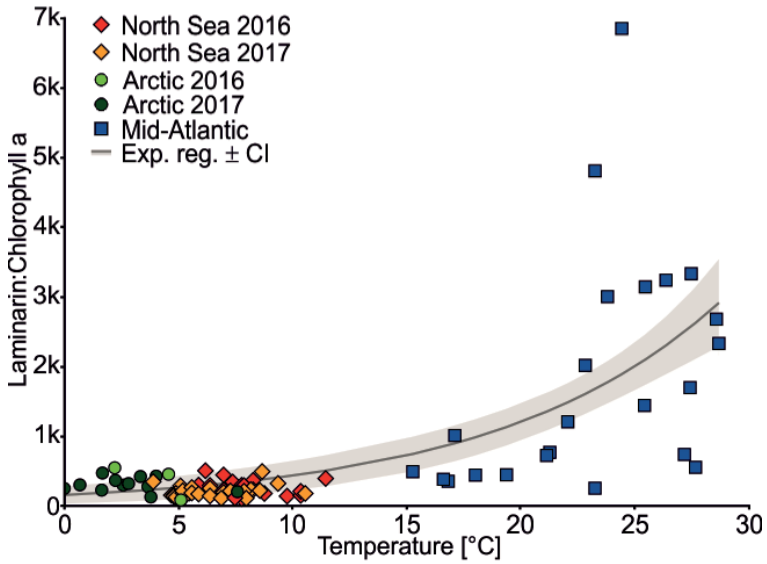
Supplementary FIG 2.4.2 Station and laminarin surface concentration overview. The entire dataset comprised samples from 8 different cruises and campaigns indicated by different symbols. The colors represent mean laminarin concentrations from surface samples (max. 50 m water depth). The enlarged Arctic region is using a polar stereographic projection of the world, whereas the rest of the map and all other enlarged areas are using an equirectangular projection. The black lines in the grey landmasses mark country borders. The map was created using QGIS (v.2.18.14) and the Natural Earth free vector and raster map data.



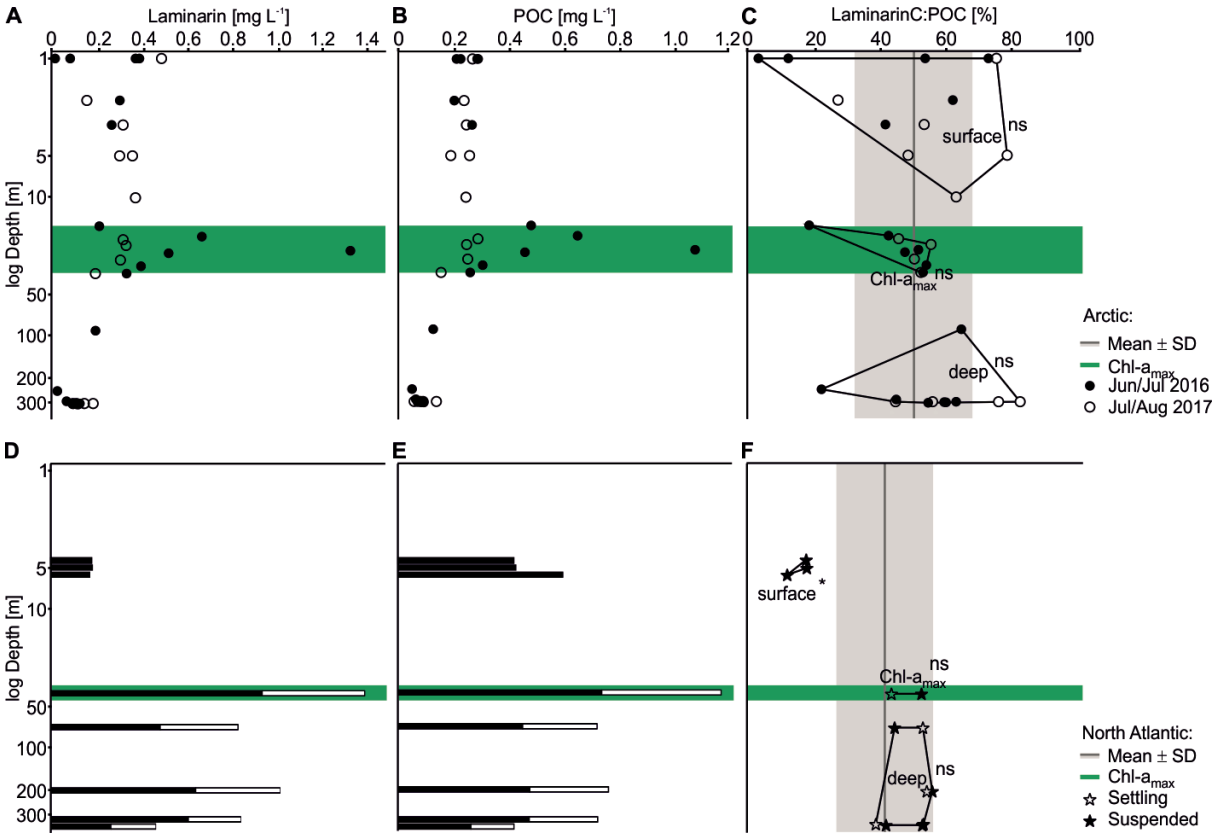
Supplementary FIG 2.4.3 Laminarin and chlorophyll a comparison on different filter fractions during a phytoplankton spring bloom in the North Sea 2017. Particulate organic matter was concentrated twice per week from 100 L seawater for over three month by sequential filtration through filters of decreasing size and laminarin was quantified by the biocatalytic assay. Multi-wavelength spectrofluorometric analysis of unfiltered seawater, resolved the contribution of four phytoplankton groups to the bloom. **(A)** Linear regression was applied to the laminarin and chlorophyll a signal relationship of all filter fractions. **(B)** The combined signal from sequentially used polycarbonate filters (PC) with pore sizes of 10 and 3 µm was compared to the signal of the glass fiber filters (GF/F) with a pore size of 0.7 µm, which was used without prefiltration.



Supplementary FIG 2.4.4 Correlations between different laminarin and chlorophyll a signals. Particulate organic matter was concentrated by filtering seawater. Laminarin was quantified by the biocatalytic assay and chlorophyll a was determined either by spectrofluorometric or chromatographic analyses. **(A)** Linear and logarithmic regression was applied to the chlorophyll a and laminarin (>10 µm fraction) signal relationship during the North Sea 2017 phytoplankton bloom. **(B)** Linear and logarithmic regression was applied to the relationship between the laminarin (>10 µm fraction) and chlorophyll a signal from diatoms and during the North Sea 2017 phytoplankton bloom. **(C)** Linear regression was applied to the laminarin and chlorophyll a signal relationship in the Arctic 2016 and 2017 dataset. **(D)** Linear and logarithmic regression was applied to the chlorophyll a and total laminarin (>3 µm) signal relationship during the Mid-Atlantic meridional transect.



Supplementary FIG 2.4.5 Correlation between Lam:Chl-a ratio and water temperature. Particulate organic matter was concentrated by filtering seawater. Laminarin was quantified by the biocatalytic assay and chlorophyll a was determined either by spectrofluorometric or chromatographic analyses. Exponential regression was applied to the relationship (R^2=0. 49; P<0.001, n=82). The confidence interval in grey (CI) was calculated at level 0.95.



Supplementary FIG 2.4.6 Vertical profiles of laminarin, POC and laminarin:POC ratios in the Arctic and Northern Atlantic datasets. During both Arctic cruises we measured vertical profiles of laminarin and POC via in situ pumping of up to 560 L at different stations and depths to 300 m. In the North Atlantic sampling, a marine snow catcher was used additionally to fractionate suspended and settling particle samples. Settling particles in the sampled water body were allowed to sediment and both fraction were drained and filtered separately. The laminarin in the concentrated particulate organic matter was quantified by the biocatalytic assay and POC with an elemental analyzer using the same filters. **(A)** Laminarin over depth during both Arctic samplings. **(B)** POC over depth during both Arctic samplings. **(C)** LaminarinC:POC ratio over depth during both Arctic samplings. Significant deviations from the overall mean LaminarinC:POC ratio of 50±18% (SD), were tested (Kruskal-Wallis-Test; ns, not significant). **(D)** Laminarin over depth during the North Atlantic sampling. **(E)** POC over depth during the North Atlantic sampling. **(F)** LaminarinC:POC ratio over depth during the North Atlantic sampling. Significant deviations from the overall mean LaminarinC:POC ratio of 41±14% (SD), were tested (Kruskal-Wallis-Test; *, P<0.05; ns, not significant).

2.4.5.4 Supplementary tables

Supplementary TAB 2.4.1 Sample overview of the entire dataset. Sums are shown in the last row.

Dataset	# Depths	# Days of time series	# Stations	# Laminarin samples			# POC samples	# Chlorophyll a samples		# suspended/settling particle samples	# Protein/lipid samples
				PC 10/3µm	GF/D 3µm	GF/F 0.7µm		Total	Group fractions		
Arctic 2016	3	*NA*	5	*NA*	15	0	15	14	*NA*	*NA*	15
Arctic 2017	3	*NA*	7	*NA*	18	0	18	18	*NA*	*NA*	*NA*
North Sea 2016	1	93	1	18	18	0	*NA*	61	*NA*	*NA*	*NA*
North Sea 2017	1	106	1	28	0	28	28	72	72	*NA*	*NA*
North Atlantic	4	*NA*	3	*NA*	13	0	13	*NA*	*NA*	10	*NA*
Mid-Atlantic	1	*NA*	23	23	*NA*	*NA*	*NA*	23	*NA*	*NA*	*NA*
Canary upwelling	1	*NA*	5	*NA*	5	0	5	*NA*	*NA*	*NA*	*NA*
Peru upwelling	1	*NA*	9	*NA*	9	0	9	*NA*	*NA*	*NA*	*NA*
			54	69	78	28	88	188	72	10	15

3 Conclusions, general discussion and outlook

3.1 Further method development and possible new applications

The aforementioned overall goal of this project was the development and application of a novel enzymatic method for the specific quantification of the algal storage glucan laminarin in order to shed more light on the global distribution of this polysaccharide and its importance in the marine carbon cycle.

The first part of this goal was achieved by establishing a new enzyme-based method for specific laminarin quantification. The method proved to be specific and fast, robust and easily applicable. We demonstrated its utility on various marine samples, including the quantification of macroalgal laminarin (Scheschonk et al., in preparation). In the spirit of the open science movement and in order to promote its further application, we made the access to the method as easy as possible for other researchers by contributing to the Bio-Protocol platform and the Addgene plasmid repository. Moreover, the method was further improved after its initial publication. The two enzymes FbGH30 and FaGH17A hydrolyzed about 50% of the glycosidic linkages compared to the yield that was obtained by the unspecific acid hydrolysis (FIG 3.1). By adding a second endo-acting β-1,3-glucanase of the GH family 17, the hydrolysis yield was increased to about 80%. By increasing the hydrolysis yield, the sensitivity of the method was increased as well. The third enzyme was the FbGH17A from *Formosa* B, of which the biochemistry and substrate specificity was invested as part of the second publication. The laminarin PULs of *Formosa* A and B and other organisms often contain enzymes from GH family 2, 3 or 5 in addition to the GH families 16, 17 and 30. These families are known to exhibit exo-β-1,3-glucanase activity (Harvey et al., 2000; Sinnott and Souchard, 1973; Wallenfels, 1951). Our assay is still missing an activity like this and we assume that an activity like this would improve the hydrolysis yield even further. Nonetheless, it has to be stated that we do not need total hydrolysis in order to quantify laminarin, since we hydrolyzed commercial laminarin at different concentrations with the enzymes and produced a linear calibration curve to account for incomplete hydrolysis. Our own search for such a complementary enzyme in the genomes of marine organisms and metagenomes remained unsuccessful, due to solubility or activity issues after heterologous expression.

However, we found the exo-β-1,3-glucanase activity in a recent study on several terrestrial GH3 (CE Nelson et al., 2017). Bgl3C from *Cellvibrio japonicus*, a plant polysaccharide degrading soil bacterium, exhibits a laminarin specific exo-β-1,3-glucanase activity that might be exploited in our assay for the further degradation of laminarin oligos into glucose monomers.

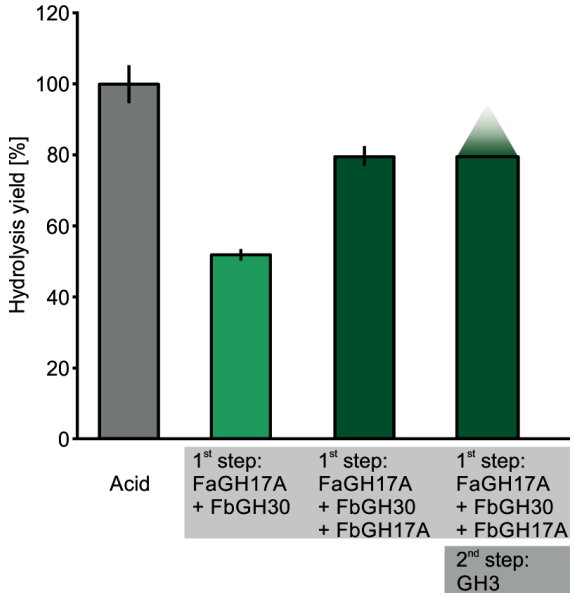


FIG 3.1 Comparison of laminarin hydrolysis yields using the old enzyme cocktail (FaGH17A+FbGH30), the new cocktail (FaGH17A+FbGH30+FbGH17A), the new cocktail including the GH3 from *Cellvibrio japonicus* (preliminary data) and total acid hydrolysis.

