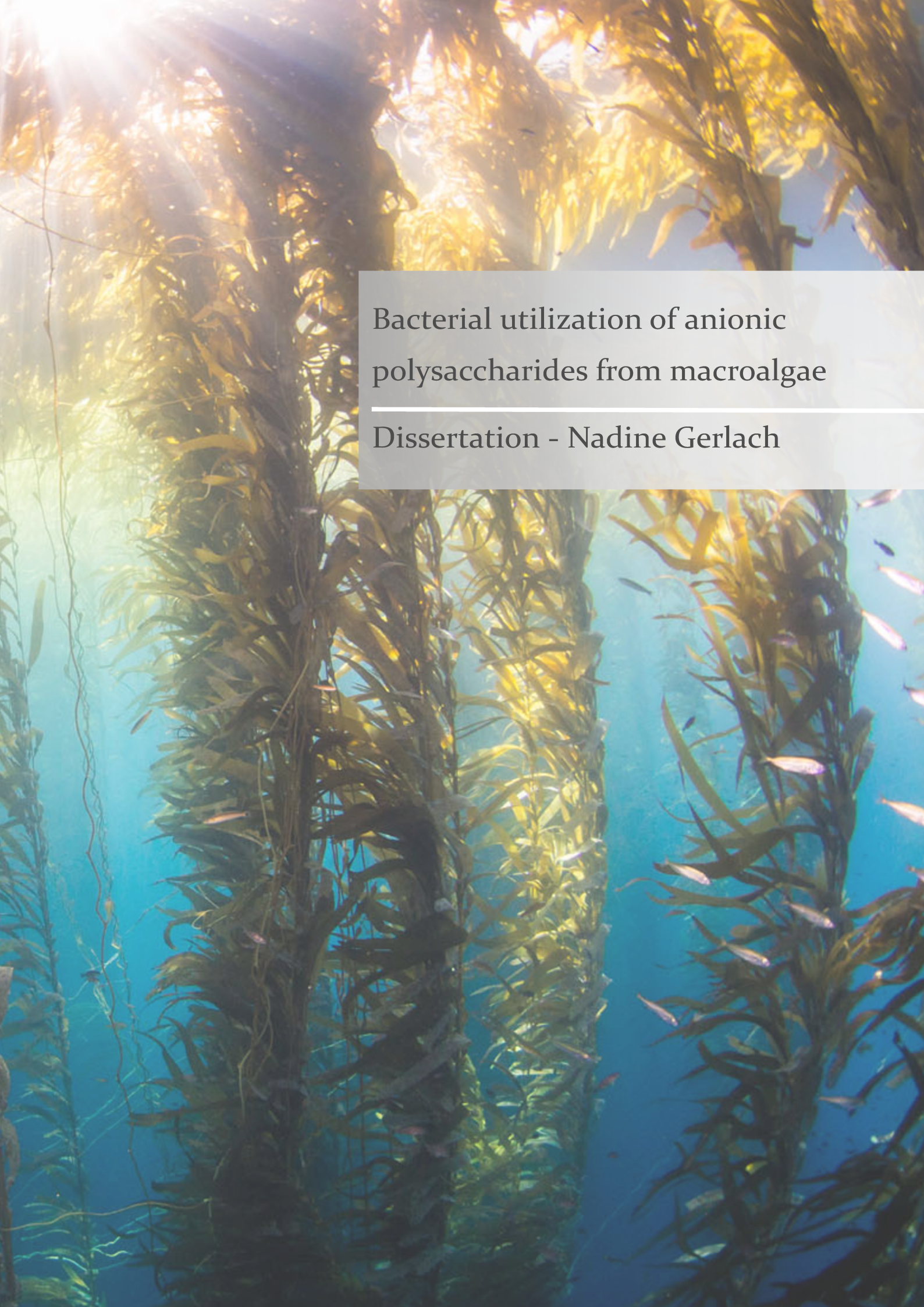


Bacterial utilization of anionic polysaccharides from macroalgae

Dissertation - Nadine Gerlach

Bacterial utilization of anionic polysaccharides from macroalgae

Dissertation

Zur Erlangung des Doktorgrades der Naturwissenschaften

– Dr.rer.nat –

Im Fachbereich der Geowissenschaften der

Vorgelegt von Nadine Gerlach

Bremen, Juli 2020

The presented doctoral thesis© was conducted in the department Marine Glycobiology at the Max-Planck-Institute for Marine Microbiology and the MARUM – Center for Marine Environmental Sciences at the University of Bremen from July 2017 to July 2020. The thesis was constructed as a draft version of scientific papers that base on the present study as well as additional background information, general discussion and outlook section for further research.

Autor: M.Sc. **Nadine Gerlach**

1. Reviewer: Dr. Jan-Hendrik Hehemann

Department of Marine Glycobiology at the Max-Planck-Institute for Marine Microbiology and the MARUM – Center for Marine Environmental Sciences at the University of Bremen

2. Reviewer: Prof. Dr. Anne S. Meyer

Department of Biotechnology and Biomedicine at the Technical University of Denmark

Promotion colloquium: October 8th 2020

Every time I slip into the ocean, it is like going home.

Silvia Alice Earl – one of the most inspiring pioneers in marine biology, oceanography and conversation. Thank you for being such an amazing role model.

Table of content

Acknowledgments

First things first – this thesis would not have been possible without the outstanding support of so many people!

Foremost, I must thank Dr. Mathias Wietz who supervised my master thesis and without him, I probably would not have met Dr. Jan-Hendrik Hehemann.

Jan-Hendrik, if we had met at the beginning of my studies, I would have bet never ending up in biochemistry for sure. I have never learnt as much as in the time in your lab. You are one of the most enthusiastic and curious scientists I have met so far.

Many thanks to my thesis and defense committee Prof. Dr. Anne Meyer, Dr. Bernhard Fuchs and Dr. Mirjam Czjzek for helpful advices and comments.

I also thank the whole department of Marine Glycobiology for their warm welcome and the nice working atmosphere – in particular: Mandy Knutzen, Alek Bolte and Tina Trautmann for lab assistance in any way, my office buddy Andi, Agata for German coffee breaks, Mikkel for beer Fridays as well as Jaagni, Tao, Silvia and Hagen for many laughter.

Many thanks go to GLOMAR – Bremen International Graduate School for Marine Sciences for financial support and my collaborators – especially Tatjana von Rosen, Dr. Craig S. Robb, Dr. Sylvain Engilberge and Dr. Tristan Wagner.

Thank you, team of SciHub – a constant source for knowledge!

Lastly and most importantly, I am more than thankful for the never-ending encouragement and love of my family and my friends. Djoana, Svenja, Viola, Diandra, Stefanie and Julia – Mädels, ihr seid meine Konstante und meine Familie. Ich wüsste nicht, wer ich ohne euch wäre. Auf die nächsten schönen Jahrzehnte mit euch!

Stefan, du warst meine Bezugsperson in meinem PhD. Deine Ermutigungen und Unterstützung haben mir oft den Tag mehr als nur erleichtert. Ich bin gespannt auf die kommende Zeit ohne Uni und Doktorarbeit.

List of abbreviations

Δ	hexenuronic acid residue
ABC	ATP-binding cassette
Ac	acetate
AGE	affinity gel electrophoresis
Aly	alginate lyase
AMAC	2-aminoacridone
ASU	asymmetric unit
BLAST	basic local alignment search tool
BSA	bovine serum albumin
CAPSO	N-cyclohexyl-2-hydroxyl-3-aminopropanesulfonic acid
CAZyme	carbohydrate active enzyme
CBM	carbohydrate binding module
CE	carbohydrate esterase
cNP	*2*-Chloro-4-nitrophenyl
C-PAGE	carbohydrate-polyacrylamide gel electrophoresis
DAPI	4′,6-Diamidin-2-phenylindol
DEH	4-deoxy- L-erythro-5-hexoseulose
DLS	dynamic light scattering
DNA	deoxyribonucleic acid
DOC	dissolved organic carbon
DOM	dissolved organic matter
DIC	dissolved inorganic carbon
EC no.	enzyme commission number
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme linked immunosorbent assay
FACE	fluorophore-assisted carbohydrate electrophoresis
FCSP	fucose-containing sulfated polysaccharides
FGE	formylglycine-generating enzyme
fGly	formylglycine
G	guluronate
GH	glycoside hydrolase
GM	guluronate and mannuronate
HPAEC-PAD	high performance anion exchange chromatography - pulsed amperometric detection
HEPES	4-(2-hydroyethyl)-1-piperazineethanesulfonic acid

IPTG	isopropyl β-D-1-thiogalactopyranoside
IMAC	immobilized metal affinity chromatography
KDG	2-dehydro-3-deoxy-d-gluconate
LB	lysogeny broth
M	mannuronate
MES	2-(N-morpholino)ethane sulfonic acid
MS	mass spectrometry
$NAD(P)^+$	nicotinamide adenine dinucleotide (phosphate), oxidized
NAD(P)H	nicotinamide adenine dinucleotide (phosphate), reduced
nc	non-classified
NMR	nuclear magnetic resonance
OM	organic matter
PAHBAH	4-hydroxybenzoic acid hydrazide
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
PIC	particulate inorganic carbon
PL	polysaccharide lyase
pNP	4-nitrophenol
POC	particulate organic carbon
POM	particulate organic matter
PorSS	por secretion system
rsmd	root-square-mean-deviation
PUL	polysaccharide utilization loci (or locus)
S	sulfate, sulfatase
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
Sus	starch utilization system
SBS	secondary binding site
T9SS	type IX secretion system
TEP	transparent exopolymer particles
TGS	Tris-glycine-SDS
TLC	thin layer chromatography
TRIS	tris(hydroxymethyl)aminomethane

Numeric values and proteogenic amino acids (aa) are given in SI units as well as the one and three letter codes. Polysaccharides are drawn according to the Symbol Nomenclature of Glycans (SNFG).