A total enzymatic hydrolysis of laminarin was not achieved and most probably, this is not possible in a batch reaction like the one that we are using. The reason for that lies in the catalytic mechanism of the enzymes (FIG 1.5). All GH families that are involved in the degradation of laminarin use the classical Koshland retaining mechanism, where a glycosyl enzyme intermediate is being formed in between the two-step reaction (Koshland, 1953). The possibility of transglycosylation reactions increases in a batch reaction over time. The equilibrium of the laminarin hydrolysis reaction is continuously being shifted. Smaller oligos and glucose accumulate in the reaction mix, since they are not being removed. The chances are increasing that an oligo or glucose monomer is replacing the water molecule during the second step of the reaction, yielding a transglycosylation. As a result, this reaction creates again longer oligos and by doing so it counteracts the total hydrolysis. Some retaining GHs are more prone to this transglycosylation activity than others (Bissaro et al., 2015; Vocadlo and Withers, 2000). The transglycosylation activity of one of our GH17 enzymes is in fact limiting the maximum hydrolysis yield that we can reach in a one-step reaction. Several observations pointed towards this conclusion. The structure of FbGH17A from *Formosa* B, which was included into the enzyme mix later, was solved by molecular replacement with its closest structurally characterized GH17 homolog from *R. miehei* (Qin et al., 2015). This homolog also exhibits a transglycosylation activity.

Furthermore, preliminary results showed that the new GH3 from *Cellvibrio japonicus* does not increase the hydrolysis yield of the enzyme mix in a one-step reaction. However, we used the common FaGH17A, FbGH30 and FbGH17A reaction mix and stopped the reaction via boiling and inactivation of the enzymes. If the new GH3 is afterwards added to the cooled sample, the hydrolysis yield in fact increases, suggesting that its effect is being masked by the transglycosylation reaction of one of the β-1,3-active GH17 in the mix. After inactivating the enzymes in the first mix, the GH3 can act alone and the result is not equalized anymore. This result suggests that the sensitivity of the assay might be improved even further, if the protocol is extended by a second hydrolysis reaction using this enzyme (FIG 3.1).

The observed transglycosylation activity led also to another idea, which will be investigated in the future. Currently, the function of CAZymes that exhibit both glycoside hydrolase and transglycosylation activity is mostly unknown (Bissaro et al., 2015). In case of the laminarinases in our assay it remains to be tested whether transglycosylation is happening in the environment or if this is an artifact of the unbalanced equilibrium that is being created in the batch reaction. However, in the marine environment, transglycosylation might be one source for the molecular diversity of recalcitrant dissolved organic carbon. The microbial carbon pump sequesters diverse molecules in the deep sea, which are recalcitrant against rapid degradation and persist in the ocean for thousands of years (Osterholz et al., 2015; Jiao and Zheng, 2011; Middelburg and Meysman, 2007) (FIG 1.1). Transglycosylating 'side reactions' might not only increase the degree of polymerization of certain polysaccharides but it could also add completely new functional groups to the reducing end of the glycan. Experimentally, this could be tested by adding simple molecules with hydroxyl groups, e.g. methanol or glycerol, to the reaction mix. Afterwards the products of these side reactions would be analyzed by chromatography and mass spectrometry. The laminarinases from our assay can be used as a model system in this experimental setup, but it can also be extended by laminarinases that were already described to be prone to transglycosylation activity (Damao Wang et al., 2016).

A different kind of analysis is possible due to the fact that laminarin differs in the degree of branching and polymerization depending on the organism which produces the laminarin. Preliminary results showed that the commercially available laminarin from the two different species *Laminaria digitata* and *Eisenia bicyclis* differs in its degree of branching (data not shown). Nonetheless, it has been stated that this difference does not influence the calibration in any way, since both calibration curves look the same. The way the enzymes act together on the different linkages of the molecule, opens up the possibility to analyze the DB in a relative easy manner. Only FbGH30 is able to hydrolyze the β-1,6-linkages of the monomeric branches. If the yield of a FBGH30 digestion is being compared to the yield of the enzymatic cocktail comprising all enzymes, a relative DB can be estimated. An analysis like this was not applicable on environmental samples with only very low concentrations of laminarin and a considerably reduced sensitivity of

the assay by using only one enzyme. Therefore, we want to apply this method to a collection of densely cultivated and pure microalgae strains.

3.2 Specialized laminarin degraders and prospective proteinbiochemical analyses

The biochemical and structural analysis of laminarinases, which eventually resulted in the second manuscript, was intended to complement the entire project. Several important findings were made in this study. The two *Formosa* strains A and B exhibit recurrently high abundances during diatom blooms in the North Sea. Furthermore, an expanded set of laminarin specific glycoside hydrolases and transporters belong to the most abundant proteins in the environmental samples and during growth experiments. Compared to *Formosa agariphila*, which is considered to follow a generalist heterotrophic strategy (Mann et al., 2013), the specialized laminarin degrader *Formosa* B has a smaller genome, less PULs and GHs, but at the same time it expresses more laminarinases in more laminarin utilization loci (LUL)(TAB 3.1).

TAB 3.1 Comparison of genome sizes and genomic content of the generalist *Formosa agariphila* and the specialized laminarin degrader *Formosa* strain B.

	Formosa agariphila	***Formosa* B**
Genome size	4.48 Mbp	2.74 Mbp
LULs/PULs	1/13	3/6
Putative laminarinases/all GHs	5/84	9/21

Besides, FLA-laminarin was imported into the periplasm of the cells by a selfish uptake mechanism. Finally, the co-induction of proteins related to nitrogen metabolism suggested the importance of a balanced carbon and nitrogen diet in these extraordinarily specialized laminarin degrading bacteria. Our contribution to this was the biochemical description of three glycoside hydrolases from PUL 2 of *Formosa* B via Michaelis-Menten-Kinetics and HPLC-based product analysis. Our finding of a multi-modular protein combining and most probably boosting the activity of a laminarinase with a transmembrane, oligosaccharide transporter (FIG 2.2.4E) has not yet been described for other bacteria. Its conservation in nature suggests a beneficial mechanism for laminarin utilization. Hence, it represents another subject worth further investigation in future studies, for example by structural analysis of the multi-modular transporter-enzyme-complex. Aside from that, the structure, which was retrieved from the crystallization of FbGH17A, needs to be further elucidated by co-crystallization using laminarin oligos as substrates and a mutated, inactive enzyme. So far, the molecular basis of the substrate specificity is deduced only from the structural modeling of the substrate. In addition, the oligomerization into trimers needs to be

confirmed by size exclusion chromatography. Finally, the localization of the participating enzymes needs to be improved experimentally, since the working model of localizations of the laminarin utilizing proteins in *Formosa* B is mainly based on *in silico* predictions in combination with indicative subproteome analyses (FIG 2.2.6).

3.3 Implications of the sweet ocean for future studies

The last part of the project can be described as the attempt to establish a laminarin inventory of the ocean by applying the newly developed enzymatic method to environmental samples primarily from the Atlantic Ocean and adjacent regions. The most essential result of this study was that laminarin accounts on average for 37±19% of the carbon in algae-derived organic matter, and that different oceanic regions vary in their bioenergetic state indicated by the abundance of this molecule. Due to the ubiquitous abundance of marine phytoplankton in general and the measured laminarin in specific, our perception of this molecule is that it plays a key role in the carbon cycling in the upper water layers. Based on our measurements, we were able to refine Alderkamp and colleagues' back-of-the-envelope calculation more precisely to an estimated 18±9 gigatons of the 47.5 gigatons of annual phytoplankton organic carbon production that is in the form of laminarin (Alderkamp, van Rijssel, et al., 2007; Field et al., 1998). Based on the limited number of samples and stations, it was not reasonable to further advance this into a more sophisticated global modelling of laminarin abundance in surface waters. Further sampling in more oceanic regions would be necessary to be able to do this kind of simulation.

One aspect of laminarin cycling that has not been focused so far in this work is the diel changes in cellular laminarin content that might influence the results of our study and need to be investigated more thoroughly. Based on acid hydrolysis measurements, it has been shown that the glucan content of microalgae changes with varying light conditions (Granum et al., 2002; Vårum and Myklestad, 1984; Barlow, 1982; Handa, 1969). The hypothesis is that storage glycans like laminarin are being produced throughout the day when high irradiance is fueling photosynthetic activity, and it is being consumed by the microalgae during the night to cover its energy and carbon requirements in this period of low irradiance. Although our preliminary data needs to be replicated again in order to achieve smaller error bars, we were able to show the exact same progression based on our laminarin assay and *Thalassiosira weissflogii* cultures (FIG 3.2). It seems that this effect needs to be taken into account in future environmental samplings. A higher resolution sampling of an algae bloom with more than one or two time points over several days, accompanied by irradiance measurements, would give insights into this phenomenon and whether it is as dominant in nature as it is in the lab. In regard to the past study, this result implicates that the comparably low Lam:POC values in the North Sea dataset might be explained

simply by the fact that the sampling took place at approximately 9 o'clock in the morning (FIG 2.4.3).

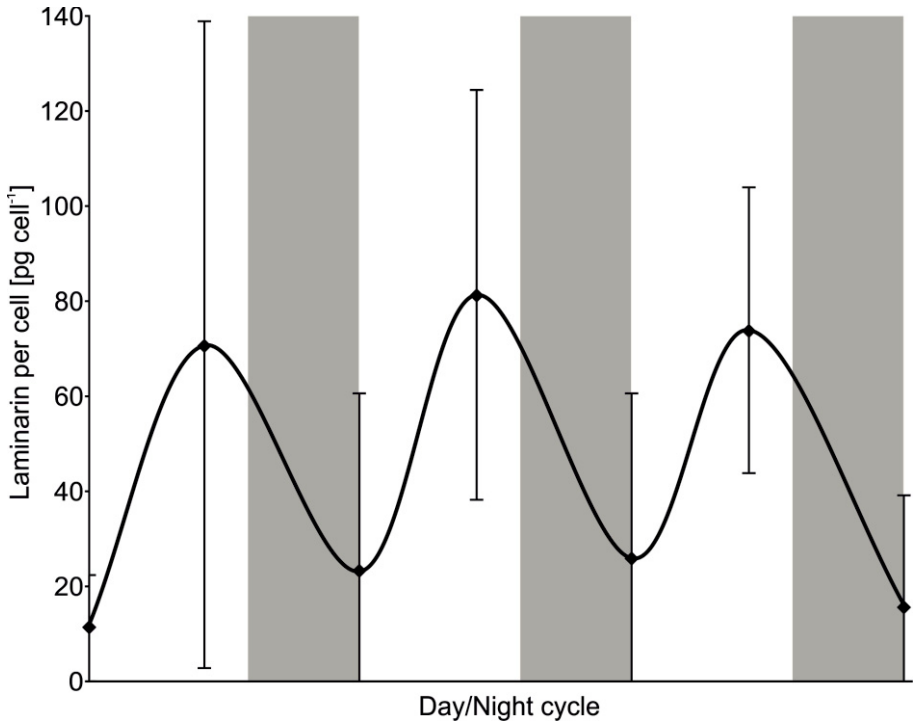


FIG 3.2 Preliminary data of the laminarin cycling of *Thalassiosira weissflogii* over several day and night cycles. The 12 h of light are indicated in white, whereas the 12 h of darkness are indicated in grey. Error bars indicate standard deviations of the biological triplicates. The diatom strain was constantly held in exponential growth conditions. Laminarin was quantified by the biocatalytic assay and cells were counted microscopically.

Although were able to highlight the importance of the marine polysaccharide laminarin by its sheer abundance in the ocean, this project should only be a start for further investigations of both polysaccharides in the marine carbon cycle in general and the cycling of laminarin in specific. In the present study, we measured the potential of laminarin degradation. We assume that the vast majority of the laminarin that was detected by our assay is derived from the intracellular compartment of microalgae that are still intact during sampling. Laminarin degradation or turnover rates were not addressed. By using techniques like isotope-labelled substrate incubations, upcoming studies have to investigate the question whether laminarin is actually the fast food of the ocean, despite our analyses of vertical laminarin profiles that did not support this notion. Despite the limitations of this approach, we are confident to say that we identified laminarin as one of Earth's most synthesized bioenergy molecules.

Another reason why laminarin is only the beginning of the glycobiological approach of broadening our understanding of the marine carbon cycle is the need for an expansion of the biocatalytic toolkit for other marine polysaccharides. Measuring a single polysaccharide in the ocean has never been tried before, but we were able to demonstrate that it is possible for a rather uncomplex polysaccharide. There are no obvious reasons why a similar approach based on more

enzyme activities should not be possible for more complex and maybe also more recalcitrant polysaccharides, e.g. the sulfated fucoidan in brown algae and diatoms. Apart from our impulse to better understand the marine carbon cycle on a basic molecular level, commercial tools for the dissection of marine polysaccharides are still poorly developed in comparison to the advanced toolkits for terrestrial carbohydrates. This is despite the fact that marine glycans are becoming more relevant in several fields of applications. Potentially beneficial health impacts have been reported for both laminarin (Han et al., 2017; Schmidt et al., 2017; Xue Liu et al., 2017; Nguyen et al., 2016; Devillé et al., 2007; Miyanishi et al., 2003) and fucoidan (Zaporozhets and Besednova, 2016; Fitton et al., 2015; Ale et al., 2011). Furthermore, in recent years microalgae in general and laminarin in specific have been proven to be a potent substrate for the production of biofuels (Manns et al., 2016; Montingelli et al., 2016; Motone et al., 2016; Sharma and Horn, 2016; Enquist-Newman et al., 2014; Ho et al., 2013; Wei et al., 2013; Wargacki et al., 2012; JMM Adams et al., 2009, 2011; Horn et al., 2000). Where starch is one of the most abundant terrestrial bioenergy sources, it seems that in this respect laminarin plays a very similar role in the marine system, making it the starch of the ocean.

BIBLIOGRAPHY

Adams, J. M. M., Gallagher, J. A., & Donnison, I. S. 2009. Fermentation study on Saccharina latissima for bioethanol production considering variable pre-treatments. *Journal of Applied Phycology*, *21*(5), 569–574. DOI: 10.1007/s10811-008-9384-7

Adams, J. M. M., Toop, T. A., Donnison, I. S., & Gallagher, J. A. 2011. Seasonal variation in Laminaria digitata and its impact on biochemical conversion routes to biofuels. *Bioresource Technology*, *102*(21), 9976–9984. DOI: 10.1016/j.biortech.2011.08.032

Adams, P., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., … Zwart, P. H. 2010. PHENIX : a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D Biological Crystallography*, *66*(2), 213–221. DOI: 10.1107/S0907444909052925

Alderkamp, A., Buma, A., & van Rijssel, M. 2007. The carbohydrates of Phaeocystis and their degradation in the microbial food web. In *Phaeocystis, major link in the biogeochemical cycling of climate-relevant elements* (pp. 99–118). Dordrecht: Springer Netherlands. DOI: 10.1007/978-1-4020-6214-8_9

Alderkamp, A., Nejstgaard, J. C., Verity, P. G., Zirbel, M. J., Sazhin, A. F., & van Rijssel, M. 2006. Dynamics in carbohydrate composition of Phaeocystis pouchetii colonies during spring blooms in mesocosms. *Journal of Sea Research*, *55*(3), 169–181. DOI: 10.1016/j.seares.2005.10.005

Alderkamp, A., van Rijssel, M., & Bolhuis, H. 2007. Characterization of marine bacteria and the activity of their enzyme systems involved in degradation of the algal storage glucan laminarin. *FEMS Microbiology Ecology*, *59*(1), 108–17. DOI: 10.1111/j.1574-6941.2006.00219.x

Ale, M. T., Maruyama, H., Tamauchi, H., Mikkelsen, J. D., & Meyer, A. S. 2011. Fucose-Containing Sulfated Polysaccharides from Brown Seaweeds Inhibit Proliferation of Melanoma Cells and Induce Apoptosis by Activation of Caspase-3 in Vitro. *Marine Drugs*, *9*(12), 2605–2621. DOI: 10.3390/md9122605

Allison, S. D. 2005. Cheaters, diffusion and nutrients constrain decomposition by microbial enzymes in spatially structured environments. *Ecology Letters*, *8*(6), 626–635. DOI: 10.1111/j.1461-0248.2005.00756.x

Alonso-Sáez, L., & Gasol, J. M. 2007. Seasonal variations in the contributions of different bacterial groups to the uptake of low-molecular-weight compounds in northwestern Mediterranean coastal waters. *Applied and Environmental Microbiology*, *73*(11), 3528–35. DOI: 10.1128/AEM.02627-06

Aluwihare, L. I., & Repeta, D. 1999. A comparison of the chemical characteristics of oceanic DOM and extracellular DOM produced by marine algae. *Marine Ecology Progress Series*, *186*(8), 105–117. DOI: 10.3354/meps186105

Aluwihare, L. I., Repeta, D. J., & Chen, R. F. 1997. A major biopolymeric component to dissolved organic carbon in surface sea water. *Nature*, *387*(6629), 166–169. DOI: 10.1038/387166a0

Antelmann, H., Tjalsma, H., Voigt, B., Ohlmeier, S., Bron, S., van Dijl, J. M., & Hecker, M. 2001. A proteomic view on genome-based signal peptide predictions. *Genome Research*, *11*(9), 1484–502. DOI: 10.1101/gr.182801

Arístegui, J., Barton, E. D., Álvarez-Salgado, X. A., Santos, A. M. P., Figueiras, F. G., Kifani, S., … Demarcq, H. 2009. Sub-regional ecosystem variability in the Canary Current upwelling. *Progress in Oceanography*, *83*(1–4), 33–48. DOI: 10.1016/j.pocean.2009.07.031

Arnosti, C. 1995. Measurement of depth- and site-related differences in polysaccharide hydrolysis rates in marine sediments. *Geochimica et Cosmochimica Acta*, *59*(20), 4247–4257. DOI: 10.1016/0016-7037(95)00247-W

Arnosti, C. 2003. Fluorescent derivatization of polysaccharides and carbohydrate-containing biopolymers for measurement of enzyme activities in complex media. *Journal of Chromatography B*, *793*(1), 181–191. DOI: 10.1016/S1570-0232(03)00375-1

Arnosti, C. 2011. Microbial extracellular enzymes and the marine carbon cycle. *Annual Review of Marine Science*, *3*, 401–25. DOI: 10.1146/annurev-marine-120709-142731

Arnosti, C., Durkin, S., & Jeffrey, W. 2005. Patterns of extracellular enzyme activities among pelagic marine microbial communities: implications for cycling of dissolved organic carbon. *Aquatic Microbial Ecology*, *38*(2), 135–145. DOI: 10.3354/ame038135

Arnosti, C., Fuchs, B. M., Amann, R., & Passow, U. 2012. Contrasting extracellular enzyme activities of particle-associated bacteria from distinct provinces of the North Atlantic Ocean. *Frontiers in Microbiology*, *3*(DEC), 425. DOI: 10.3389/fmicb.2012.00425

Arnosti, C., & Holmer, M. 1999. Carbohydrate dynamics and contributions to the carbon budget of an organic-rich coastal sediment. *Geochimica et Cosmochimica Acta*, *63*(3–4), 393–403. DOI: 10.1016/S0016-7037(99)00076-9

Azam, F. 1998. Microbial Control of Oceanic Carbon Flux: The Plot Thickens. *Science*, *280*(5364), 694–696. DOI: 10.1126/science.280.5364.694

Azam, F., Fenchel, T., Field, J., Gray, J., Meyer-Reil, L., & Thingstad, F. 1983. The Ecological Role of Water-Column Microbes in the Sea. *Marine Ecology Progress Series*, *10*(11), 257–263. DOI: 10.3354/meps010257

Barbeyron, T., Thomas, F., Barbe, V., Teeling, H., Schenowitz, C., Dossat, C., … Michel, G. 2016. Habitat and taxon as driving forces of carbohydrate catabolism in marine heterotrophic bacteria: example of the model algae-associated bacterium Zobellia galactanivorans Dsij T. *Environmental Microbiology*, *18*(12), 4610–4627. DOI: 10.1111/1462-2920.13584

Barlow, R. G. 1982. Phytoplankton ecology in the Southern Benguela current. II. Carbon assimilation patterns. *Journal of Experimental Marine Biology and Ecology*, *63*(3), 229–237. DOI: 10.1016/0022-0981(82)90180-0

Bauer, M., Kube, M., Teeling, H., Richter, M., Lombardot, T., Allers, E., … Glöckner, F. O. 2006. Whole genome analysis of the marine Bacteroidetes "Gramella forsetii" reveals adaptations to degradation of polymeric organic matter. *Environmental Microbiology*, *8*(12), 2201–2213. DOI: 10.1111/j.1462-2920.2006.01152.x

Beattie, A., Hirst, E. L., & Percival, E. 1961. Studies on the metabolism of the Chrysophyceae. Comparative structural investigations on leucosin (chrysolaminarin) separated from diatoms and laminarin from the brown algae. *Biochemical Journal*, *79*(3), 531–537. DOI: 10.1042/bj0790531

Becker, K. W., Elling, F. J., Yoshinaga, M. Y., Söllinger, A., Urich, T., & Hinrichs, K. U. 2016. Unusual butane- and pentanetriol-based tetraether lipids in Methanomassiliicoccus Luminyensis, a representative of the seventh order of methanogens. *Applied and Environmental Microbiology*, *82*(15), 4505–4516. DOI: 10.1128/AEM.00772-16

Becker, S., Scheffel, A., Polz, M. F., & Hehemann, J.-H. 2017. Accurate Quantification of Laminarin in Marine Organic Matter with Enzymes from Marine Microbes. *Applied and Environmental Microbiology*, *83*(9), e03389-16. DOI: 10.1128/AEM.03389-16

Benner, R., Pakulski, J. D., McCarthy, M., Hedges, J., & Hatcher, P. 1992. Bulk chemical characteristics of dissolved organic matter in the ocean. *Science (New York, N.Y.)*, *255*(5051), 1561–4. DOI: 10.1126/science.255.5051.1561

Bennke, C. M., Krüger, K., Kappelmann, L., Huang, S., Gobet, A., Schüler, M., … Amann, R. 2016. Polysaccharide utilisation loci of Bacteroidetes from two contrasting open ocean sites in the North Atlantic. *Environmental Microbiology*, *18*(12), 4456–4470. DOI: 10.1111/1462-2920.13429

Berlemont, R., & Martiny, A. C. 2016. Glycoside Hydrolases across Environmental Microbial Communities. *PLoS Computational Biology*, *12*(12), e1005300. DOI: 10.1371/journal.pcbi.1005300

Biddanda, B., & Benner, R. 1997. Carbon, nitrogen, and carbohydrate fluxes during the production of particulate and dissolved organic matter by marine phytoplankton. *Limnology and Oceanography*, *42*(3), 506–518. DOI: 10.4319/lo.1997.42.3.0506

Biersmith, A., & Benner, R. 1998. Carbohydrates in phytoplankton and freshly produced dissolved organic matter. *Marine Chemistry*, *63*(1–2), 131–144. DOI: 10.1016/S0304-4203(98)00057-7

Bissaro, B., Monsan, P., Fauré, R., & O'Donohue, M. J. 2015. Glycosynthesis in a waterworld: new insight into the molecular basis of transglycosylation in retaining glycoside hydrolases. *Biochemical Journal*, *467*(1), 17–35. DOI: 10.1042/BJ20141412

Borch, N. H., & Kirchman, D. L. 1997. Concentration and composition of dissolved combined neutral sugars (polysaccharides) in seawater determined by HPLC-PAD. *Marine Chemistry*, *57*(1–2), 85–95. DOI: 10.1016/S0304-4203(97)00002-9

Bourgeois, S., Kerhervé, P., Calleja, M. L., Many, G., & Morata, N. 2016. Glacier inputs influence organic matter composition and prokaryotic distribution in a high Arctic fjord (Kongsfjorden, Svalbard). *Journal of Marine Systems*, *164*, 112–127. DOI: 10.1016/j.jmarsys.2016.08.009

Brunt, K., Sanders, P., & Rozema, T. 1998. The Enzymatic Determination of Starch in Food, Feed and Raw Materials of the Starch Industry. *Starch - Stärke*, *50*(10), 413–419. DOI: 10.1002/(SICI)1521-379X(199810)50:10<413::AID-STAR413>3.0.CO;2-F

Brussaard, C., Riegman, R., Noordeloos, A., Cadée, G., Witte, H., Kop, A., … Bak, R. 1995. Effects of grazing, sedimentation and phytoplankton cell lysis on the structure of a coastal pelagic food web. *Marine Ecology Progress Series*, *123*(1–3), 259–271. DOI: 10.3354/meps123259

Buchan, A., LeCleir, G. R., Gulvik, C. A., & González, J. M. 2014. Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nature Reviews Microbiology*, *12*(10), 686–698. DOI: 10.1038/nrmicro3326

Bunse, C., & Pinhassi, J. 2017. Marine Bacterioplankton Seasonal Succession Dynamics. *Trends in Microbiology*, *25*(6), 494–505. DOI: 10.1016/j.tim.2016.12.013

Bushnell, B. 2016. BBMap short read aligner. Berkeley, CA, USA.