Author's contributions

Publications

1. Reisky L, Préchoux, A, Zühlke M, Bäumgen M, Robb, CS, **Gerlach N**, *et al.* A marine bacterial enzymatic cascade degrades the algal polysaccharide ulvan. *Nature Chemical Biology* **15,** 803–812 (2019). https://doi.org/10.1038/s41589-019-0311-9

Author contributions:
J.-H.H., T.S., G.M. and U.T.B. initiated the study and directed the project. L.R., A.P., R.L. and M.B. cloned the genes and expressed and purified the enzymes for the degradation reactions. M.B., J.-H.H. and L.R. isolated ulvan and purified oligomers. Metabolites were analyzed by C.S. via NMR and HPLC–ELS–MS for which M.D.M. provided resources. L.R. and M.B. performed biocatalyses for the analyses in gel-based assays. A.P. together with M.B. performed HPAEC–PAD analyses. M.-K.Z. with support from S.M., F.U. and A.T.-S. performed the proteome analyses for which D.B. provided the resources. N.G., C.S.R. and T.R. performed crystallographic experiments and solved the protein structures. G.M. analyzed the crystal structure of the l-rhamnose mutarotase and of the sulfatases. S.T. performed the computational analyses of PUL predictions. J.-H.H. and L.R. wrote the paper with input from U.T.B., G.M., S.M., M.-K.Z. and T.S. All authors read and approved the final manuscript.

Contribution of NG:
Data acquisition and analyses: protein expression, purification, crystallization and X-ray data collection at the synchrotron (P10_PLnc, P17_GH2, P29_PDnc, P30_PL28, P31_GH39, P32_S1_8, P33_GH105, P36_S1_25), structure determination (P32_S1_8, P17_GH2).
Writing: methods - protein crystallization and X-ray data collection at the synchrotron of P17_GH2.

Manuscripts

1. **Gerlach N**, Vanni C, Hehemann JH. Novel carbohydrate-binding module from *Formosa agariphila* KMM3901^T binds to macroalgal polysaccharides. To be submitted to Protein Science.

Author contributions: contribution of NG (%) to the current state of the manuscript
JHH and NG (50 %) designed the study. NG cloned, expressed, purified and crystallized the enzyme. NG collected X-ray diffraction data, performed binding studies and structural analyses. NG produced and purified the enzyme ligand. NG (80 %) and CV performed metagenomic analyses. NG (95 %) and CV visualized the data. NG wrote the initial manuscript.

2. **Gerlach N**, Sichert A, Hehemann JH. Crystal structure of a putative marine bacterial sulfatase of family S1_15. To be submitted to Protein Science.

Author contribution: contribution of NG (%) to the current state of the manuscript
NG (50 %), AS and JHH designed the study. AS designed primer and performed phylogenetic analysis. NG cloned, expressed, purified and crystallized the enzyme. NG collected X-ray diffraction data, conducted biochemical and structural analyses. NG (95 %) and AS visualized the data. NG (80 %), AS and JHH wrote the manuscript.

3. **Gerlach N**, Hehemann JH. Marine bacterial fucosidase of family GH95 degrades macroalgal fucoidans. Submitted to Applied Microbiology and Biotechnology.

Author contribution: contribution of NG (%) to the current state of the manuscript
JHH and NG (50 %) designed the study. NG cloned, expressed and purified the enzyme. NG conducted biochemical experiments, analyzed and visualized the data. NG (80 %) and JHH wrote the manuscript.

4. **Gerlach N**, Sichert A, Keegan, R., Patharkar, A, Von Rosen T, Robb CS, Hehemann JH. Biochemical and structural comparison of marine fucosidases of family GH29. In preparation for The Journal of Biological Chemistry.

Author contribution: contribution of NG (%) to the current state of the manuscript
NG (50 %), AS and JHH designed the study. NG (50 %), AS and TvR cloned and expressed the enzymes. NG (80 %) and AP performed biochemical analyses. NG (95 %) and TvR crystallized the enzymes. NG (95 %) and CSR collected X-ray diffraction data. NG (70 %), RK and CSR determined the protein structures. AS conducted phylogenetic analysis. NG designed mutants. NG (95 %) and AS visualized the data. NG wrote the initial manuscript.

Poster and oral presentations

1. Poster presentation at Summer Course Glycoscience (2018) in Wageningen, The Netherlands

2. Poster presentation at MIMAS II conference (2018) in Greifswald, Germany

3. Oral presentation at YOUMARES 9 conference (2018) in Oldenburg, Germany

4. Oral presentation and organization of PhD Days in Marine Science (2019) in Bremen, Germany

5. Poster presentation and session host at ICYMARE conference (2019) in Bremen, Germany

Reviews and reports

1. The CAZypedia Consortium. Ten years of CAZypedia: a living encyclopedia of carbohydrate-active enzymes. *Glycobiology*, Volume 28, Issue 1, January 2018, Pages 3–8, https://doi.org/10.1093/glycob/cwx089

Contribution to the online encyclopedia: Polysaccharide Lyase Family 7
https://www.cazypedia.org/index.php/Polysaccharide_Lyase_Family_7

2. ICYMARE 2019 conference proceedings

In the chapter *How microbes affect the marine carbon cycle – the role of bacteria in degrading algae-derived organic matter*, I summarize state-of-the-art knowledge of microbial communities influencing the marine carbon cycle, focusing on heterotrophic bacteria in the water column utilizing algal polysaccharides. This manuscript has been peer reviewed by two anonymous reviewers and will be published in Springer Link in 2020.

3. Schewe, I. (2018): The Expedition PS107 of the Research Vessel POLARSTERN to the Fram Strait and the AWI-HAUSGARTEN in 2017, Berichte zur Polar- und Meeresforschung = Reports on polar and marine research, Bremerhaven, Alfred Wegener Institute for Polar and Marine Research, 717, 120 p.. https://doi.org/10.2312/BzPM_0717_2018

Contribution: Chapter 8 – Functional glycomics of algal blooms in the arctic ocean

4. Safety evaluation for genetic engineering work at MPI Bremen – Algal polysaccharide degrading enzymes and associated proteins from marine bacteria

Contribution: NG wrote the safety evaluation with inputs of JHH and Anke Meyerdierks

Proposals for macromolecular crystallographic beamtime

1. Structural analysis of bacterial CAZymes and associated enzymes utilizing algal polysaccharides

Nadine Gerlach[1,2], Jan-Hendrik Hehemann[1,2]

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

Synchrotron: EMBL – European Molecular Biology Laboratory, Hamburg, Germany
Proposal ID: MX-686 at EMBL
Submitted: 16th December 2018

2. Structural analysis of bacterial CAZymes and associated enzymes utilizing algal polysaccharides

Nadine Gerlach[1,2], Jan-Hendrik Hehemann[1,2]

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

Synchrotron: DESY – Deutsches Elektronen Synchrotron, Hamburg, Germany

Proposal ID: I-20180855
Submitted: 31st August 2018

Proposal ID: I-20190288 (Continuation of Proposal ID: I-20180855)
Submitted: 28th February 2019

Proposal ID: I-20191001 (Continuation of Proposal ID: I-20190288)
Submitted: 30th August 2019

Proposal ID: I-20200368 (Continuation of Proposal ID: I-20191001)
Submitted: 28th February 2020

Summary

About two thirds of Earth is covered by oceans of which less than 2 % constitute marine vegetation. Yet, marine vegetation serves crucial ecosystem functioning and is especially vital to the health of coastal ecosystems. Macroalgae display one group of marine vegetation which can form ecological habitats that are not only impressive by their dimensions but also with regards to biodiversity and primary production. Among algal compounds, polysaccharides play an important role as bioactive polymers and food source for heterotrophic organisms. Especially anionic polysaccharides are of high scientific interest as they define the 3D structure of the algae and sequester carbon in form of sinking particles from the photic zone to the deep sea. Detailed insights into the enzymatic machinery of heterotrophic bacteria capable to utilize algal polysaccharides contributes to a better understanding of carbon cycling and food web dynamics in the marine environment. In the present dissertation, I investigated different aspects of bacterial utilization of three anionic polysaccharides from macroalgae – ulvan, alginate and fucoidan.

For the first project, I determined the protein structures of ulvan-degrading enzymes from *Formosa agariphila* KMM3901^T, an algae-associated bacteria of the phylum Bacteroidetes. Bacteroidetes are globally abundant and efficient carbohydrate utilizers and therefore, key degraders of algae-derived organic matter. The combination of biochemical, cultivational and proteomic approaches alongside with the novel protein structures revealed molecular insights into ulvan utilization. We solved the structures of an endo-acting GH39 with a novel activity towards α-1,4-L-rhamnose, an exo-acting β-glucuronidase of family GH2, and an arylsulfatase of subfamily S1_8 which removes the sulfate group of an ulvan disaccharide.

In the second project, I combined structural and biochemical characterization of fucosidases to elucidate fucoidan utilization by the remarkable carbohydrate-degrader '*Lentimonas'* sp. CC4. Fucoidan is a compositional and structural versatile class of fucose-containing sulfated polysaccharides. This complex class of glycans has been considered more resistant to microbial degradation than others, and thereby may act as potential carbon sink.

The research in the present dissertation delivered molecular insights into the function and structure of glycan-degrading enzymes. These findings will contribute to a greater understanding of the utilization of anionic polysaccharides by marine heterotrophic bacteria. Consequently, using bacterial enzymes as analytical tools to investigate the turnover rates of anionic polysaccharides will elucidate their fate in the marine carbon cycle.

Zusammenfassung

Etwa zwei Drittel der Erde sind von Ozeanen bedeckt, von denen Vegetation lediglich 2 % ausmacht. Nichtsdestotrotz ist diese unerlässlich für die Funktionsweise von marinen Ökosystemen und vor allem der Integrität von Küstenhabitaten. Die Gruppe der Makroalgen zeichnen sich durch das Ausbilden großer Habitate aus, die sowohl mit einer hohen Biodiversität als auch intensiver Primärproduktion einher gehen. Unter den Bestandteilen der Algen nehmen die Polysacharide der Makroalgen als bioaktive Polymere and Nahrungsquelle für heterotrophe Organismen eine essenzielle Rolle ein. Insbesondere anionische Zucker sind dabei von hohem wissenschaftlichem Interesse, da sie nicht nur die 3D-Struktur der Makroalgen definieren, sondern auch die Sedimentierung von organischer Materie beeinflussen indem sie Partikel bilden, welche bis in die Tiefsee absinken.

Detaillierte Einblicke in die enzymatischen Abbaumechanismen heterotropher Bakterien tragen zu einem besseren Verständnis von der Dynamik der Nahrungsnetzte und dem Kohlenstoffkreislauf im Meer bei. In der vorliegenden Dissertation habe ich verschiedene Aspekte der bakteriellen Verwertung der folgenden anionischen Polysaccharide untersucht: Ulvan, Alginat und Fucoidan.