Bycroft, M. 1999. The structure of a PKD domain from polycystin-1: implications for polycystic kidney disease. *The EMBO Journal*, *18*(2), 297–305. DOI: 10.1093/emboj/18.2.297

Caballero, M. A., Jallet, D., Shi, L., Rithner, C., Zhang, Y., & Peers, G. 2016. Quantification of chrysolaminarin from the model diatom Phaeodactylum tricornutum. *Algal Research*, *20*, 180–188. DOI: 10.1016/j.algal.2016.10.008

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. 2009. BLAST+: architecture and applications. *BMC Bioinformatics*, *10*(1), 421. DOI: 10.1186/1471-2105-10-421

Candiano, G., Bruschi, M., Musante, L., Santucci, L., Ghiggeri, G. M., Carnemolla, B., … Righetti, P. G. 2004. Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis*, *25*(9), 1327–33. DOI: 10.1002/elps.200305844

Cardman, Z., Arnosti, C., Durbin, A., Ziervogel, K., Cox, C., Steen, A. D., & Teske, A. 2014. Verrucomicrobia Are Candidates for Polysaccharide-Degrading Bacterioplankton in an Arctic Fjord of Svalbard. *Applied and Environmental Microbiology*, *80*(12), 3749–3756. DOI: 10.1128/AEM.00899-14

Carlson, C. A. 2002. Production and Removal Processes. In D. A. Hansell & C. A. Carlson (Eds.), *Biogeochemistry of Marine Dissolved Organic Matter* (pp. 91–151). Elsevier. DOI: 10.1016/B978-012323841-2/50006-3

Chen, L., Sadek, M., Stone, B. a, Brownlee, R. T. C., Fincher, G. B., & Høj, P. B. 1995. Stereochemical course of glucan hydrolysis by barley (1 → 3)- and (1 → 3,1 → 4)-β-glucanases. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, *1253*(1), 112–116. DOI: 10.1016/0167-4838(95)00157-P

Chiovitti, A., Molino, P., Crawford, S. a, Teng, R., Spurck, T., & Wetherbee, R. 2004. The glucans extracted with warm water from diatoms are mainly derived from intracellular chrysolaminaran and not extracellular polysaccharides. *European Journal of Phycology*, *39*(2), 117–28. DOI: 10.1080/0967026042000201885

Chizhov, A. O., Dell, A., Morris, H. R., Reason, A. J., Haslam, S. M., McDowell, R. A., … Usov, A. I. 1998. Structural analysis of laminarans by MALDI and FAB mass spectrometry. *Carbohydrate Research*, *310*(3), 203–10. DOI: 10.1016/S0008-6215(98)00177-3

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., … Thornton, P. 2013. Carbon and Other Biogeochemical Cycles. In Intergovernmental Panel on Climate Change (Ed.), *Climate Change 2013 - The Physical Science Basis* (pp. 465–570). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781107415324.015

Cloern, J. E. 1996. Phytoplankton bloom dynamics in coastal ecosystems: A review with some general lessons from sustained investigation of San Francisco Bay, California. *Reviews of Geophysics*, *34*(2), 127–168. DOI: 10.1029/96RG00986

Cockburn, D. W., & Koropatkin, N. M. 2016. Polysaccharide Degradation by the Intestinal Microbiota and Its Influence on Human Health and Disease. *Journal of Molecular Biology*, *428*(16), 3230–3252. DOI: 10.1016/j.jmb.2016.06.021

Cottrell, M. T., & Kirchman, D. L. 2000a. Community Composition of Marine Bacterioplankton Determined by 16S rRNA Gene Clone Libraries and Fluorescence In Situ Hybridization. *Applied and Environmental Microbiology*, *66*(12), 5116–5122. DOI: 10.1128/AEM.66.12.5116-5122.2000

Cottrell, M. T., & Kirchman, D. L. 2000b. Natural Assemblages of Marine Proteobacteria and Members of the Cytophaga-Flavobacter Cluster Consuming Low- and High-Molecular-Weight Dissolved Organic Matter. *Applied and Environmental Microbiology*, *66*(4), 1692–1697. DOI: 10.1128/AEM.66.4.1692-1697.2000

Cowie, G. L., & Hedges, J. 1984. Carbohydrate sources in a coastal marine environment. *Geochimica et Cosmochimica Acta*, *48*(10), 2075–2087. DOI: 10.1016/0016-7037(84)90388-0

Cowtan, K. 2006. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallographica Section D Biological Crystallography*, *62*(9), 1002–1011. DOI: 10.1107/S0907444906022116

Cox, J., & Mann, M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, *26*(12), 1367–1372. DOI: 10.1038/nbt.1511

Cuskin, F., Lowe, E. C., Temple, M. J., Zhu, Y., Cameron, E. a, Pudlo, N. a, … Gilbert, H. 2015. Human gut Bacteroidetes can utilize yeast mannan through a selfish mechanism. *Nature*, *517*(7533), 165–169. DOI: 10.1038/nature13995

d'Ippolito, G., Sardo, A., Paris, D., Vella, F. M., Adelfi, M. G., Botte, P., … Fontana, A. 2015. Potential of lipid metabolism in marine diatoms for biofuel production. *Biotechnology for Biofuels*, *8*(1), 28. DOI: 10.1186/s13068-015-0212-4

Davies, G., & Henrissat, B. 1995. Structures and mechanisms of glycosyl hydrolases. *Structure (London, England : 1993)*, *3*(9), 853–859. DOI: 10.1016/S0969-2126(01)00220-9

Davies, G., Wilson, K. S., & Henrissat, B. 1997. Nomenclature for sugar-binding subsites in glycosyl hydrolases. *The Biochemical Journal*, *321 (Pt 2*, 557–9. DOI: 10.1007/s007920050009

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., … Velayoudon, D. 2015. Eukaryotic plankton diversity in the sunlit ocean. *Science*, *348*(6237), 1261605–1261605. DOI: 10.1126/science.1261605

Devillé, C., Gharbi, M., Dandrifosse, G., & Peulen, O. 2007. Study on the effects of laminarin, a polysaccharide from seaweed, on gut characteristics. *Journal of the Science of Food and Agriculture*, *87*(9), 1717–1725. DOI: 10.1002/jsfa.2901

Ducklow, H., Steinberg, D., & Buesseler, K. 2001. Upper Ocean Carbon Export and the Biological Pump. *Oceanography*, *14*(4), 50–58. DOI: 10.5670/oceanog.2001.06

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–7. DOI: 10.1093/nar/gkh340

Elifantz, H., Dittel, A., Cottrell, M., & Kirchman, D. L. 2007. Dissolved organic matter assimilation by heterotrophic bacterial groups in the western Arctic Ocean. *Aquatic Microbial Ecology*, *50*(1), 39–49. DOI: 10.3354/ame01145

Elifantz, H., Malmstrom, R. R., Cottrell, M. T., & Kirchman, D. L. 2005. Assimilation of polysaccharides and glucose by major bacterial groups in the Delaware Estuary. *Applied and Environmental Microbiology*, *71*(12), 7799–805. DOI: 10.1128/AEM.71.12.7799-7805.2005

Engel, A., & Händel, N. 2011. 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*, *127*(1–4), 180–91. DOI: 10.1016/j.marchem.2011.09.004

Engel, A., Harlay, J., Piontek, J., & Chou, L. 2012. Contribution of combined carbohydrates to dissolved and particulate organic carbon after the spring bloom in the northern Bay of Biscay (North-Eastern Atlantic Ocean). *Continental Shelf Research*, *45*, 42–53. DOI: 10.1016/j.csr.2012.05.016

Engel, A., Thoms, S., Riebesell, U., Rochelle-Newall, E., & Zondervan, I. 2004. Polysaccharide aggregation as a potential sink of marine dissolved organic carbon. *Nature*, *428*(6986), 929–932. DOI: 10.1038/nature02453

Enquist-Newman, M., Faust, A. M. E., Bravo, D. D., Santos, C. N. S., Raisner, R. M., Hanel, A., … Yoshikuni, Y. 2014. Efficient ethanol production from brown macroalgae sugars by a synthetic yeast platform. *Nature*, *505*(7482), 239–243. DOI: 10.1038/nature12771

Fabiano, M., & Danovaro, R. 1994. Composition of organic matter in sediments facing a river estuary (Tyrrhenian Sea): relationships with bacteria and microphytobenthic biomass. *Hydrobiologia*, *277*(2), 71–84. DOI: 10.1007/BF00016755

Falkowski, P. G., & Chen, Y.-B. 2003. Photoacclimation of Light Harvesting Systems in Eukaryotic Algae. In B. Green & W. Parson (Eds.), *Light-Harvesting Antennas in Photosynthesis* (pp. 423–447). Dordrecht, Netherlands: Springer Science+Business Media. DOI: 10.1007/978-94-017-2087-8_15

Feller, G., Narinx, E., Arpigny, J. L., Aittaleb, M., Baise, E., Genicot, S., & Gerday, C. 1996. Enzymes from psychrophilic organisms. *FEMS Microbiology Reviews*, *18*(2–3), 189–202. DOI: 10.1111/j.1574-6976.1996.tb00236.x

Fernández-Gómez, B., Richter, M., Schüler, M., Pinhassi, J., Acinas, S. G., González, J. M., & Pedrós-Alió, C. 2013. Ecology of marine Bacteroidetes: a comparative genomics approach. *The ISME Journal*, *7*(5), 1026–37. DOI: 10.1038/ismej.2012.169

Fichez, R. 1991. Composition and fate of organic-matter in submarine cave sediments-implications for the biogeochemical cycle of organic-carbon. *Oceanologica Acta*, *14*(4), 369–377.

Field, R. A., Behrenfeld, M. J., Randerson, & Falkowski, P. G. 1998. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science (New York, N.Y.)*, *281*(5374), 237–40. DOI: 10.1126/science.281.5374.237

Fitton, J., Stringer, D., & Karpiniec, S. 2015. Therapies from Fucoidan: An Update. *Marine Drugs*, *13*(12), 5920–5946. DOI: 10.3390/md13095920

Fogg, G. E. 1983. The Ecological Significance of Extracellular Products of Phytoplankton Photosynthesis. *Botanica Marina*, *26*(1), 3–14. DOI: 10.1515/botm.1983.26.1.3

Fuhrman, J. A. 1999. Marine viruses and their biogeochemical and ecological effects. *Nature*, *399*(6736), 541–8. DOI: 10.1038/21119

Gal, A., Wirth, R., Kopka, J., Fratzl, P., Faivre, D., & Scheffel, A. 2016. Macromolecular recognition directs calcium ions to coccolith mineralization sites. *Science (New York, N.Y.)*, *353*(6299), 590–3. DOI: 10.1126/science.aaf7889

Gattuso, J.-P., Frankignoulle, M., & Wollast, R. 1998. Carbon and Carbonate Metabolism in Coastal Aquatic Ecosystems. *Annual Review of Ecology and Systematics*, *29*(1), 405–434. DOI: 10.1146/annurev.ecolsys.29.1.405

Gibson, D., Young, L., Chuang, R., Venter, J., Hutchison, C., & Smith, H. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), 343–5. DOI: 10.1038/nmeth.1318

Gilbert, H. 2010. The biochemistry and structural biology of plant cell wall deconstruction. *Plant Physiology*, *153*(2), 444–55. DOI: 10.1104/pp.110.156646

Glöckner, F. O., Fuchs, B. M., & Amann, R. 1999. Bacterioplankton compositions of lakes and oceans: a first comparison based on fluorescence in situ hybridization. *Applied and Environmental Microbiology*, *65*(8), 3721–6.

Graiff, A., Ruth, W., Kragl, U., & Karsten, U. 2015. Chemical characterization and quantification of the brown algal storage compound laminarin — A new methodological approach. *Journal of Applied Phycology*. DOI: 10.1007/s10811-015-0563-z

Granum, E., Kirkvold, S., & Myklestad, S. 2002. Cellular and extracellular production of carbohydrates and amino acids by the marine diatom Skeletonema costatum: diel variations and effects of N depletion. *Marine Ecology Progress Series*, *242*(Werner 1977), 83–94. DOI: 10.3354/meps242083

Granum, E., & Myklestad, S. 2002. A simple combined method for determination of β-1,3-glucan and cell wall polysaccharides in diatoms. *Hydrobiologia*, *477*(1972), 155–61. DOI: 10.1023/A:1021077407766

Groisillier, A., Labourel, A., Michel, G., & Tonon, T. 2015. The Mannitol Utilization System of the Marine Bacterium Zobellia galactanivorans. *Applied and Environmental Microbiology*, *81*(5), 1799–1812. DOI: 10.1128/AEM.02808-14

Grondin, J. M., Tamura, K., Déjean, G., Abbott, D. W., & Brumer, H. 2017. Polysaccharide Utilization Loci: Fueling Microbial Communities. *Journal of Bacteriology*, *199*(15), JB.00860-16. DOI: 10.1128/JB.00860-16

Guillard, R., & Ryther, J. 1962. Studies of marine planktonic diatoms. I. Cyclotella nana Hustedt, and Detonula confervacea (cleve) Gran. *Canadian Journal of Microbiology*, *8*(1140), 229–39. DOI: 10.1139/m62-029

Guillard, R., & Wangersky, P. J. 1958. The Production of Extracellular Carbohydrates by Some Marine Flagellates1. *Limnology and Oceanography*, *3*(4), 449–454. DOI: 10.4319/lo.1958.3.4.0449

Guschina, I. A., & Harwood, J. L. 2006. Lipids and lipid metabolism in eukaryotic algae. *Progress in Lipid Research*, *45*(2), 160–86. DOI: 10.1016/j.plipres.2006.01.001

Guy, L., Roat Kultima, J., & Andersson, S. G. E. 2010. genoPlotR: comparative gene and genome visualization in R. *Bioinformatics*, *26*(18), 2334–2335. DOI: 10.1093/bioinformatics/btq413

Hagström, Å., Pinhassi, J., & Zweifel, U. 2000. Biogeographical diversity among marine bacterioplankton. *Aquatic Microbial Ecology*, *21*(3), 231–244. DOI: 10.3354/ame021231

Hahnke, R. L., Bennke, C. M., Fuchs, B. M., Mann, A. J., Rhiel, E., Teeling, H., … Harder, J. 2015. Dilution cultivation of marine heterotrophic bacteria abundant after a spring phytoplankton bloom in the North Sea. *Environmental Microbiology*, *17*(10), 3515–26. DOI: 10.1111/1462-2920.12479

Han, M. H., Lee, D.-S., Jeong, J.-W., Hong, S.-H., Choi, I.-W., Cha, H.-J., … Hyun Choi, Y. 2017. Fucoidan Induces ROS-Dependent Apoptosis in 5637 Human Bladder Cancer Cells by Downregulating Telomerase Activity via Inactivation of the PI3K/Akt Signaling Pathway. *Drug Development Research*, *78*(1), 37–48. DOI: 10.1002/ddr.21367

Handa, N. 1969. Carbohydrate metabolism in the marine diatom Skeletonema costatum. *Marine Biology*, *4*(3), 208–214. DOI: 10.1007/BF00393894

Harrison, P. J., Waters, R. E., & Taylor, F. J. R. 1980. A Broad Spectrum Artificial Sea Water Medium for Coastal and Open Ocean Phytoplankton. *Journal of Phycology*, *16*(1), 28–35. DOI: 10.1111/j.0022-3646.1980.00028.x

Harvey, A. J., Hrmova, M., De Gori, R., Varghese, J. N., & Fincher, G. B. 2000. Comparative modeling of the three-dimensional structures of family 3 glycoside hydrolases. *Proteins: Structure, Function, and Genetics*, *41*(2), 257–269. DOI: 10.1002/1097-0134(20001101)41:2<257::AID-PROT100>3.0.CO;2-C

He, F. 2011. Laemmli-SDS-PAGE. *BIO-PROTOCOL*, *1*(11). DOI: 10.21769/BioProtoc.80

Hedges, J., Baldock, J., Gélinas, Y., Lee, C., Peterson, M., & Wakeham, S. 2001. Evidence for non-selective preservation of organic matter in sinking marine particles. *Nature*, *409*(6822), 801–4. DOI: 10.1038/35057247

Hehemann, J.-H., Arevalo, P., Datta, M. S., Yu, X., Corzett, C. H., Henschel, A., … Polz, M. F. 2016. Adaptive radiation by waves of gene transfer leads to fine-scale resource partitioning in marine microbes. *Nature Communications*, *7*, 12860. DOI: 10.1038/ncomms12860

Hehemann, J.-H., Boraston, A. B., & Czjzek, M. 2014. A sweet new wave: structures and mechanisms of enzymes that digest polysaccharides from marine algae. *Current Opinion in Structural Biology*, *28*(1), 77–86. DOI: 10.1016/j.sbi.2014.07.009

Hehemann, J.-H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., & Michel, G. 2010. Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature*, *464*(7290), 908–912. DOI: 10.1038/nature08937

Hehemann, J.-H., Correc, G., Thomas, F., Bernard, T., Barbeyron, T., Jam, M., … Czjzek, M. 2012. Biochemical and structural characterization of the complex agarolytic enzyme system from the marine bacterium Zobellia galactanivorans. *The Journal of Biological Chemistry*, *287*(36), 30571–30584. DOI: 10.1074/jbc.M112.377184

Hehemann, J.-H., Kelly, A. G., Pudlo, N. a, Martens, E. C., & Boraston, A. B. 2012. Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(48), 19786–91. DOI: 10.1073/pnas.1211002109

Hehemann, J.-H., Truong, L. Van, Unfried, F., Welsch, N., Kabisch, J., Heiden, S. E., … Schweder, T. 2017. Aquatic adaptation of a laterally acquired pectin degradation pathway in marine Gammaproteobacteria. *Environmental Microbiology*. DOI: 10.1111/1462-2920.13726

Heinz, E., Williams, T. A., Nakjang, S., Noël, C. J., Swan, D. C., Goldberg, A. V., … Embley, T. M. 2012. The Genome of the Obligate Intracellular Parasite Trachipleistophora hominis: New Insights into Microsporidian Genome Dynamics and Reductive Evolution. *PLoS Pathogens*, *8*(10), e1002979. DOI: 10.1371/journal.ppat.1002979

Hellebust, J. A. 1965. Excretion of some organic compounds by marine phytoplankton. *Limnology and Oceanography*, *10*(2), 192–206. DOI: 10.4319/lo.1965.10.2.0192

Ho, S.-H., Huang, S.-W., Chen, C.-Y., Hasunuma, T., Kondo, A., & Chang, J.-S. 2013. Bioethanol production using carbohydrate-rich microalgae biomass as feedstock. *Bioresource Technology*, *135*, 191–198. DOI: 10.1016/j.biortech.2012.10.015

Hoagland, K. D., Rosowski, J. R., Gretz, M. R., & Roemer, S. C. 1993. Diatom extracellular polymeric substances: Function, fine structure, chemistry, and physiology. *Journal of Phycology*, *29*(5), 537–566. DOI: 10.1111/j.0022-3646.1993.00537.x

Hofmann, J., Hahm, H. S., Seeberger, P. H., & Pagel, K. 2015. Identification of carbohydrate anomers using ion mobility–mass spectrometry. *Nature*, *526*(7572), 241–244. DOI: 10.1038/nature15388

Hoppe, H.-G. 1983. Significance of exoenzymatic activities in the ecology of brackish water: measurements by means of methylumbelliferyl-substrates. *Marine Ecology Progress Series*, *11*, 299–308. DOI: 10.3354/meps011299

Horn, S. J., Aasen, I. M., & Ostgaard, K. 2000. Ethanol production from seaweed extract. *Journal of Industrial Microbiology and Biotechnology*, *25*(5), 249–254. DOI: 10.1038/sj.jim.7000065

Jackson, P. 1990. The use of polyacrylamide-gel electrophoresis for the high-resolution separation of reducing saccharides labelled with the fluorophore 8-aminonaphthalene-1,3,6-trisulphonic acid. Detection of picomolar quantities by an imaging system based on a cooled cha. *The Biochemical Journal*, *270*(3), 705–13.

Janse, I., van Rijssel, M., Hall, P.-J., Gerwig, G. J., Gottschal, J. C., & Prins, R. A. 1996. The storage glucan of Phaeocystis globosa (Prymnesiophyceae) cells. *Journal of Phycology*, *32*(3), 382–387. DOI: 10.1111/j.0022-3646.1996.00382.x

Jeng, W.-Y., Wang, N.-C., Lin, C.-T., Shyur, L.-F., & Wang, A. H. J. 2011. Crystal Structures of the Laminarinase Catalytic Domain from Thermotoga maritima MSB8 in Complex with Inhibitors: ESSENTIAL RESIDUES FOR -1,3- and -1,4-GLUCAN SELECTION. *Journal of Biological Chemistry*, *286*(52), 45030–45040. DOI: 10.1074/jbc.M111.271213

Jiao, N., Herndl, G. J., Hansell, D. A., Benner, R., Kattner, G., Wilhelm, S. W., … Azam, F. 2010. Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. *Nature Reviews Microbiology*, *8*(8), 593–599. DOI: 10.1038/nrmicro2386

Jiao, N., & Zheng, Q. 2011. The Microbial Carbon Pump: from Genes to Ecosystems. *Applied and Environmental Microbiology*, *77*(21), 7439–7444. DOI: 10.1128/AEM.05640-11

Jones, D. T., Taylor, W. R., & Thornton, J. M. 1992. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics*, *8*(3), 275–282. DOI: 10.1093/bioinformatics/8.3.275

Jooste, P. J., & Hugo, C. J. 1999. The taxonomy, ecology and cultivation of bacterial genera belonging to the family Flavobacteriaceae. *International Journal of Food Microbiology*, *53*(2–3), 81–94. DOI: 10.1016/S0168-1605(99)00162-2

Kabisch, A., Otto, A., König, S., Becher, D., Albrecht, D., Schüler, M., … Schweder, T. 2014. Functional characterization of polysaccharide utilization loci in the marine Bacteroidetes "Gramella forsetii" KT0803. *The ISME Journal*, *8*(7), 1492–502. DOI: 10.1038/ismej.2014.4

Kaoutari, A. El, Armougom, F., Gordon, J. I., Raoult, D., & Henrissat, B. 2013. The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nature Reviews Microbiology*, *11*(7), 497–504. DOI: 10.1038/nrmicro3050

Keith, S., & Arnosti, C. 2001. Extracellular enzyme activity in a river-bay-shelf transect: variations in polysaccharide hydrolysis rates with substrate and size class. *Aquatic Microbial Ecology*, *24*(3), 243–53. DOI: 10.3354/ame024243

Kelly, A. E., Ou, H. D., Withers, R., & Dötsch, V. 2002. Low-conductivity buffers for high-sensitivity NMR measurements. *Journal of the American Chemical Society*, *124*(40), 12013–9. DOI: 10.1021/ja026121b

Kirchman, D. L. 2002. The ecology of Cytophaga-Flavobacteria in aquatic environments. *FEMS Microbiology Ecology*, *39*(2), 91–100. DOI: 10.1111/j.1574-6941.2002.tb00910.x

Koshland, D. 1953. Stereochemistry and the mechanism of enzymatic reactions. *Biological Reviews*, *28*(4), 416–36. DOI: 10.1111/j.1469-185X.1953.tb01386.x

Labourel, A., Jam, M., Jeudy, A., Hehemann, J.-H., Czjzek, M., & Michel, G. 2014. The β-glucanase ZgLamA from Zobellia galactanivorans evolved a bent active site adapted for efficient degradation of algal laminarin. *The Journal of Biological Chemistry*, *289*(4), 2027–42. DOI: 10.1074/jbc.M113.538843

Labourel, A., Jam, M., Legentil, L., Sylla, B., Hehemann, J. H., Ferrieres, V., … Michel, G. 2015. Structural and biochemical characterization of the laminarinase ZgLamCGH16 from Zobellia galactanivorans suggests preferred recognition of branched laminarin. *Acta Crystallographica. Section D, Biological Crystallography*, *71*(Pt 2), 173–184. DOI: 10.1107/S139900471402450X

Laine, R. A. 1994. A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05 x 10(12) structures for a reducing hexasaccharide: the Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology*, *4*(6), 759–67. DOI: 10.1093/glycob/4.6.759

Landa, M., Blain, S., Christaki, U., Monchy, S., & Obernosterer, I. 2016. Shifts in bacterial community composition associated with increased carbon cycling in a mosaic of phytoplankton blooms. *The ISME Journal*, *10*(1), 39–50. DOI: 10.1038/ismej.2015.105

Legendre, L., Rivkin, R. B., Weinbauer, M. G., Guidi, L., & Uitz, J. 2015. The microbial carbon pump concept: Potential biogeochemical significance in the globally changing ocean. *Progress in Oceanography*, *134*, 432–450. DOI: 10.1016/j.pocean.2015.01.008

Lever, M. 1972. A new reaction for colorimetric determination of carbohydrates. *Analytical Biochemistry*, *47*(1), 273–9. DOI: 10.1016/0003-2697(72)90301-6

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. DOI: 10.1093/bioinformatics/btp352

Liu, X., Liu, H., Zhai, Y., Li, Y., Zhu, X., & Zhang, W. 2017. Laminarin protects against hydrogen peroxide-induced oxidative damage in MRC-5 cells possibly via regulating NRF2. *PeerJ*, *5*, e3642. DOI: 10.7717/peerj.3642

Liu, Z., Kobiela, M. E., McKee, G. A., Tang, T., Lee, C., Mulholland, M. R., & Hatcher, P. 2010. The effect of chemical structure on the hydrolysis of tetrapeptides along a river-to-ocean transect: AVFA and SWGA. *Marine Chemistry*, *119*(1–4), 108–120. DOI: 10.1016/j.marchem.2010.01.005

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., & Henrissat, B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research*, *42*(Database issue), D490-5. DOI: 10.1093/nar/gkt1178

Lowry, O. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. 1951. Protein measurement with the Folin phenol reagent. *The Journal of Biological Chemistry*, *193*(1), 265–75. DOI: 10.1016/0304-3894(92)87011-4

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, … Schleifer, K.-H. 2004. ARB: a software environment for sequence data. *Nucleic Acids Research*, *32*(4), 1363–71. DOI: 10.1093/nar/gkh293

Maeda, M., & Nishizawa, K. 1968. Fine structure of laminaran of Eisenia bicyclis. *Journal of Biochemistry*, *63*(2), 199–206.