Im ersten Projekt habe ich die Proteinstrukturen von Ulvan-abbauenden Enzymen von *Formosa agariphila* KMM3901^{T}, einem algenassoziierten Bakterium des Stammes Bacteroidetes, bestimmt. Bacteroideten sind weltweit vorkommende, effiziente Kohlenhydratabbauer. Folglich spielen sie eine Schlüsselrolle im Abbau der aus Algen stammenden organischen Materie. Die Kombination von Kultivierungsansätzen, Proteomik und biochemischer Charakterisierung von Enzymen zusammen mit den Proteinstrukturen ergab neuartige Einblicke in den Abbauprozess von Ulvan auf molekularer Ebene. Wir haben die Strukturen einer neuen endo-α-1,4-L-Rhamnosidase der Familie GH39, einer exo-β-Glucuronidase der Familie GH2 und einer Arylsulfatase der Unterfamilie S1_8, welche die Sulfatgruppe eines Ulvandissachrids entfernt, gelöst.

Im zweiten Kapitel befasste ich mich mit dem Abbau von Fucoidan von '*Lentimonas*' sp. CC4 mit dem Fokus auf der biochemischen und strukturellen Charakterisierung von Fucosidasen. Fucoidane bilden eine komplexe Klasse von Fucose-haltigen sulfatisierten Polysacchariden, die aufgrund ihrer strukturellen Komplexität eine höhere Resistenz gegenüber mikrobieller Verwertung aufweist und somit potenziell als Kohlenstoffsenke dient.

Die Forschung dieser Dissertation lieferte neue Erkenntnisse in die molekulare Funktionsweise und 3D-Struktur von Polysaccharid-abbauenden Enzymen. Somit tragen diese Ergebnisse zu einem besseren Verständnis der Polysaccharidnutzung durch heterotrophe Bakterien bei. Die Anwendung dieser bakteriellen Enzyme als Werkzeuge zur Bestimmung der Umsatzraten kann zukünftig dazu beitragen, die Rolle von Algenzuckern im marinen Kohlenstoffkreislauf aufzuschlüsseln.

Introduction

The impact of microbes on the marine carbon cycle

Since the industrial revolution, humans have introduced more greenhouse gases into the atmosphere than ever. These increasing numbers mainly derive from fossil fuel burning and are of great concern due to their effect on weather and climate (IPCC report 2013) but also in global biochemical cycles. The oceans play a crucial role in the global carbon cycle. They are one of the main reservoirs and the most important sink for the greenhouse gas carbon dioxide (CO_2), holding 50 times more carbon than the atmosphere (Falkowski et al. 2000).

Carbon is an important component of both oceanic inorganic and organic matter (OM), which is not exclusively of marine origin. Terrestrial inputs from rivers in coastal areas (Blair and Aller 2012), and windblown dust (Mahowald et al. 2005) affect food web dynamics and marine communities which in consequence complicate investigations on the composition and fate of marine organic matter. Both organic and inorganic carbon occurs in its dissolved or particulate form. Dissolved inorganic carbon (DIC) – including carbonate (CO_3^{2-}), bicarbonate (HCO_3^-), carbonic acid H_2CO_3 and carbon dioxide (CO_2) – precipitate into particulate inorganic carbon (PIC) in the form of $CaCO_3$ via both abiotic and biotic processes. Vice versa, PIC can be retransferred into DIC via dissolution or can be converted to particulate organic carbon (POC) via chemoautotrophy or photosynthesis by primary producers. POC becomes DOC due to abiotic and biotic fragmentation and disaggregation, whereas PIC is produced from POC only via consumption and respiration by heterotrophs (Wallmann and Aloisi 1992; Emerson and Hedges 2008; Ciais et al. 2013). Spontaneous assembly of DOC can result in POC (Chin et al. 2003), while UV exposure leads to DOC (e.g. Mopper and Kieber 2002; Cory et al. 2015; Shen and Benner 2018) but also the formation of DIC (Porcal et al. 2015).

Two main processes are driving the flux of carbon from the atmosphere to the oceans – the biological and the physical carbon pump, consisting of the solubility and the carbonate pump. The biggest oceanic carbon pool is the DIC, which main source is the abiotic dissolution of atmospheric CO_2 into the surface water. The transformation of CO_2 into its aqueous forms carbonic acid determines seawater's pH. The counter process to the so-called solubility pump is the carbonate pump, which describes the biotic conversion of DIC to PIC. Marine organisms such as corals and coccolithophores – a globally distributed eukaryotic phytoplankton group – produce calcium carbonate, which is the main compound of their skeletons and shells, respectively (Wallmann and Aloisi 1992; Emerson and Hedges 2008; Ciais et al. 2013).

In the oceans, inorganic carbon is mainly fixed as CO_2 by phytoplankton. Phytoplankton are photoautotrophic floating eukaryotic (e.g. diatoms) and bacterial (e.g. cyanobacteria) organisms

that are unable to actively swim against a water current. Phytoplankton constitutes the basis of the marine food web. Like plants, they convert CO_2 into oxygen and biomass, which serves as carbon and energy source for all trophic levels (Fig. 1). About one-half of the global net primary production (NPP) is performed by marine microalgae (Field et al. 1998), yet they only account for ~ 1% of the global plant biomass (Behrenfeld 2014). The turnover rate of organic matter in the oceans, especially in the photic zone (first ~ 200 m) is remarkable. About 60-99 % of all assimilated carbon is directly used by marine bacteria (Fuhrman and Azam, 1982; Azam and Malfatti, 2007 and references therein), while the rest is either processed by other organisms of the food web or is exported into the deep sea via the biological pump (Buchan et al. 2014). Marine bacteria transform photosynthetically derived organic matter so that it becomes again accessible to other organisms. This process is known as the microbial loop (Azam et al. 1983). The most frequent and best studied bacteria involved in this process belong to the classes of Flavobacteriia, Alphaproteobacteria and Gammaproteobacteria (Buchan et al. 2014 and references therein). But also auto-catalyzed cell death (Berges and Falkowski 1998), weathering, viral infection (Suttle et al. 1990) and sloppy feeding of zooplankton contribute to the two main carbon pools – dissolved (DOM) and particulate organic matter (POM).

The fate of DOM and POM is currently intensively studied. Global assessment of the turnover rates and the degradation of both OM pools help us to identify sinks and sources for carbon as well as to gain a better understanding of biochemical cycles and food web dynamics. Certain fractions of POM and DOM are more resistant to microbial degradation than others, leading to the transport of organic matter into the deep sea (Fig. 1). These processes are known as biological and microbial carbon pump (Jiao et al. 2010; Jiao and Zheng 2011; Jiao et al. 2014; Jiao et al. 2018). The long-term storage of OM of up to thousands of years in the deep sea has an important effect on food web dynamics. Nevertheless, microbial degradation of DOM and POM by free-living and particle-associated bacteria is taking place throughout the water column and even in deep sea sediments (Buchan et al. 2014 and references therein). Recent studies suggest that the fate of organic matter might be determined due to inorganic matrix shielding these potential substrates from further degradation (Hedges et al. 2001). However, the oceans are a constantly changing environment. Water masses in the deep sea are transported around the globe. OM from one location gets exposed to a different microbial community in another location which might possess a different strategy capable to utilize these substances (Shen and Benner 2018).

The fate of organic matter and the role of microbes is still not completely understood, but what seems clear is that the circulation of water masses and the metabolic diversity of bacteria influence the type of molecule that is consumed but also its lifetime in the marine environment (Shen and Benner 2018).

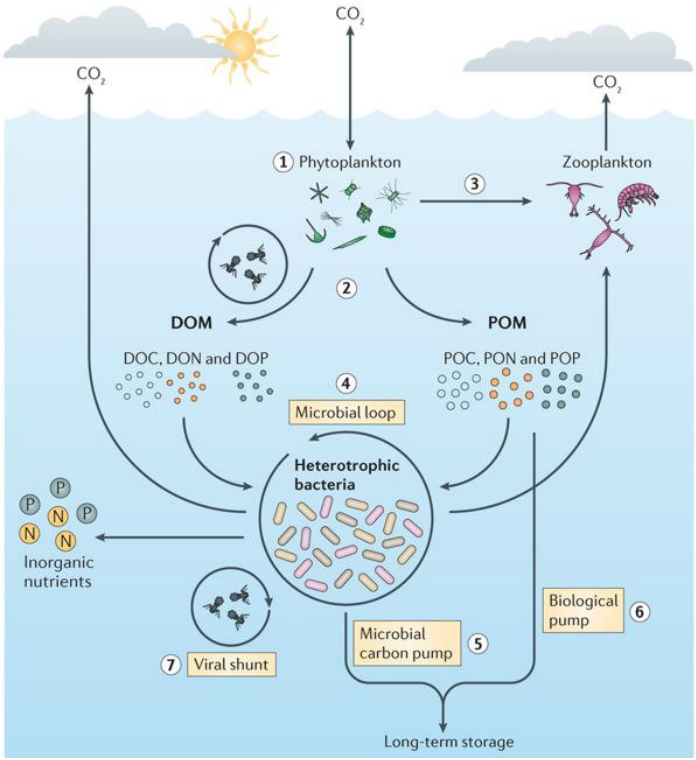


Figure 1. The marine carbon cycle (Buchan et al. 2014). (1) Inorganic carbon is fixed by phytoplankton and (2) converted into organic material, which can either be secreted as dissolved organic matter (DOM) or released as particulate organic matter (POM) due to weathering and (3) grazing by zooplankton. (4) Heterotrophic bacteria remineralize both carbon pools, (5) while a certain proportion cannot be further catabolized and becomes recalcitrant (RDOC). (6) Some compounds of the phytoplankton-derived DOM can aggregate and become POM, which is extensively colonized by microbes while sinking to the deep sea (defined as >1,000 m depth). (7) Viruses affect the release of POM and DOM from the phytoplankton and bacterial community due to cell lysis.

Ecological relevance of macroalgae blooms

"Every second breath one takes comes from the oceans[1]"

"Algae contribute to half of the global carbon fixation[2]"

"Algae as primary producers are the foundation of aquatic ecosystems[3]"

"Algae blooms trigger a subsequently succession of bacterioplankton species[4]"

The above-mentioned statements mainly refer to microalgae or phytoplankton, but what is the ecological relevance of macroalgae and how do they effect marine ecosystems and food web dyna-

[1] Data from: Emerson *et al*., 2008 and 2010 in Global Biogeochemical Cycles and Deep-Sea Research I
[2] Data from: Field *et al*., 1998 in Science
[3] Quay *et al*., 2008 in Global Biogeochemical Cycles
[4] Data from: Teeling *et al*., 2012 in Science Reports

mics? To what extent are macroalgae contributing to the marine carbon cycle?

About two thirds of Earth is covered by oceans of which less than 2 % are covered by marine vegetation (Duarte and Cebrián 1996). Marine vegetation is comprised of seagrasses, mangroves, salt marshes and algae. Despite their name, seagrasses are the only marine representatives of flowering plants (angiosperms). In contrast, algae do not resemble vascular plants but belong to the protists (Macreadie et al. 2017). Algae can be found all over the world and are classified into two major groups – macroalgae or seaweed and microalgae or phytoplankton (El Gamal 2010). Macroalgae are multicellular or siphonocladous protistan that are divided into three classes – green algae (Chlorophyta), brown algae (Phaeophyta) and red algae (Rhodophyta). Today there are more than 17000 species within almost 1500 genera known. Macroalgae are more dominant and taxonomically diverse in tropical regions but can be found in all latitudes including the polar regions, where vertical distribution is limited by light penetration (Garbary et al. 2001; Macreadie et al. 2017).

Contrary to microalgae, the impact of macroalgae on the marine carbon cycle is often not considered as they make less contribution to primary production and respiration (Duarte et al. 2004). Yet, they serve important ecological roles and are vital to the functioning and health of coastal ecosystems: (i) Ecosystem engineers: macroalgae form habitats and nursery grounds for diverse organisms. Fishes and invertebrates can be found in kelp forests and beds dominated by the order Laminariales and Fucales (Kirkman and Kendrick 1997; Macreadie et al. 2017; Gower and King 2019). By doing so they also increase and sustain marine biodiversity (Duarte et al. 2004). (ii) Together with seagrass macroalgae are the world's primary fishing grounds (Macreadie et al. 2017). (iii) Nutrient source: algae are an important nutrient source for many herbivore (grazers), but also for marine bacteria as they actively and passively release external organic metabolites (Khailov and Burlakova 1969; Neumann et al. 2015; Macreadie et al. 2017). (iv) In vegetated ecosystems macroalgae have a high contribution to respiration and gross primary production (Duarte et al. 2004) and thereby oxygenate the water, recycle nutrients, trap and stabilize sediments (Macreadie et al. 2017). (v) Influence the sedimentation of organic material: vegetated marine habitats contribute to almost half of the carbon burial in the coastal and global ocean (Duarte et al. 2004). Recent studies propose that they contributed to carbon sequestration for more than 2.1 billion years and about 173 TgC per year of which ~ 88% is sequestered in the deep sea (Krause-Jensen and Duarte 2016). About 43 % of their NPP is exported to sediments (Duarte and Cebrián 1996), especially due to phenols and refractory carbon compounds (Krause-Jensen and Duarte 2016). Hence, macroalgae act as important carbon donors in coastal ecosystems and make a meaningful impact on global blue carbon sequestration (Nellemann et al. 2009).

Within recent years, macroalgae gained popularity not only as food source but also as interesting candidates for applications in various sectors. They serve as phycocollides in biotechnology (e.g. agar for microbial growth) and industry (e.g. ingredient in food and cosmetics), but are also used in pharmacy as source for vitamins or drug candidates (El Gamal 2010 and references therein). Moreover, they are known as source of a number of oil deposits (Krause-Jensen and Duarte 2016), as fertilizer by local farmers (Smetacek and Zingone 2013), for biogas- and fuel (Wargacki et al. 2012a; Kawai and Murata 2016; Lara et al. 2020) and biomaterial production (Cesário et al. 2018).

Due to climate change and other anthropogenic activities leading to higher temperatures and eutrophication of coastal waters, algae must adapt to changing ecosystems. On one hand loss of kelp forests and consequently the loss of habitat for many marine species increases, but on the other hand biomass proliferation of specific algal species increases (Krause-Jensen and Duarte 2016). Seaweed tides like microalgae blooms are normal ecological phenomena known for centuries. For instance, golden tides composed of *Sargassum* sp. are known to occur along beaches between the Gulf of Mexico and Bermuda in summer, whereas green tides of *Ulva* sp. occur all over the world. Within the last years, seaweed tides have extended dramatically, reaching an unprecedented degree of dimensions and geographical areas (Liu et al. 2013; Smetacek and Zingone 2013; Wang et al. 2019a). The exact impacts of such events remain unclear, but certain shore-based activities including tourism already note negative effects (Gower and King 2019). Seaweed tides can drift long-distances on a horizontal but also vertical scale (Krause-Jensen and Duarte 2016), thereby acting as carbon donors to other long-term sequestration habitats (Hill et al. 2015). Macro- but also microalgae are capable of outcompeting other marine groups due to higher nutrient up-take rates. Especially coastal ecosystems under high nutrient input from e.g. rivers due to agriculture have an increased risk of shift in primary producers, which can lead to alterations in community and food web structures and ultimately in a loss of biodiversity and carbon sequestration function (Macreadie et al. 2017).

As for many other ecosystems such as corals or the intestinal system of ruminants, the health of macroalgae depends on its microbiota (Fernandes et al. 2012). Macroalgae interact with microbes located in the surrounded seawater or on the algae itself. The microbial surface colonization is characterized by a strong taxonomically diversity and depends greatly on host specificity as well as spatial and temporal variations (Martin et al. 2014 and references therein). Epiphytic communities include Alpha- and Gammaproteobacteria, Bacteroidetes, Cyanobacteria and members of the PCV (Planctomycetes-Clamydia-Verrucomicrobia) superphylum (Lachnit et al. 2011; Mann et al. 2013; Barbeyron et al. 2016b; Avcı et al. 2017). The so-called seaweed-holobiont represent a symbiosis, in which the algae provides oxygen, nutrients and carbon sources (e.g. in form of polysaccharides) for associated microbes, which in return improve the performance and the resilience of algae by providing CO_2, fixed nitrogen, inhibitors against other biofouling organisms (Egan et al. 2013 and

references therein). Epibiotic microbes are often fast colonizers, highly adaptive and enriched in polysaccharide-degrading bacteria which play a key role in the rapid metabolization of algal biomass and exudates (Goecke et al. 2010; Martin et al. 2015; Ihua et al. 2019). Polysaccharide-degrading bacteria are specialized for certain algal glycans to deal with intra- and interspecies competition. Marine Vibrios repeatedly acquired and expanded their metabolic repertoire leading to a fine scale ecophysiological differentiation among species. Their macroalgae associated lifestyle specialized on the same macroalgal glycan has led to a specialization into three ecophysiological types: (i) the pioneers which degrade the glycan into smaller oligosaccharides, which can directly be taken up by the (ii) scavengers or (iii) further extracellular degraded by harvesters (Hehemann et al. 2016; Corzett et al. 2018). This concept of coexistence of microbes has been further expanded to Bacteroidetes, Planctomycetes and Gammaproteobacteria. Selfish bacteria degrade glycans which are directly taken up, whereas sharing bacteria provide hydrolysis byproducts to scavenging bacteria which do not possess hydrolytic capacity (Reintjes et al. 2018).

Anionic polysaccharides and the degrading enzymatic machinery

Macroalgae are taxonomically diverse, consisting of several lineages with complex, independent evolutionary histories. Like plants they have evolved structurally versatile, dynamic, carbohydrate-rich cell walls over millions of years. Because of the unique biochemistry as well as the arrangement and interlinkage of polymers, the biosynthesis and also the metabolism of algae cell walls is still not fully understood (Popper et al. 2011). In general, algal cell walls are composed of a fibrillar skeletal framework with phenols and an extensive polysaccharide matrix (Macreadie et al. 2017). Especially the structural complexity of cell wall polysaccharides leads to a high variability in decomposition rates (Hill et al. 2015). The polysaccharide composition differs depending on the algal species, the part of the algae, the developmental and life-cycle stage but also abiotic factors such as temperature and geographical location (Popper et al. 2011).

Polysaccharides or glycans are monosaccharides linked by glycosidic bonds, resulting in either homoglycans (i.e. single sugar constitute) or heteroglycans constituted of multiple different monosaccharides. Both types are characterized and distinguished by the ring size of the monomer, anomeric configuration, position of linkage, absence or presence of branching and their charge – neutral (e.g. laminarin), positively or cationic (e.g. chitosan) or negatively charged or anionic polysaccharides (e.g. fucoidan, ulvan, alginate). The classification is based on the repetitive features (Aspinall 1983), which in theory leads to an almost infinite number of possible structures (Laine 1994). Polysaccharides are polymeric macromolecules that are synthesized by fungi, plants, bacteria and algae. Most marine organisms produce heteroglycans (Bajpai et al. 2014). Algae produce polysaccharides as energy and carbon source or to structure their cell wall. Due to their

location in coastal areas macroalgae are exposed to rapid changes of large amplitude in physical and chemical parameters because of the tidal cycles. Therefore, they have evolved structural versatile and complex cell wall systems (Fig. 2) to support the thallus (Deniaud-Bouët et al. 2017 and references therein). The cell wall rigidity of the brown algae order Fucales is caused by fucose-containing sulfated polysaccharides (FCSPs) and alginate, which is cross-linked with polyphenols and ions such as calcium (Deniaud-Bouët et al. 2014). In addition, cell wall polymers have specialized functions in algal development, cell adhesion and cell differentiation as well as protectant against pathogens (Deniaud-Bouët et al. 2017 and references therein) and heavy metals (Andrade et al. 2010). Especially anionic, i.e. negatively charged, polysaccharides seem to be of great importance. For instance, they can assist in the passive cellular ion transport by selective cation binding (Deniaud-Bouët et al. 2017 and references therein). Carboxylated polysaccharides like alginate are known to elicit the expression of defense responses in *Laminaria digitate* (Küpper et al. 2001), whereas sulfated polysaccharides cope with desiccation and osmotic stress at low tide. They facilitate water and ion retention in extracellular matrices (Mabeau and Kloareg 1987; Shepherd et al. 1999) and are excreted as slime to protect the algae from excessive UV exposure (Aspinall 1983).