Mäki-Arvela, P., Salmi, T., Holmbom, B., Willför, S., & Murzin, D. Y. 2011. Synthesis of sugars by hydrolysis of hemicelluloses--a review. *Chemical Reviews*, *111*(9), 5638–66. DOI: 10.1021/cr2000042

Mann, A. J., Hahnke, R. L., Huang, S., Werner, J., Xing, P., Barbeyron, T., … Teeling, H. 2013. The genome of the alga-associated marine flavobacterium Formosa agariphila KMM 3901T reveals a broad potential for degradation of algal polysaccharides. *Applied and Environmental Microbiology*, *79*(21), 6813–22. DOI: 10.1128/AEM.01937-13

Manns, D., Andersen, S. K., Saake, B., & Meyer, A. S. 2016. Brown seaweed processing: enzymatic saccharification of Laminaria digitata requires no pre-treatment. *Journal of Applied Phycology*, *28*(2), 1287–1294. DOI: 10.1007/s10811-015-0663-9

Marker, A. F. H. 1965. Extracellular carbohydrate liberation in the flagellates Isochrysis galbana and Prymnesium parvum. *Journal of the Marine Biological Association of the United Kingdom*, *45*(3), 755. DOI: 10.1017/S002531540001657X

Martens, E. C., Koropatkin, N. M., Smith, T. J., & Gordon, J. I. 2009. Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. *The Journal of Biological Chemistry*, *284*(37), 24673–7. DOI: 10.1074/jbc.R109.022848

Martens, E. C., Lowe, E. C., Chiang, H., Pudlo, N. a, Wu, M., McNulty, N. P., … Gordon, J. I. 2011. Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. *PLoS Biology*, *9*(12). DOI: 10.1371/journal.pbio.1001221

McCarthy, M., Hedges, J., & Benner, R. 1996. Major biochemical composition of dissolved high molecular weight organic matter in seawater. *Marine Chemistry*, *55*(3–4), 281–97. DOI: 10.1016/S0304-4203(96)00041-2

McConville, M. J., Bacic, A., & Clarke, A. E. 1986. Structural studies of chrysolaminaran from the ice diatom Stauroneis amphioxys (Gregory). *Carbohydrate Research*, *153*(2), 330–333. DOI: 10.1016/S0008-6215(00)90276-3

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P., Winn, M. D., Storoni, L. C., & Read, R. J. 2007. Phaser crystallographic software. *Journal of Applied Crystallography*, *40*(4), 658–674. DOI: 10.1107/S0021889807021206

Menshova, R., Ermakova, S., Anastyuk, S., Isakov, V., Dubrovskaya, Y., Kusaykin, M., … Zvyagintseva, T. 2014. Structure, enzymatic transformation and anticancer activity of branched high molecular weight laminaran from brown alga Eisenia bicyclis. *Carbohydrate Polymers*, *99*, 101–9. DOI: 10.1016/j.carbpol.2013.08.037

Middelburg, J. J., & Meysman, F. J. R. 2007. Burial at Sea. *Science*, *316*(5829), 1294–1295. DOI: 10.1126/science.1144001

Miyanishi, N., Iwamoto, Y., Watanabe, E., & Odaz, T. 2003. Induction of TNF-α production from human peripheral blood monocytes with β-1,3-glucan oligomer prepared from laminarin with β-1,3-glucanase from Bacillus clausii NM-1. *Journal of Bioscience and Bioengineering*, *95*(2), 192–195. DOI: 10.1016/S1389-1723(03)80128-7

Montingelli, M. E., Benyounis, K. Y., Stokes, J., & Olabi, A. G. 2016. Pretreatment of macroalgal biomass for biogas production. *Energy Conversion and Management*, *108*, 202–209. DOI: 10.1016/j.enconman.2015.11.008

Moran, M. A., Kujawinski, E. B., Stubbins, A., Fatland, R., Aluwihare, L. I., Buchan, A., … Waldbauer, J. R. 2016. Deciphering ocean carbon in a changing world. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(12), 3143–51. DOI: 10.1073/pnas.1514645113

Moretti, R., & Thorson, J. 2008. A comparison of sugar indicators enables a universal high-throughput sugar-1-phosphate nucleotidyltransferase assay. *Analytical Biochemistry*, *377*(2), 251–8. DOI: 10.1016/j.ab.2008.03.018

Motone, K., Takagi, T., Sasaki, Y., Kuroda, K., & Ueda, M. 2016. Direct ethanol fermentation of the algal storage polysaccharide laminarin with an optimized combination of engineered yeasts. *Journal of Biotechnology*, *231*, 129–135. DOI: 10.1016/j.jbiotec.2016.06.002

Murray, A., Arnosti, C., De La Rocha, C., Grossart, H., & Passow, U. 2007. Microbial dynamics in autotrophic and heterotrophic seawater mesocosms. II. Bacterioplankton community structure and hydrolytic enzyme activities. *Aquatic Microbial Ecology*, *49*(4), 123–141. DOI: 10.3354/ame01139

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., … Vagin, A. A. 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica. Section D, Biological Crystallography*, *67*(Pt 4), 355–67. DOI: 10.1107/S0907444911001314

Myklestad, S. 1974. Production of carbohydrates by marine planktonic diatoms. I. Comparison of nine different species in culture. *Journal of Experimental Marine Biology and Ecology*, *15*(3), 261–74. DOI: 10.1016/0022-0981(74)90049-5

Myklestad, S. 1977. Production of carbohydrates by marine planktonic diatoms. II. Influence of the ratio in the growth medium on the assimilation ratio, growth rate, and production of cellular and extracellular carbohydrates by Chaetoceros affinis var. willei (Gran) Hustedt. *Journal of Experimental Marine Biology and Ecology*, *29*(2), 161–179. DOI: 10.1016/0022-0981(77)90046-6

Myklestad, S. 1995. Release of extracellular products by phytoplankton with special emphasis on polysaccharides. *Science of The Total Environment*, *165*(1–3), 155–164. DOI: 10.1016/0048-9697(95)04549-G

Myklestad, S. 2000. Dissolved Organic Carbon from Phytoplankton. In W. PJ (Ed.), *Marine Chemistry* (pp. 111–148). Berlin/Heidelberg: Springer-Verlag. DOI: 10.1007/10683826_5

Ndeh, D., Rogowski, A., Cartmell, A., Luis, A. S., Baslé, A., Gray, J., … Gilbert, H. 2017. Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature*, *544*(7648), 65–70. DOI: 10.1038/nature21725

Needham, D. M., & Fuhrman, J. A. 2016. Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. *Nature Microbiology*, *1*(4), 16005. DOI: 10.1038/nmicrobiol.2016.5

Nelson, C., Attia, M. A., Rogowski, A., Morland, C., Brumer, H., & Gardner, J. G. 2017. Comprehensive functional characterization of the glycoside hydrolase family 3 enzymes from Cellvibrio japonicus reveals unique metabolic roles in biomass saccharification. *Environmental Microbiology*, *19*(12), 5025–5039. DOI: 10.1111/1462-2920.13959

Nelson, D., Tréguer, P., Brzezinski, M. A., Leynaert, A., & Quéguiner, B. 1995. Production and dissolution of biogenic silica in the ocean: Revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biogeochemical Cycles*, *9*(3), 359–372. DOI: 10.1029/95GB01070

Nguyen, S. G., Kim, J., Guevarra, R. B., Lee, J.-H., Kim, E., Kim, S., & Unno, T. 2016. Laminarin favorably modulates gut microbiota in mice fed a high-fat diet. *Food & Function*, *7*(10), 4193–4201. DOI: 10.1039/C6FO00929H

Obernosterer, I., & Herndl, G. 1995. Phytoplankton extracellular release and bacterial growth: dependence on the inorganic N: P ratio. *Marine Ecology Progress Series*, *116*(1–3), 247–257. DOI: 10.3354/meps116247

Osterholz, H., Niggemann, J., Giebel, H.-A., Simon, M., & Dittmar, T. 2015. Inefficient microbial production of refractory dissolved organic matter in the ocean. *Nature Communications*, *6*(May), 7422. DOI: 10.1038/ncomms8422

Otto, A., Bernhardt, J., Meyer, H., Schaffer, M., Herbst, F. A., Siebourg, J., … Becher, D. 2010. Systems-wide temporal proteomic profiling in glucose-starved Bacillus subtilis. *Nature Communications*, *1*(9). DOI: 10.1038/ncomms1137

Oyama, S., Yamagata, Y., Abe, K., & Nakajima, T. 2002. Cloning and expression of an endo-1,6-beta-D-glucanase gene (neg1) from Neurospora crassa. *Bioscience, Biotechnology, and Biochemistry*, *66*(6), 1378–81. DOI: 10.1271/bbb.66.1378

Painter, T. J. 1983. Algal Polysaccharides. In *The Polysaccharides* (pp. 195–285). Elsevier. DOI: 10.1016/B978-0-12-065602-8.50009-1

Pakulski, J. D., & Benner, R. 1994. Abundance and distribution of carbohydrates in the ocean. *Limnology and Oceanography*. DOI: 10.4319/lo.1994.39.4.0930

Panschin, I., Huang, S., Meier-Kolthoff, J. P., Tindall, B. J., Rohde, M., Verbarg, S., … Hahnke, R. L. 2016. Comparing polysaccharide decomposition between the type strains Gramella echinicola KMM 6050T (DSM 19838T) and Gramella portivictoriae UST040801-001T (DSM 23547T), and emended description of Gramella echinicola Nedashkovskaya et al. 2005 emend. Shahina et. *Standards in Genomic Sciences*, *11*(1), 37. DOI: 10.1186/s40793-016-0163-9

Paulsen, B. S., & Myklestad, S. 1978. Structural studies of the reserve glucan produced by the marine diatom Skeletonema costatum (grev.) Cleve. *Carbohydrate Research*, *62*(2), 386–8. DOI: 10.1016/S0008-6215(00)80888-5

Pedersen, M., & Borum, J. 1996. Nutrient control of algal growth in estuarine waters. Nutrient limitation and the importance of nitrogen requirements and nitrogen storage among phytoplankton and species of macroalgae. *Marine Ecology Progress Series*, *142*(1), 261–272. DOI: 10.3354/meps142261

Pedrotti, M., Beauvais, S., Kerros, M., Iversen, K., & Peters, F. 2009. Bacterial colonization of transparent exopolymeric particles in mesocosms under different turbulence intensities and nutrient conditions. *Aquatic Microbial Ecology*, *55*(3), 301–312. DOI: 10.3354/ame01308

Peplies, J., Kottmann, R., Ludwig, W., & Glöckner, F. O. 2008. A standard operating procedure for phylogenetic inference (SOPPI) using (rRNA) marker genes. *Systematic and Applied Microbiology*, *31*(4), 251–257. DOI: 10.1016/j.syapm.2008.08.003

Percival, E., & Ross, A. 1951. 156. The constitution of laminarin. Part II. The soluble laminarin of Laminaria digitata. *Journal of the Chemical Society (Resumed)*, (0), 720–6. DOI: 10.1039/jr9510000720

Pérez, V., Fernández, E., Marañón, E., Serret, P., & García-Soto, C. 2005. Seasonal and interannual variability of chlorophyll a and primary production in the Equatorial Atlantic: In situ and remote sensing observations. *Journal of Plankton Research*, *27*(2), 189–197. DOI: 10.1093/plankt/fbh159

Peterson, G. L. 1977. A simplification of the protein assay method of Lowry et al. which is more generally applicable. *Analytical Biochemistry*, *83*(2), 346–56. DOI: 10.1016/0003-2697(77)90043-4

Pinhassi, J., Sala, M. M., Havskum, H., Peters, F., Guadayol, O., Malits, A., & Marrasé, C. 2004. Changes in bacterioplankton composition under different phytoplankton regimens. *Applied and Environmental Microbiology*, *70*(11), 6753–66. DOI: 10.1128/AEM.70.11.6753-6766.2004

Pommier, T., Canbäck, B., Riemann, L., Boström, K. H., Simu, K., Lundberg, P., … Hagström, A. 2007. Global patterns of diversity and community structure in marine bacterioplankton. *Molecular Ecology*, *16*(4), 867–80. DOI: 10.1111/j.1365-294X.2006.03189.x

Poulton, A. J., Holligan, P. M., Hickman, A., Kim, Y.-N., Adey, T. R., Stinchcombe, M. C., … Woodward, E. M. S. 2006. Phytoplankton carbon fixation, chlorophyll-biomass and diagnostic pigments in the Atlantic Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, *53*(14–16), 1593–1610. DOI: 10.1016/j.dsr2.2006.05.007

Pruesse, E., Peplies, J., & Glöckner, F. O. 2012. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics*, *28*(14), 1823–1829. DOI: 10.1093/bioinformatics/bts252

Qi, X., Sun, Y., & Xiong, S. 2015. A single freeze-thawing cycle for highly efficient solubilization of inclusion body proteins and its refolding into bioactive form. *Microbial Cell Factories*, *14*(1), 24. DOI: 10.1186/s12934-015-0208-6

Qin, Z., Yan, Q., Lei, J., Yang, S., Jiang, Z., & Wu, S. 2015. The first crystal structure of a glycoside hydrolase family 17 β-1,3-glucanosyltransferase displays a unique catalytic cleft. *Acta Crystallographica. Section D, Biological Crystallography*, *71*(Pt 8), 1714–24. DOI: 10.1107/S1399004715011037

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., … Glöckner, F. O. 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), D590–D596. DOI: 10.1093/nar/gks1219

Read, S. M., Currie, G., & Bacic, A. 1996. Analysis of the structural heterogeneity of laminarin by electrospray-ionisation-mass spectrometry. *Carbohydrate Research*, *281*(2), 187–201. DOI: 10.1016/0008-6215(95)00350-9

Reese, E., Parrish, F., & Mandels, M. 1962. Beta-d-1,6-Glucanases in fungi. *Canadian Journal of Microbiology*, *8*(3), 327–34. DOI: 10.1139/m62-045