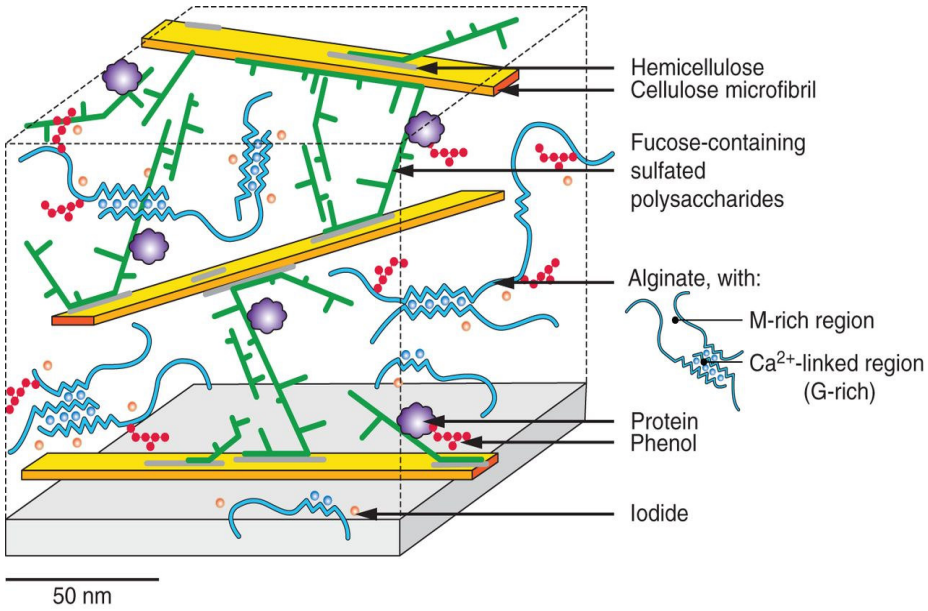


Figure 2. Cell wall structure of brown algae (Deniaud-Bouët et al. 2014). The cell walls of algae species from the order Fucales are characterized by hemicellulose and cellulose microfibrils, which are interlinked by the polysaccharides fucoidan and alginates.

Anionic polysaccharides are widely used in industry and pharmacy. Their main applications are stabilizers, thickeners and emulsifiers in food and cosmetics due to their gelling property (Kraan 2012; Bajpai et al. 2014). Recent clinical investigations demonstrated antioxidant, anticancer,

anticoagulant, anti-inflammatory, immunostimulatory, anticoagulant, antiviral, antiproliferative, anti-apoptic and antitumor properties (Laurienzo 2010; Croci et al. 2011; Ustyuzhanina et al. 2014; De Jesus Raposo et al. 2015; Thirumurugan and Dhanaraju 2017). Especially fucoidans are heavily studied for their potential drug usage treating several diseases (Gurpilhares et al. 2016; Garcia-Vaquero et al. 2017).

Nonetheless, the role of anionic polysaccharides in the marine carbon cycle is still barely understood. In the presence of divalent cations such as calcium, anionic polysaccharide form gel particles, which are three-dimensional networks of biopolymers imbedded in seawater. They can be formed from DOM or polymer chains released by phytoplankton or bacteria. Marine gels vary in composition, size and formation time. Anionic polysaccharides increase the stickiness of particles, which in turn influences the fate of sinking particles. Micro- and nanogels can gain size by continued annealing and collision forming macrogels such as transparent exopolymers (TEP), which is important regarding sedimentation processes of organic matter and thereby affecting deep sea communities. Hence, marine gels play an important role in the transition of DOM to POM, carbon cycling and the microbial loop of the marine food web (Passow et al. 1994; Chin et al. 1998; Passow 2000; Engel 2000; Passow 2002; Engel et al. 2004; Verdugo et al. 2004).

Marine heterotrophic bacteria have evolved highly specific and efficient enzymatic machinery to utilize all manner of polysaccharides, resulting into the key degrader of DOM and regulator of organic matter cycling in the marine environment.

Glycan degrading enzymes

Many heterotrophic bacteria possess operon-like or regulator-like gene clusters for glycan utilization. These gene cluster are characterized by transporters, carbohydrate-active enzymes (CAZymes) and associated protein such as binding proteins and sulfatases specific for a certain type of polysaccharide. Depending on the presence of SusCD-like proteins, which bind and transport glycans, the gene clusters are called polysaccharide utilization clusters (PULs) (Bjursell et al. 2006) or carbohydrate utilizing system (CUTs) (Hemsworth et al. 2016) such as the ulvan utilization loci in *Alteromonas* sp. LOR (Foran et al. 2017) or the alginate utilizing system in *Alteromonas macleodii* (Neumann et al. 2015). In CUTs, SusCD-like proteins are replaced by functional-equal proteins such as ABC transporter system (Poretsky et al. 2010) or carbohydrate-binding modules (CBMs) (e.g. Boraston et al. 2004; Mystkowska et al. 2018).

CAZymes catalyze the breakdown, biosynthesis or modification of carbohydrates and glycoconjugates and are therefore classified into glycoside hydrolases (GHs), polysaccharide lyases (PLs), glycosyl transferases (GTs), carbohydrate esterases (CEs) and auxiliary activities (AAs).

Each enzyme class harbors up to hundreds of different families and subfamilies, which are defined by amino acid sequence identity. The degradation of polysaccharides is carried out by PLs, which perform β-elimination (i.e. a non-hydrolytic cleavage) and GHs, which hydrolyze glyosidic bonds (Lombard et al. 2014a). PLs cleave uronic acid containing polysaccharides such as alginate and ulvan via a syn- or anti-elimination mechanism (Fig. 3), sharing the same three steps: First, a basic amino acid side chain or charge stabilizing cation such as Ca^{2+} or Mn^{2+} is removing the C5 proton of the uronic acid. The resulting anion is stabilized, and the charge is delocalized onto the C6 carbonyl group. The cleavage of the O-4:C-4 bonding that is facilitated by proton donation from a catalytic acid results into an unsaturated hexenuronic acid residue and a new reducing end at the point of cleavage. Glx/Asx act as neutralizer assisted by other charged amino acids (His, Arg or Glu) in the active site. The catalytic residues tyrosine or histidine/tyrosine act as Brønsted acid and base during syn or anti-elimination, respectively (Yip and Withers 2006; Garron and Cygler 2010; Garron and Cygler 2014).

Compared to β-elimination, hydrolysis is used much more frequently in glycan-degrading microbes (Garron and Cygler 2014). GHs perform the hydrolysis of the glycoside bond at C-1:O-4 either via a retaining or inverting mechanism. Both mechanisms are characterized by a pair of negatively charged amino acids as catalytic residues, which are either separated by ~ 5 or 6-12 Å, respectively. The retention or inversion of the anomeric configuration is either achieved via a direct displacement of by a nucleophilic and a general acid/base catalyst pair or a double-displacement of the aglycone moiety by a general acid and base catalyst pair (Yip and Withers 2004; Hemsworth et al. 2016).

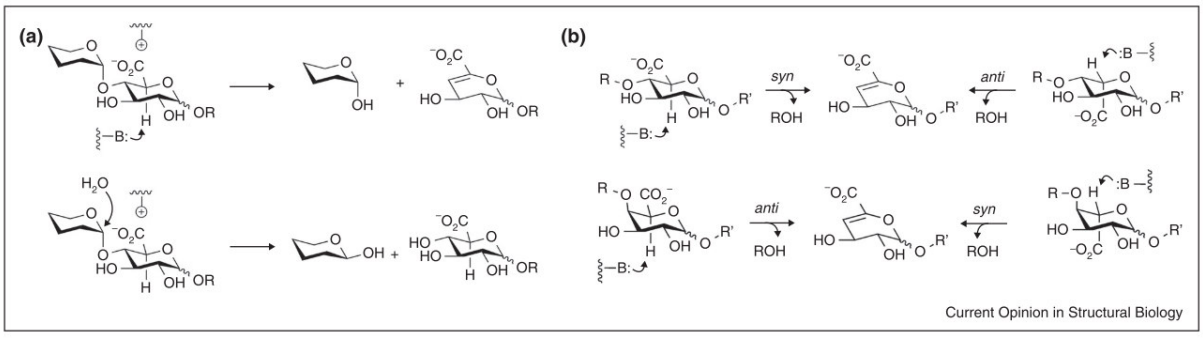


Figure 3. Enzymatic cleavage of glycans by GHs and PLs. A) cleavage of a hexose-uronate disaccharide by an inverting glycoside hydrolase and lyase, respectively. (B) Syn- and anti-pathways of D-glucoronide and L-iduronine, respectively (Garron and Cygler 2014).

Like CAZymes, sulfatases have been grouped into families based on amino acid sequence identity and classified into subfamilies corresponding to their substrate specificity (Barbeyron et al. 2016a). Till date, there are four sulfatase families: Family S1 (EC 3.1.6 and .10.1) is characterized by its

catalytic residue cysteine or serine, which is post translationally modified into the Cα-formylglycine (FGly), arylsulfatases of family S2 (EC 1.14.11.-) which catalyze the oxygenolytic cleavage of alkyl sulfate esters, and families S3 (EC 3.1.6.- and 3.1.6.19) and S4 (EC 3.1.6.1), which both belong to the metallo-beta-lactamase superfamily characterized by its zinc-binding motif. Even though sulfatases have been identified in many glycan degrading pathways, the number of biochemically and structurally characterized sulfatases are limited compared to CAZymes.

For a better understanding of the mechanistic motif, sulfatases, PLs and GHs have been distinguished into endo- or exo-acting enzymes. These terms describe where the enzymes act on the backbone of a poly- or oligosaccharide. Endo-acting enzymes hydrolyze or cleave within the backbone generating oligosaccharides, whereas exo-acting enzymes cleave mono- or disaccharides off the reducing or non-reducing end of the glycan chain, respectively (Fig. 4A). In both cases, the positions of the glycan's building blocks bound closely to the active site are called sugar-binding subunits (n), which are numbered (i.e. 1,2,3) according to their position towards the non-reducing (-) or reducing (+) end, respectively. The difference between sulfatases and CAZymes is the position where the catalytic mechanism takes place. In CAZymes, the cleavage occurs between the +1 and -1 subsite of the glycan. In case of sulfatases, the monosaccharide which is harboring the sulfate group is bound to the 0 subsite, and the cleavage occurs between the 0- and sulfate-binding subsite (Davies et al. 1997; Hettle et al. 2018). Nevertheless, the mode of action in sulfatases and CAZymes results in similar structural features. The active site of endo-acting enzymes is characterized by an open cleft, whereas exo-acting enzymes have a small pocket architecture (Fig. 4B). Structural techniques like X-ray crystallography can provide valid indications on the enzyme's function and putative substrates even if protein-ligand complexes cannot be achieved: (i) Identification of catalytic residues: With very few exceptions, catalytic residues possess polar and charged side chains, which are involved in proton shuttling (via protonation and deprotonation) and stabilization of the transition state or other intermediates mainly through electrostatic interactions (Zhou and Pang 2018). (ii) Interbiopolymer interactions, i.e. the attraction and repulsion between unlike macromolecules, are responsible for complex formation, which is mainly influenced by pH, ionic strength, confirmation, charge density and the ratio of protein to polysaccharide. Thus, the surface of the enzyme can provide valuable information on potential complex formation. The surface of proteins is non-uniform, allowing both attraction and repulsion interaction with its substrate by specific local patches (Schmitt et al. 1998; Doublier et al. 2000; Turgeon 2003; Turgeon et al. 2007). Many GH families possess secondary binding sites (SBSs), which are located relative to the active site and bind one or more carbohydrates. They are not conserved among related enzymes but are characterized by aromatic residues that recognize the substrate, accompanied by one or more other sites which act as "molecular tweezers" that orientate the side chains of the glycan prior hydrolysis in the active site. Their functional motif is very similar to CBMs. Hence, it is likely

that GHs compensate for the absence of CBMs by having one or multiple SBSs (Cuyvers et al. 2012). Another example besides hydrophobicity is the amino acid electrostatic potential of the enzyme's surface. Electrostatic interactions are the fundamental interaction between oppositely charged proteins and polysaccharides (Dubin 1988; Zhou and Pang 2018). For instance, anionic polysaccharides are bound by oppositely charged regions of the enzyme's surface, which results in a neutral net charge (de Kruif et al. 2004). Charged residues are crucial for the binding of not only ligands but also metals such as Ca^{2+} as cofactor certain for alginate lyases (Wong et al. 2000).

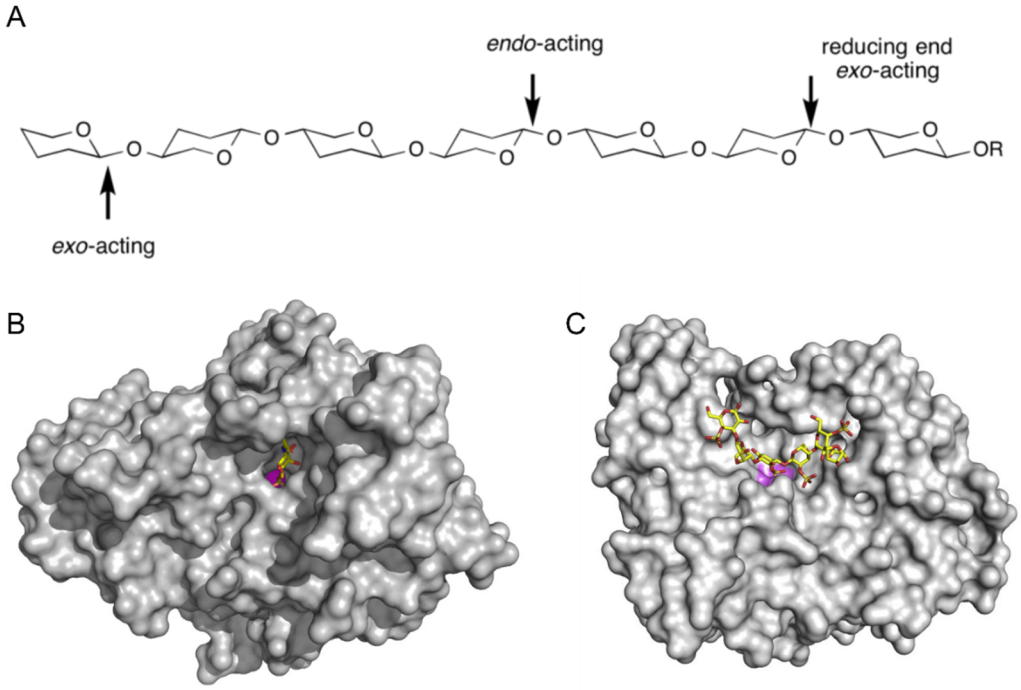


Figure 4. Structural insights of endo- and exo-acting enzymes. (A) Glycan degrading enzymes either act within the polysaccharide chain or at its reducing and non-reducing ends. (B) Exo-acting enzymes possess a small pocket architecture, whereas the active site of (C) endo-acting enzymes is characterized by an open cleft. Modified from (Davies et al. 2018; Hettle et al. 2018; Ndeh et al. 2020).

The structural and compositional variety of (anionic) polysaccharides influences the enzymatic machinery needed to degrade the glycans. The more complex the polysaccharide, the more and diverse enzymes are found in heterotrophic bacteria and/or several utilizing steps by various bacteria are needed. For instance, the catabolism of the neutral polysaccharide laminarin involves less enzymatic steps and CAZyme families compared to the anionic polysaccharides ulvan, fucoidan or carrageenan (Ficko-Blean et al. 2017; Unfried et al. 2018; Reisky et al. 2019; Sichert et al. 2020). The following sections compare the rather simple anionic polysaccharide alginate with the more complex glycans ulvan and fucoidans.

Alginate – a carboxylated and acetylated glycan from algae and bacteria

Alginate is the main cell wall component of brown algae (Phaeophyceae) and can contribute > 45 % to the dry weight (Mabeau and Kloareg 1987). Alginate is an unbranched, binary copolymer firstly discovered in the 1880s (Kraan 2012) and available in the acid and salt form. It is composed of uronic acids ß-D-mannuronate (M) and its C5-epimer α-L-guluronate (G) occurring in 1,4-linked blocks of either polyguluronate (polyG), polymannuronate (polyM), or alternating polyguluronate and polymannuronate (polyGM). The number and the ratio of M and G in the blocks depend on the species and affect the mechanical strength and flexibility of the algae. Alginate is an anionic polysaccharide containing one carboxylic acid per sugar monomer (Draget et al. 2005). Besides brown seaweed, alginate is also synthesized by two heterotrophic bacterial families, the Azotobacteriaceae and the Pseudomonadaceae (Rehm and Valla 1997; Sabra et al. 2001). In contrast to algal alginate, bacterial alginate is often substituted with O-acetyl groups at the C2 and/or C3 of D-mannuronate (Davidson et al. 2009), which affects the biochemical properties of alginate such as the binding of selective ions (Wong et al. 2000).

Alginate is widely used in industry, pharmacy and waste water treatment (Yabur et al. 2007; Langkamp and Uijterlinde 2015; Bedê et al. 2017). It is used as stabilizer, viscosifier (Draget et al. 2005) and hydrocolloid (Kraan 2012) due to its ability to retain water and from gels because of selective binding of multivalent cations such as Ca^{2+} and Mg^{2+}. Salts of alginate are soluble in water at low ionic strengths, but in aquatic environments alginate is mainly found in its insoluble form due to cross-linking by Ca^{2+} and to lesser extent Sr^{2+} ions in seawater (Aspinall 1983). Especially the interactions of Ca^{2+} ions with guluronate and MG residues leads to the formation of the so-called egg model in which the cations bind to the carboxyl groups and thereby form helical Ca-alginate fibers (Grant et al. 1973; Li et al. 2007b).

In algae, alginates are an essential component of the cell wall. They maintain the 3D structure due to the so-called intracellular egg-box junctions, which contribute to tissue stiffness and cell wall resistance during mechanical stress such as osmotic pressure during tides (Rabillé et al. 2019). The egg-box consist of cross linkages with guluronate residues from opposing chains of Ca-alginate (Grant et al. 1973; Li et al. 2007b). High L-guluronic acid content is found in old and tough tissue, but the overall ratio of M:G building blocks is also affected by the algal species as well as seasonal and spatial factors (Draget et al. 2005).

Marine heterotrophic bacterial lineages known for alginate degradation are Bacteroidetes and Proteobacteria (Thomas et al. 2012; Neumann et al. 2015). In-depth phylogenetic analysis has shown that the ancestral origin of alginolytic operons was probably a marine flavobacterium. Via independent lateral gene transfer, alginolytic operons were transferred to marine Proteobacteria (Alpha- and Gammaproteobacteria) and Japanese gut Bacteroides as a result of algae-based dietary

(Thomas et al. 2012). Alginate lyases (Aly) are typically extracellular or periplasmic (Wong et al. 2000), classified into the PL families 5-7, 14, 15, 17 and 18 (Lombard et al. 2014a). They cleave the glycosidic bond resulting in the oligomers L-guluronate or D-mannuronate and a nonreducing end with a 4-deoxy-L-erythro-hex-4en pyranosyl uronate residue (Fig. 5). The 4,5-unsaturated monosaccharide is converted to the intermediate 4-deoxy- L-erythro-5-hexoseulose (DEH) by KdgF, which is finally catalyzed to 2-keto-3-deoxy-gluconate (KDG) by DEH reductases. Alginate degrader gain energy from the conversion of KDG to pyruvate and 3-phosphoglyceraldehyde via the Enter-Doudoroff pathway (Hobbs et al. 2016). An alginate deacetylase to cleave acetyl groups of bacterial alginates has not been reported so far.