Reintjes, G., Arnosti, C., Fuchs, B. M., & Amann, R. 2017. An alternative polysaccharide uptake mechanism of marine bacteria. *The ISME Journal*, *11*(7), 1640–1650. DOI: 10.1038/ismej.2017.26

Repeta, D. J., & Aluwihare, L. I. 2006. Radiocarbon analysis of neutral sugars in high-molecular-weight dissolved organic carbon: Implications for organic carbon cycling. *Limnology and Oceanography*, *51*(2), 1045–53. DOI: 10.4319/lo.2006.51.2.1045

Rich, J. H., Ducklow, H. W., & Kirchman, D. L. 1996. Concentrations and uptake of neutral monosaccharides along 14°W in the equatorial Pacific: Contribution of glucose to heterotrophic bacterial activity and the DOM flux. *Limnology and Oceanography*, *41*(4), 595–604. DOI: 10.4319/lo.1996.41.4.0595

Robinson, C., Poulton, A. J., Holligan, P. M., Baker, A. R., Forster, G., Gist, N., … Zubkov, M. V. 2006. The Atlantic Meridional Transect (AMT) Programme: A contextual view 1995–2005. *Deep Sea Research Part II: Topical Studies in Oceanography*, *53*(14–16), 1485–1515. DOI: 10.1016/j.dsr2.2006.05.015

Romine, M. F. 2011. Genome-wide protein localization prediction strategies for gram negative bacteria. *BMC Genomics*, *12*(Suppl 1), S1. DOI: 10.1186/1471-2164-12-S1-S1

Rouwenhorst, R. J., Jzn, J. F., Scheffers, W. A., & van Dijken, J. P. 1991. Determination of protein concentration by total organic carbon analysis. *Journal of Biochemical and Biophysical Methods*, *22*(2), 119–28. DOI: 10.1016/0165-022X(91)90024-Q

Sabine, C. L. 2004. The Oceanic Sink for Anthropogenic CO2. *Science*, *305*(5682), 367–371. DOI: 10.1126/science.1097403

Sambrotto, R. N., Niebauer, H. J., Goering, J. J., & Iverson, R. L. 1986. Relationships among vertical mixing, nitrate uptake, and phytoplankton growth during the spring bloom in the southeast Bering Sea middle shelf. *Continental Shelf Research*, *5*(1–2), 161–198. DOI: 10.1016/0278-4343(86)90014-2

Sato, K., Naito, M., Yukitake, H., Hirakawa, H., Shoji, M., McBride, M. J., … Nakayama, K. 2010. A protein secretion system linked to bacteroidete gliding motility and pathogenesis. *Proceedings of the National Academy of Sciences*, *107*(1), 276–281. DOI: 10.1073/pnas.0912010107

Satyanarayana, U., & Chakrapani, U. 2013. *Biochemistry* (4th ed.).

Schmidt, J., Bischoff, A. A., WeiB, M., Kim, S. K., Frickenhaus, S., Slater, M. J., & Buck, B. H. 2017. Effect of Beta-1-3-Glucan and Mannans on Growth and Fitness of Starry Flounder (Platichthys Stellatus): A Potential New Candidate for Aquaculture in Temperate Regions. *Journal of FisheriesSciences.com*, *11*(3), 17–25. DOI: 10.21767/1307-234X.1000125

Schneider, T., Schmid, E., de Castro, J. V., Cardinale, M., Eberl, L., Grube, M., … Riedel, K. 2011. Structure and function of the symbiosis partners of the lung lichen (Lobaria pulmonaria L. Hoffm.) analyzed by metaproteomics. *Proteomics*, *11*(13), 2752–6. DOI: 10.1002/pmic.201000679

Senni, K., Pereira, J., Gueniche, F., Delbarre-Ladrat, C., Sinquin, C., Ratiskol, J., … Colliec-Jouault, S. 2011. Marine Polysaccharides: A Source of Bioactive Molecules for Cell Therapy and Tissue Engineering. *Marine Drugs*, *9*(12), 1664–1681. DOI: 10.3390/md9091664

Sharma, S., & Horn, S. J. 2016. Enzymatic saccharification of brown seaweed for production of fermentable sugars. *Bioresource Technology*, *213*, 155–161. DOI: 10.1016/j.biortech.2016.02.090

Shin, H., Oh, S., Kim, S., Won Kim, H., & Son, J. 2009. Conformational characteristics of β-glucan in laminarin probed by terahertz spectroscopy. *Applied Physics Letters*, *94*(11), 111911. DOI: 10.1063/1.3100778

Shin, J.-B., Krey, J. F., Hassan, A., Metlagel, Z., Tauscher, A. N., Pagana, J. M., … Barr-Gillespie, P. G. 2013. Molecular architecture of the chick vestibular hair bundle. *Nature Neuroscience*, *16*(3), 365–374. DOI: 10.1038/nn.3312

Sinnott, M. L., & Souchard, I. J. L. 1973. The mechanism of action of β-galactosidase. Effect of aglycone nature and α-deuterium substitution on the hydrolysis of aryl galactosides. *Biochemical Journal*, *133*(1), 89–98. DOI: 10.1042/bj1330089

Skoog, A., Biddanda, B., & Benner, R. 1999. Bacterial utilization of dissolved glucose in the upper water column of the Gulf of Mexico. *Limnology and Oceanography*, *44*(7), 1625–1633. DOI: 10.4319/lo.1999.44.7.1625

Slocombe, S. P., Ross, M., Thomas, N., McNeill, S., & Stanley, M. S. 2013. A rapid and general method for measurement of protein in micro-algal biomass. *Bioresource Technology*, *129*, 51–7. DOI: 10.1016/j.biortech.2012.10.163

Somville, M., & Billen, G. 1983. A method for determining exoproteolytic activity in natural waters. *Limnology and Oceanography*, *28*(1), 190–193. DOI: 10.4319/lo.1983.28.1.0190

Sonnenburg, E. D., Zheng, H., Joglekar, P., Higginbottom, S. K., Firbank, S. J., Bolam, D. N., & Sonnenburg, J. L. 2010. Specificity of Polysaccharide Use in Intestinal Bacteroides Species Determines Diet-Induced Microbiota Alterations. *Cell*, *141*(7), 1241–1252. DOI: 10.1016/j.cell.2010.05.005

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. DOI: 10.1093/bioinformatics/btu033

Størseth, T. R., Hansen, K., Reitan, K. I., & Skjermo, J. 2005. Structural characterization of β-D-(1→3)-glucans from different growth phases of the marine diatoms Chaetoceros mülleri and Thalassiosira weissflogii. *Carbohydrate Research*, *340*(6), 1159–1164. DOI: 10.1016/j.carres.2004.12.036

Størseth, T. R., Kirkvold, S., Skjermo, J., & Reitan, K. I. 2006. A branched beta-D-(1-->3,1-->6)-glucan from the marine diatom Chaetoceros debilis (Bacillariophyceae) characterized by NMR. *Carbohydrate Research*, *341*(12), 2108–14. DOI: 10.1016/j.carres.2006.05.005

Strom, S. L., Benner, R., Ziegler, S., & Dagg, M. J. 1997. Planktonic grazers are a potentially important source of marine dissolved organic carbon. *Limnology and Oceanography*, *42*(6), 1364–1374. DOI: 10.4319/lo.1997.42.6.1364

Studier, F. W. 2005. Protein production by auto-induction in high density shaking cultures. *Protein Expression and Purification*, *41*(1), 207–34. DOI: 10.1016/j.pep.2005.01.016

Suttle, C. A. 2005. Viruses in the sea. *Nature*, *437*(7057), 356–61. DOI: 10.1038/nature04160

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, *30*(12), 2725–2729. DOI: 10.1093/molbev/mst197

Tancula, E., Feldhaus, M. J., Bedzyk, L. A., & Salyers, A. A. 1992. Location and characterization of genes involved in binding of starch to the surface of Bacteroides thetaiotaomicron. *Journal of Bacteriology*, *174*(17), 5609–16. DOI: 10.1128/jb.174.17.5609-5616.1992

Tang, K., Jiao, N., Liu, K., Zhang, Y., & Li, S. 2012. Distribution and Functions of TonB-Dependent Transporters in Marine Bacteria and Environments: Implications for Dissolved Organic Matter Utilization. *PLoS ONE*, *7*(7), e41204. DOI: 10.1371/journal.pone.0041204

Teeling, H., Fuchs, B. M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C. M., … Amann, R. 2012. Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science (New York, N.Y.)*, *336*(6081), 608–11. DOI: 10.1126/science.1218344

Teeling, H., Fuchs, B. M., Bennke, C. M., Krüger, K., Chafee, M., Kappelmann, L., … Amann, R. 2016. Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. *eLife*, *5*(April 2016), 1–31. DOI: 10.7554/eLife.11888

Thomas, F., Lundqvist, L. C. E., Jam, M., Jeudy, A., Barbeyron, T., Sandström, C., … Czjzek, M. 2013. Comparative Characterization of Two Marine Alginate Lyases from Zobellia galactanivorans Reveals Distinct Modes of Action and Exquisite Adaptation to Their Natural Substrate. *Journal of Biological Chemistry*, *288*(32), 23021–23037. DOI: 10.1074/jbc.M113.467217

Thornton, D. C. O. 2014. Dissolved organic matter (DOM) release by phytoplankton in the contemporary and future ocean. *European Journal of Phycology*, *49*(1), 20–46. DOI: 10.1080/09670262.2013.875596

Tippenhauer, S., Torres-Valdes, S., Fong, A. A., Krauß, F., Huchler, M., & Wisotzki, A. 2017. Physical oceanography measured on water bottle samples during POLARSTERN cruise PS99.2 (ARK-XXX/1.2). PANGAEA. DOI: 10.1594/PANGAEA.871952

Turner, J. T. 2015. Zooplankton fecal pellets, marine snow, phytodetritus and the ocean's biological pump. *Progress in Oceanography*, *130*(4), 205–248. DOI: 10.1016/j.pocean.2014.08.005

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., … Cox, J. 2016. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods*, *13*(9), 731–740. DOI: 10.1038/nmeth.3901

Usui, T., Toriyama, T., & Mizuno, T. 1979. Structural Investigation of Laminaran of Eisenia bicyclis. *Agricultural and Biological Chemistry*, *43*(3), 603–11. DOI: 10.1080/00021369.1979.10863488

van Boekel, W., Hansen, F., Riegman, R., & Bak, R. 1992. Lysis-induced decline of a Phaeocystis spring bloom and coupling with the microbial foodweb. *Marine Ecology Progress Series*, *81*(7057), 269–276. DOI: 10.3354/meps081269

van Dongen-Vogels, V., Seymour, J. R., Middleton, J. F., Mitchell, J. G., & Seuront, L. 2012. Shifts in picophytoplankton community structure influenced by changing upwelling conditions. *Estuarine, Coastal and Shelf Science*, *109*, 81–90. DOI: 10.1016/j.ecss.2012.05.026

van Oijen, T., van Leeuwe, M. a., & Gieskes, W. W. C. 2003. Variation of particulate carbohydrate pools over time and depth in a diatom-dominated plankton community at the Antarctic Polar Front. *Polar Biology*, *26*(3), 195–201. DOI: 10.1007/s00300-002-0456-x

Varghese, J. N., Garrett, T. P., Colman, P. M., Chen, L., Høj, P. B., & Fincher, G. B. 1994. Three-dimensional structures of two plant beta-glucan endohydrolases with distinct substrate specificities. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(7), 2785–9. DOI: 10.1073/pnas.91.7.2785

Vårum, K. M., & Myklestad, S. 1984. Effects of light, salinity and nutrient limitation on the production of β-1,3-d-glucan and exo-d-glucanase activity in Skeletonema costatum (Grev.) Cleve. *Journal of Experimental Marine Biology and Ecology*, *83*(1), 13–25. DOI: 10.1016/0022-0981(84)90114-X

Verdugo, P., Alldredge, A. L., Azam, F., Kirchman, D. L., Passow, U., & Santschi, P. H. 2004. The oceanic gel phase: a bridge in the DOM–POM continuum. *Marine Chemistry*, *92*(1–4), 67–85. DOI: 10.1016/j.marchem.2004.06.017

Vidal-Melgosa, S., Pedersen, H. L., Schückel, J., Arnal, G., Dumon, C., Amby, D. B., … Willats, W. G. T. 2015. A new versatile microarray-based method for high throughput screening of carbohydrate-active enzymes. *The Journal of Biological Chemistry*, *290*(14), 9020–36. DOI: 10.1074/jbc.M114.630673

Vizcaíno, J. A., Côté, R. G., Csordas, A., Dianes, J. A., Fabregat, A., Foster, J. M., … Hermjakob, H. 2012. The Proteomics Identifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Research*, *41*(D1), D1063–D1069. DOI: 10.1093/nar/gks1262

Vocadlo, D. J., & Davies, G. 2008. Mechanistic insights into glycosidase chemistry. *Current Opinion in Chemical Biology*, *12*(5), 539–55. DOI: 10.1016/j.cbpa.2008.05.010

Vocadlo, D. J., & Withers, S. G. 2000. Glycosidase-Catalysed Oligosaccharide Synthesis. In *Carbohydrates in Chemistry and Biology* (Vol. 2, pp. 724–844). Weinheim, Germany: Wiley-VCH Verlag GmbH. DOI: 10.1002/9783527618255.ch29