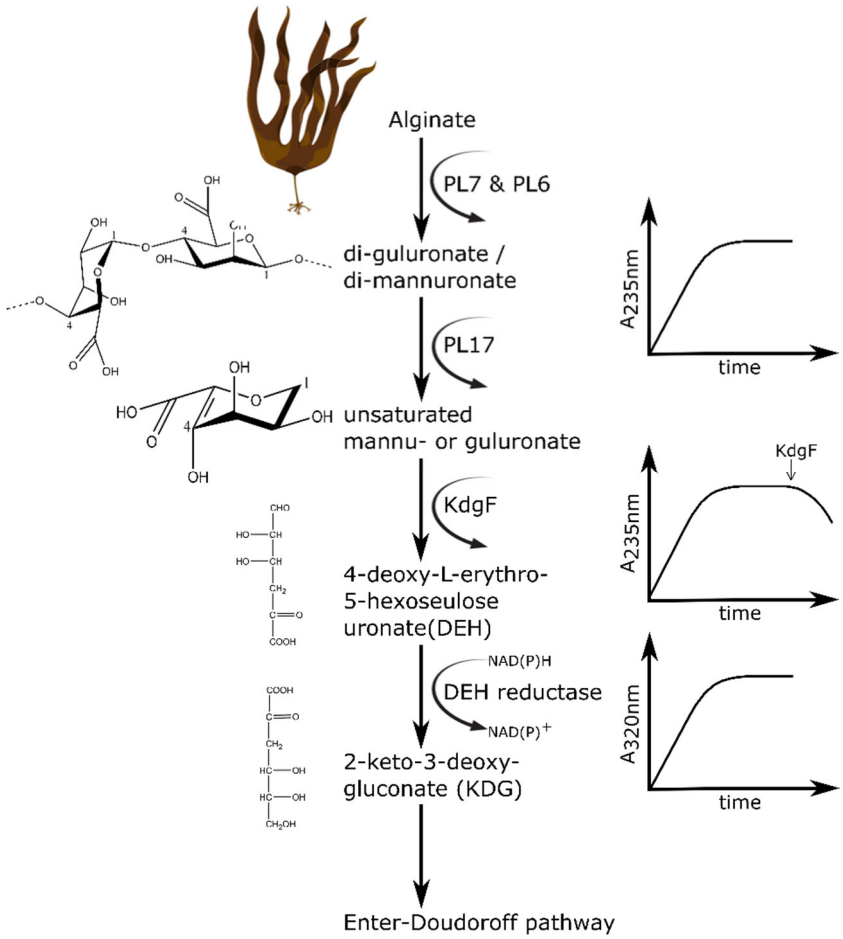


Figure 5. Enzymatic degradation of alginate. Endo-acting Alys of families PL7 and PL6 generate oligosaccharides, which are further degraded into the monosaccharides mannuronate and guluronate. Their activity can be observed at 235 nm representing the formation of unsaturated residues (double bond between C4 and C5) at the non-reducing end. KdgF further utilizes the unsaturated monosaccharide to DEH, resulting in a decrease at 235 nm. The activity of the DEH reductase can be measured at 320 nm, which represents the formation of NAD(P)+.

Ulvan – the sulfated cell well polysaccharide of green seaweed

Ulvan is the main cell wall polysaccharide of the algae order Ulvales from the globally distributed green algae (Chlorophyta) contributing up to 29 % of the algal dry weight (Lahaye and Robic 2007) with an average molecular weight of 189 – 8200 kDa (Lahaye 1998). There are two types of ulvan in nature – water-soluble and insoluble cellulose-like material (Kraan, 2012). Ulvan is structurally versatile with regards to charge density and molecular weight but consists mainly of L-rhamnose (Rha), D-xylose (Xyl) and uronic acids (Lahaye and Robic 2007). ^{13}C NMR spectroscopy identified branched glucuronic acid on O2 of rhamnose 3-sulfate (Lahaye 1998) and sulfate groups at C3 of rhamnose and/or C2 of xylose. The degree of sulfation varies among Ulvales species from 16 – 23 %. Ulvan is composed of repeating disaccharides of D-glucuronic acid (GlcA) and L-iduronic acid (IdoA) linked to 3-sulfated rhamnose (R3S). These building blocks are termed ulvanobiuronic acid A or B, respectively. There are also lower amounts of disaccharides of RS3 linked to Xyl (Lahaye and Robic 2007). In comparison to polysaccharides derived from brown and red algae such as alginate, polysaccharides of green algae are less exploited (Kopel et al. 2016).

Bacteria known for degrading ulvan belong to the orders Flavobacteriales and Alteromonadales, such as *Formosa agariphila* (Mann et al. 2013; Salinas and French 2017; Reisky et al. 2018; Reisky et al. 2019), *Nonlabens ulvanivorans* (Collen et al. 2011; Kopel et al. 2014a; Collén et al. 2014), *Pseudoalteromonas* sp. PLSV (Kopel et al. 2014b), *Alteromonas* spp. (Kopel et al. 2014c; Foran et al. 2017). Due to the large structural variability of ulvan, a multistep pathway including various endo- and exo-acting enzyme classes is required for a complete degradation (Salinas and French 2017; Reisky et al. 2019). These enzymes belong to ulvan hydrolases of families GH105, GH88, GH78, GH43 GH39, GH3 and GH2 and sulfatases S1_7 and S1_25. Ulvan hydrolases and sulfatases are highly specific towards the linkage of the sugar residue, i.e. α-L-rhamnosidases, β-xylosidases, β-galactosidases, iduronate-2-sulfatase, etc. Ulvan lyases belong to the families PL24 (Ulaganathan et al. 2018b), PL25 (Ulaganathan et al. 2017), PL28 (Collen et al. 2011; Reisky et al. 2018), PL37 and PL40 (Kopel et al. 2016; Li et al. 2020) and cleave the ß-1,4 glycosidic bond between the building blocks ulvanobiuronic acid A or B, resulting in the formation of a reducing end on one fragment and an unsaturated ring (Δ, 4-deoxy-L-threo-hex-4-enopyranosiduronic acid) on the non-reducing end of the second fragment (Ulaganathan et al. 2018a).

Fucoidan – sulfated fucans in algae and seagrass

Fucoidan or fucoidin is a synonym for fucose-containing sulfated polysaccharides (FCSPs). In some studies, fucoidan is also referred to as sulfated fucan, which rather describes the main characteristic of the macromolecule. To establish a coherent use of the term henceforth this molecule is called fucoidan. Fucoidans can be found in marine invertebrates (Ribeiro et al. 1994; Mulloy et al. 1994;

Yu et al. 2014a; Yu et al. 2014b) and in the cell walls of macroalgae (Deniaud-Bouët et al. 2017). This class of polysaccharide has been first described more than 100 years ago (Kylin 1913), but still our knowledge of its role in the marine environment is limited. This results mainly from its compositional and structural versatility which has been described for different brown algal species of the orders Fucales, Ectocarpales and Laminariales. Homofucans mainly consist of α-1,3- or alternating α-1,2/3/4-L-fucose, whereas heterofucans are additionally enriched in other monosaccharides like galactose, glucuronic acid, mannose and xylose. In both cases, fucose residues can be substituted with sulfate at C2, C3 and/or C4, 2-O-acetylation (Tab. 1). The composition does not only depend on the algal species (Fitton et al. 2015; Sichert et al. 2020), but is also influenced by seasonal and spatial factors (Anastyuk et al. 2010; Fletcher et al. 2017). In recent years, there was a significant improvement in extraction, purification and methodological approaches as well as combinations of different techniques leading to detailed insights into the 3D structure of different fucoidans compared to analyses of crude algal extracts in the past (Ale et al. 2011; Ale and Meyer 2013).

Fucoidan-like polysaccharides are also found in seagrasses (Kannan et al. 2013) and microalgae (Fig. 6). Microalgal cells and their exudates are known to contain high amounts of fucose and galactose (Gügi et al. 2015), yet little is known on the glycan composition and their 3D structure. FCSPs were detected in laboratory cultures of the diatom species *Thalassiosira weissflogii*, *Thalassiosira pseudonana* and on *Chaetoceros affinis* as well as in environmental samples from the diatoms spring bloom dominated by *Chaetoceros* spp. Moreover, putative fucoidans were also found in exudates of *T. weissflogii* (Vidal-Melgosa and Hehemann 2019), raising the questions of its functional properties in microalgae as well as its role in the marine carbon cycling.

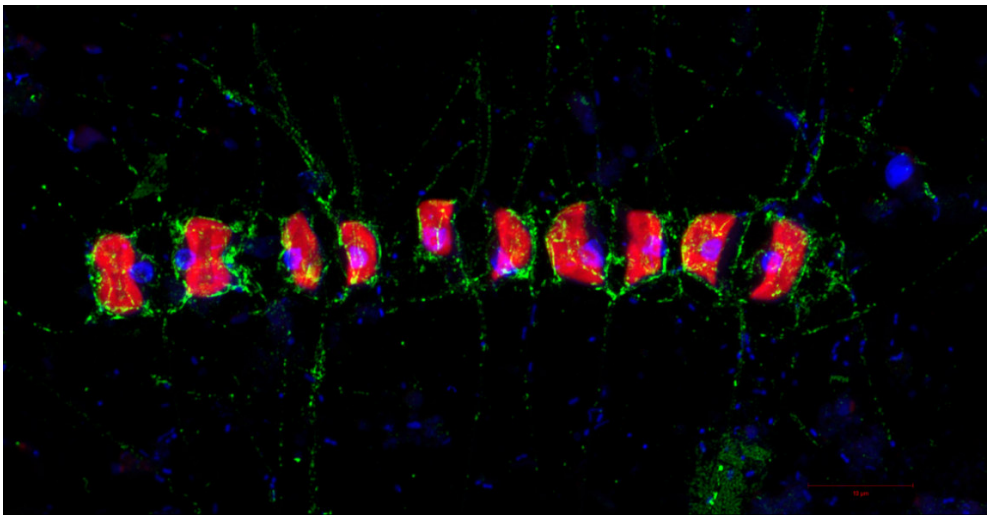


Figure 6. Fucose-containing sulfated polysaccharides in microalgae. Microscopic image of *Chaetoceros* spp. (red) from a spring bloom in Helgoland, German Bight, North Sea contain FCSPs (green). DNA staining using DAPI (blue). Image taken from (Vidal-Melgosa and Hehemann 2019).

Tab. 1. Structural complexity of macroalgal fucoidans. Presence (✓), absence (-) and no information available (/).