Wallenfels, K. 1951. Enzymatische Synthese von Oligosacchariden aus Disacchariden. *Die Naturwissenschaften*, *38*(13), 306–307. DOI: 10.1007/BF00636782

Wang, D., Kim, D. H., Seo, N., Yun, E. J., An, H. J., Kim, J.-H., & Kim, K. H. 2016. A Novel Glycoside Hydrolase Family 5 β-1,3-1,6-Endoglucanase from Saccharophagus degradans 2-40T and Its Transglycosylase Activity. *Applied and Environmental Microbiology*, *82*(14), 4340–9. DOI: 10.1128/AEM.00635-16

Wang, F. K. 2001. Confidence interval for the mean of non-normal data. *Quality and Reliability Engineering International*, *17*(4), 257–267. DOI: 10.1002/qre.400

Wang, Q., Graham, R. W., Trimbur, D., Warren, R. A. J., & Withers, S. G. 1994. Changing Enzymic Reaction Mechanisms by Mutagenesis: Conversion of a Retaining Glucosidase to an Inverting Enzyme. *Journal of the American Chemical Society*, *116*(25), 11594–11595. DOI: 10.1021/ja00104a060

Ward, O. P., Moo-young, M., & Venkat, K. 1989. Enzymatic Degradation of Cell Wall and Related Plant Polysaccharides. *Critical Reviews in Biotechnology*, *8*(4), 237–74. DOI: 10.3109/07388558909148194

Wargacki, a. J., Leonard, E., Win, M. N., Regitsky, D. D., Santos, C. N. S., Kim, P. B., … Yoshikuni, Y. 2012. An Engineered Microbial Platform for Direct Biofuel Production from Brown Macroalgae. *Science*, *335*(6066), 308–313. DOI: 10.1126/science.1214547

Warren, R. 1996. Microbial hydrolysis of polysaccharides. *Annual Review of Microbiology*, *50*(1), 183–212. DOI: 10.1146/annurev.micro.50.1.183

Wei, N., Quarterman, J., & Jin, Y.-S. 2013. Marine macroalgae: an untapped resource for producing fuels and chemicals. *Trends in Biotechnology*, *31*(2), 70–77. DOI: 10.1016/j.tibtech.2012.10.009

Weiss, M., Abele, U., Weckesser, J., Welte, W., Schiltz, E., & Schulz, G. 1991. Molecular architecture and electrostatic properties of a bacterial porin. *Science*, *254*(5038), 1627–1630. DOI: 10.1126/science.1721242

Whitaker, J. R. 1974. Analytical Applications of Enzymes. In J. R. Whitaker (Ed.), *Adv. Chem. Ser* (Vol. 136, pp. 31–78). WASHINGTON, D. C.: American Chemical Society. DOI: 10.1021/ba-1974-0136.ch002

Wiltshire, K. H., & Dürselen, C.-D. 2004. Revision and quality analyses of the Helgoland Reede long-term phytoplankton data archive. *Helgoland Marine Research*, *58*(4), 252–268. DOI: 10.1007/s10152-004-0192-4

Wiltshire, K. H., Malzahn, A. M., Wirtz, K., Greve, W., Janisch, S., Mangelsdorf, P., … Boersma, M. 2008. Resilience of North Sea phytoplankton spring bloom dynamics: An analysis of long-term data at Helgoland Roads. *Limnology and Oceanography*, *53*(4), 1294–1302. DOI: 10.4319/lo.2008.53.4.1294

Woodward, J. R., & Fincher, G. B. 1982. Substrate specificities and kinetic properties of two (1→3), (1→4)-β-d-glucan endo-hydrolases from germinating barley (Hordeum vulgare). *Carbohydrate Research*, *106*(1), 111–22. DOI: 10.1016/S0008-6215(00)80737-5

Worden, A. Z., Follows, M. J., Giovannoni, S. J., Wilken, S., Zimmerman, A. E., & Keeling, P. J. 2015. Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes. *Science*, *347*(6223), 1257594–1257594. DOI: 10.1126/science.1257594

Wunsch, C. 2002. What Is the Thermohaline Circulation? *Science*, *298*(5596), 1179–1181. DOI: 10.1126/science.1079329

Xing, P., Hahnke, R. L., Unfried, F., Markert, S., Huang, S., Barbeyron, T., … Teeling, H. 2015. Niches of two polysaccharide-degrading Polaribacter isolates from the North Sea during a spring diatom bloom. *The ISME Journal*, *9*(6), 1410–1422. DOI: 10.1038/ismej.2014.225

Zaporozhets, T., & Besednova, N. 2016. Prospects for the therapeutic application of sulfated polysaccharides of brown algae in diseases of the cardiovascular system: review. *Pharmaceutical Biology*, *54*(12), 3126–3135. DOI: 10.1080/13880209.2016.1185444

Zechel, D. L., & Withers, S. G. 2000. Glycosidase mechanisms: anatomy of a finely tuned catalyst. *Accounts of Chemical Research*, *33*(1), 11–8.

Zha, X., Xiao, J., Zhang, H., Wang, J., Pan, L., Yang, X., & Luo, J. 2012. Polysaccharides in Laminaria japonica (LP): Extraction, physicochemical properties and their hypolipidemic activities in diet-induced mouse model of atherosclerosis. *Food Chemistry*, *134*(1), 244–252. DOI: 10.1016/j.foodchem.2012.02.129

Zhou, J., Bruns, M. A., & Tiedje, J. M. 1996. DNA recovery from soils of diverse composition. *Applied and Environmental Microbiology*, *62*(2), 316–22.

Ziervogel, K., & Arnosti, C. 2008. Polysaccharide hydrolysis in aggregates and free enzyme activity in aggregate-free seawater from the north-eastern Gulf of Mexico. *Environmental Microbiology*, *10*(2), 289–99. DOI: 10.1111/j.1462-2920.2007.01451.x

Ziervogel, K., & Arnosti, C. 2009. Enzyme activities in the Delaware Estuary affected by elevated suspended sediment load. *Estuarine, Coastal and Shelf Science*, *84*(2), 253–258. DOI: 10.1016/j.ecss.2009.06.022

Ziervogel, K., Karlsson, E., & Arnosti, C. 2007. Surface associations of enzymes and of organic matter: Consequences for hydrolytic activity and organic matter remineralization in marine systems. *Marine Chemistry*, *104*(3–4), 241–252. DOI: 10.1016/j.marchem.2006.12.001

Ziervogel, K., Steen, A. D., & Arnosti, C. 2010. Changes in the spectrum and rates of extracellular enzyme activities in seawater following aggregate formation. *Biogeosciences*, *7*(3), 1007–1015. DOI: 10.5194/bg-7-1007-2010

Zlotnik, I., & Dubinsky, Z. 1989. The effect of light and temperature on DOC excretion by phytoplankton. *Limnology and Oceanography*, *34*(5), 831–839. DOI: 10.4319/lo.1989.34.5.0831

ACKNOWLEDGEMENTS

First of all, I have to thank you Jan not only for giving me the opportunity to work on this fascinating project but also for placing your trust in me when you just started your group in an empty lab. I always enjoyed the great team spirit and atmosphere in our small group. Thank you for all the support and the open door at any times.

Many thanks to my thesis committee, Thorsten and Bernhard, for the guidance during the last three years and especially to Carol for reviewing my thesis and the fruitful discussions. I am also very grateful to Morten, who gave me the opportunity to take part on an expedition and provided me with many samples and lots of scientific experience. Thank you Kai-Uwe for your input and your participation in my defense committee.

Further, I would like to thank all the TAs both at MPI and also Hilke at AWI Helgoland. You all helped me a lot when we started setting up the lab and tried to organize our first sampling campaign. Thank you for being so patient and answering all my questions.

A huge thank you goes to all the people in my group. I always loved to be a part of this great team. Not only that you always help each other but for me it also felt much more like a group of friends as a group of colleagues. Thank you Agata, Andi, Nadine, Silvia, Tao and Tina for creating such a wonderful atmosphere in the lab. Thank you Jaagni for being the best student I could have had. I am especially grateful to Andi and Nadine for the proofreading of my thesis. Melissa and Craig, thank you for sharing your french-canadian(-islander) perspective as well as your crystallographic expertise with me, thanks for all the help during long nights at DESY and of course for all the pleasant Beer-Fridays.

Helga, I could not have had a better roomy than you. Thank you for the great time, for enduring the unbearable warmth in our office and for sharing your R knowledge with me. Stay independent! Katy and Josi, you made my entire time in Bremen and also on Polarstern really special. Thank you for being such great friends.

Ein besonderer Dank gilt meiner Familie. Meinen Eltern, Karin und Reinhard, sowie meiner Schwester Svenja. Habt Dank für euren bedingungslosen Rückhalt und das Vertrauen in mich über all die Jahre. Ohne euch wäre das alles so nicht möglich gewesen. Mona, dir danke ich für deine Nähe, dein Vertrauen und dafür, dass du deine Geduld mit mir nicht verloren hast.

APPENDIX

Additional co-author contributions

High Arctic kelps maintain their photosynthetic functions throughout the polar night

Lydia Scheschonk[1], Stefan Becker[2,3], Jan-Hendrik Hehemann[2,3], Kai Bischof[1]

[1] Marine Botany at the University of Bremen
[2] MARUM – Center for Marine Environmental Sciences at the University of Bremen, Germany
[3] Max-Planck-Institute for Marine Microbiology, Bremen, Germany

In preparation.

Abstract

Kelps, seaweeds of the order Laminariales, are important ecosystem engineers in arctic coastal ecosystems. With respect to this, increasing global temperatures, which have been observed to be most severe in the wintertime Arctic, necessitate to assess the capacity of adaptive life strategies of polar kelps during polar night. However, data on seaweed ecosystem functioning under polar night conditions is scarcely available.

We assessed several physiological parameters (photosynthesis, pigment content, respiration, and carbohydrate storage) in two species of arctic kelps, the boreal-temperate *Saccharina latissima* and the arctic-endemic *Laminaria solidungula*, during the period of polar night 2016/17. Plants were sampled from Kongsfjorden, Svalbard, at 78° 55' N, shortly before the onset of the dark period in October, as well as towards the end of polar night in early February. Analyses were carried out for different sections along the phylloid (Meristem, Centre, Distal Part).

Our data suggest that kelps do largely maintain their photosynthetic functions throughout the entire winter period, as indicated by PI-curves, and matching Chl a and antenna pigment contents. The maintenance of the photosynthetic functions is most likely fueled by the storage carbohydrate laminaran. Overall laminaran content was reduced by ~96 % in *S. latissima*, and by ~90 % in *L. solidungula* by the end of the winter. However, strong differences were observed between the different phylloid regions across species, indicating specific adaptive strategies between boreal-temperate and arctic-endemic species. The data presented here form the baseline for subsequent studies investigating the effects of further increases in temperature on arctic seaweeds during the polar night.

Versicherung an Eides Statt / *Affirmation in lieu of an oath*

gem. § 5 Abs. 5 der Promotionsordnung vom 15.07.2015 /
according to § 5 (5) of the Doctoral Degree Rules and Regulations of 15 July, 2015

Ich / I, Stefan Becker, Große Annenstraße 96, 28199, Bremen
(Vorname / First *Name*, Name / *Name*, Anschrift / *Address*, ggf. Matr.-Nr. / *student ID no., if applicable)*

versichere an Eides Statt durch meine Unterschrift, dass ich die vorliegende Dissertation selbständig und ohne fremde Hilfe angefertigt und alle Stellen, die ich wörtlich dem Sinne nach aus Veröffentlichungen entnommen habe, als solche kenntlich gemacht habe, mich auch keiner anderen als der angegebenen Literatur oder sonstiger Hilfsmittel bedient habe und die zu Prüfungszwecken beigelegte elektronische Version (PDF) der Dissertation mit der abgegebenen gedruckten Version identisch ist. / *With my signature I affirm in lieu of an oath that I prepared the submitted dissertation independently and without illicit assistance from third parties, that I appropriately referenced any text or content from other sources, that I used only literature and resources listed in the dissertation, and that the electronic (PDF) and printed versions of the dissertation are identical.*

Ich versichere an Eides Statt, dass ich die vorgenannten Angaben nach bestem Wissen und Gewissen gemacht habe und dass die Angaben der Wahrheit entsprechen und ich nichts verschwiegen habe*. / I affirm in lieu of an oath that the information provided herein to the best of my knowledge is true and complete.*

Die Strafbarkeit einer falschen eidesstattlichen Versicherung ist mir bekannt, namentlich die Strafandrohung gemäß § 156 StGB bis zu drei Jahren Freiheitsstrafe oder Geldstrafe bei vorsätzlicher Begehung der Tat bzw. gemäß § 161 Abs. 1 StGB bis zu einem Jahr Freiheitsstrafe oder Geldstrafe bei fahrlässiger Begehung. / *I am aware that a false affidavit is a criminal offence which is punishable by law in accordance with § 156 of the German Criminal Code (StGB) with up to three years imprisonment or a fine in case of intention, or in accordance with § 161 (1) of the German Criminal Code with up to one year imprisonment or a fine in case of negligence.*

Bremen, 29.03.2018
Ort / *Place*, Datum / *Date*

Unterschrift / *Signature*