Order	Species	Fucoidan characteristics					
		L-Fucose	D-Galactose	Sulfation	Acetylation	Additional	References
Fucales	*Fucus vesiculosus*	α1,3 α1,4	✓	Fuc2S Fuc3S Fuc4S	-	Man, Xyl, Ara, Rha, Uronic acids	(Nishino et al. 1994; Espinel-Ingroff et al. 2007; Fitton et al. 2015)
	Fucus serratus	α1,2 α1,3 α1,4	✓	Fuc2S Fuc4S	Fuc3S Fuc4S	Man, Xyl, Uronic acids	(Bilan et al. 2006; Sichert et al. 2020)
	Durvillaea potatorum	✓	✓	✓	/	Man, Xyl, Ara, Rha, Uronic acids	(Fitton et al. 2015; Sichert et al. 2020)
	Sarassum fusiforme	α1,3 β1,4 β1,3 β1,4	α1,3 α1,4 α1,6	Fuc2S Fuc3S Fuc4S Gal2S Gal3S Man6S	-	Man, Xyl, Glc, Uronic acids	(Hu et al. 2014; Hu et al. 2016; Sichert et al. 2020)
Lamiariales	*Ecklonia maxmima*	α1,2 α1,3	α1,4	Fuc4S	✓	Man, Xyl, Uronic acids	(Sichert et al. 2020)
	Lessoia nigrescens	✓	✓	✓	/	Xyl, Ara, Rha, Uronic acids	(Fitton et al. 2015)
	Undaria pinnatifida	α1,3	β1,3 β1,4 β1,6	Fuc2S Fuc4S Gal3S Gal4S Gal6S	✓	Man, Xyl, Rha, Ara, Glc	(Lee et al. 2004; Hemmingson et al. 2006; Skriptsova et al. 2010; Synytsya et al. 2010; Fitton et al. 2015; Zhao et al. 2018)
	Macrocystis pyrifera	✓	✓	✓	/	Xyl, Rha, Uronic acids	(Fitton et al. 2015)
Ectocarpales	*Cladosiphon okamuranus*	α1,3 branches of α1,2	✓	Fuc4S	O-2 or O-4 of Fuc	Man, Xyl, Rha, Ara, Glc	(Nagaoka et al. 1999; Lim et al. 2019; Sichert et al. 2020)

Fucoidan is presumably the most complex class of marine glycans, comparable to the pectin plant polysaccharide rhamnogalacturonan-II (RG-II). RG-II contains 13 different sugars and 21 distinct glycosidic linkages. Still, it can be degraded by an individual bacteria possessing 25 highly specific enzymes (Ndeh et al. 2017). Fucoidan-degrading organisms and putative candidates have been identified by a combination of metagenomic, transcriptomic and proteomic analyses combined with classical cultivation approaches and biochemical studies. These bacteria are mainly found in the PCV (Planctomycetes-Chlamydia-Verrucomicrobia) superphylum (Sakai et al. 2003; Wegner et al. 2013; van Vliet et al. 2019; Sichert et al. 2020), *Flavobacteriaceae* (Descamps et al. 2006; Barbeyron et al. 2008; Chen et al. 2016) and Alteromonadales (Bakunina et al. 2002). Fucoidan-specialized bacteria possess hundreds of CAZymes and sulfatases to tackle the structural versatility of this class of glycans (Wegner et al. 2013; van Vliet et al. 2019; Sichert et al. 2020). The complete enzymatic pathway to degrade fucoidans has not been elucidated yet. The pronounced substrate specificity of fucoidan-degrading enzymes in combination with the structural versatility of fucoidans suggest an individual multistep utilization process depending on the type and source of fucoidan. Similar to other complex anionic glycans, the utilization process presumably involves dozens of enzymes acting in a distinct order (Ficko-Blean et al. 2017; Reisky et al. 2019). Fucoidan hydrolyzing CAZymes belong to three different classes: (i) fucoidanases, (ii) fucosidases, and (iii) sulfatases (Kusaykin et al. 2015).

Fucosidases (EC 3.2.1.51, -.111) are CAZymes belong to the glycoside hydrolase (GH) family GH29. They hydrolyze the α-1,2/3/4 or /6 fucosyl linkages of fucose-containing conjugates. Members of family GH29 are exo-acting CAZymes, i.e. they catalyze the removal of the nonreducing terminal L-fucose (6-deoxy-galactose, Fuc) (White Jr. et al. 1987; Eneyskaya et al. 2001a). The active site of GH29 is built of six highly conserved residues and is located in a $(\beta/\alpha)_8$ triosephosphate isomerase (TIM) barrel domain (Sela et al. 2012; Sulzenbacher et al. 2004; Summers et al. 2016) and harbors the catalytic residues aspartic acid (Asp) and glutamic acid (Glu) as nucleophile and Bronsted acid/base, respectively (Tarling et al. 2003; Cobucci-Ponzano et al. 2003a; Sulzenbacher et al. 2004). So far, there is only one structure known which deviates from the highly conserved active site found in family GH29. The structural analyses of a GH29 in complex with competitive inhibitors of β-D-galactosidases identified two pockets. One binds fucose, whereas the other one binds galactose. This observation is in line with biochemical analyses demonstrating that BT2192 from *B. thetaiotaomicron* is the first GH29 with a dual activity towards α-1,3/4-L-fucose and β-D-galactose. Like other fucosidases, BT2192 possesses the typical catalytic pair of Asp and Glu, which are located in between both pockets. Structural insights suggest, that BT2192 possesses a Glu as second nucleophile which can attack the intermediate to release D-galactose (Guillotin et al. 2014).

Family GH29 is highly upregulated in macroalgae associated and fucoidan degrading bacteria from the PCV superphylum (Kim et al. 2016; Sichert et al. 2020) and Flavobacteria dominating microalgae blooms (Teeling et al. 2012; Kappelmann et al. 2019). However, biochemical and structural studies revealing the linkage and substrate specificity for FCPS from macro- and presumably also microalgae are still limited. So far, only two fucosidases from marine mollusks *Haliotis gigunteu* and *Pecten maximus* have been reported to be active on macroalgal fucoidans (Tanaka and Sorai 1970; Berteau et al. 2002). In contrast, endo-acting fucoidanases have been studied more extensively.

Fucoidanases (EC 3.2.1.-) are either unclassified or belong to family GH107 (Kusaykin et al. 2006; Silchenko et al. 2013; Silchenko et al. 2014; Kim et al. 2015; Silchenko et al. 2016; Silchenko et al. 2017; Cao et al. 2018; Schultz-Johansen et al. 2018), which displays structural similarities to family GH29 (Vickers et al. 2018). They are characterized by a retaining mechanism with conserved histidine and aspartic acid to act as an acid/base and nucleophile pair. Family GH107 have been only reported to cleave α-1,4-fucosidic linkages with a pronounced substrate specificity. FFA2 from *Formosa algae* KMM 3553^T cleaves between repeating units of L-fucose-2-sulfate-α-1,3-L-fucose-2-sulfate in fucoidan from *Fucus evanescens* (Silchenko et al. 2016), whereas FcnA from *Mariniflexile fucanivorans* SW5 hydrolyses between repeating units of L-fucose-2,3-disulfate-α-1,3-L-fucose-2-sulfate in fucoidan from *Pelvetia canaliculata* (Colin et al. 2006). P5AFcnA and P19DFcnA from marine *Psychromonas* sp. SW5A and SW19D, respectively, are active on fucoidans from *Laminaria hyperborea* and *Macrocystis pyrifera* (Vickers et al. 2018).

Similar to GHs, sulfatases (EC 3.1.6) can be found in all kingdoms of life, have high specificity for the targeted carbohydrate backbone (Hettle et al. 2018) and are grouped into homologous enzyme families (Barbeyron et al. 2016a). Sulfatase family S1 is the biggest family consisting of 72 subfamilies and non-classified candidates. S1 family is characterized by a small C-terminal and a large N-terminal domain harboring the active site. The active site is consisting of ten conserved residues: five basic and four acidic/polar amino acids as well as a formylglycine (FGly) residue, which derives from post-translational oxidation of a Cys or Ser and is essential for the activation of many sulfatases (Schmidt et al. 1995; Lukatela et al. 1998; Dierks et al. 1998). Fucoidan-degrading bacteria are especially enriched in sulfatases. For example, *Rhodopirellula baltica* SWK7 has 181 sulfatases (Wegner et al. 2013), *Lentisphaera araneosa* HTCC2155T has 267 sulfatases (Thrash et al. 2010) and *Kiritimatiellales* F1 has 521 sulfatases (van Vliet et al. 2019). However, to date, only three sulfatases of families S1_17 and S1_15 from the marine mollusk *Pecten maximus* (Daniel et al. 2001; Berteau et al. 2004) and *Wenyingzhuangia fucanilytica* CZ1127^T (Silchenko et al. 2018) have been shown to act on fucoidan, whereas most uncharacterized enzymes remain potential candidates. The most abundant and upregulated candidates belong to the subfamilies S1_7, S1_8, S1_15, S1_16 and S1_25 (van Vliet et al. 2019; Sichert et al. 2020).

Based on the structural versatility of fucoidans, there are presumably fucoidan-degrading enzymes from various other families, which are capable to act on different building blocks than sulfated fucose. Such enzymes could be galactosidases, xylosidases and mannosidases as well as enzymes capable to cleave uronic acids. Family GH95 appears to be a promising candidate. Similar to GH29s, family GH95 is exo-active and enriched in Flavobacteria dominating microalgae blooms (Teeling et al. 2012; Kappelmann et al. 2019) and fucoidan-degrading bacteria (Kim et al. 2016; Sichert et al. 2020). Compared to family GH29, GH95s are inverting CAZymes that are either fucosidases (Nagae et al. 2007; Sela et al. 2012; Fan et al. 2016; Zeuner et al. 2018; Hobbs et al. 2019) or galactosidases (Rogowski et al. 2015). Both families have not been reported to be active on decorated fucose- or galactose-containing substrates. Hence, they might require enzymatic partners to remove fucoidan decoration with regards to sulfation and acetylation. To date, there is only one study reporting enzymatic deactelyation of macroalgal fucoidans. The marine bacterium *Luteolibacter algae* H18 produces an intracellular enzyme that cleaves acetate from fucoidan *C. okamuranus*. This enzyme shares sequence identity with many uncharacterized proteins as well as with a dienelactone hydrolase and an acetylxylan esterase (Nagao et al. 2017). The low sequence identity suggests that there might be many more fucoidan-degrading enzymes with novel functions from existing and maybe even novel CAZyme and sulfatases families yet to be discovered.

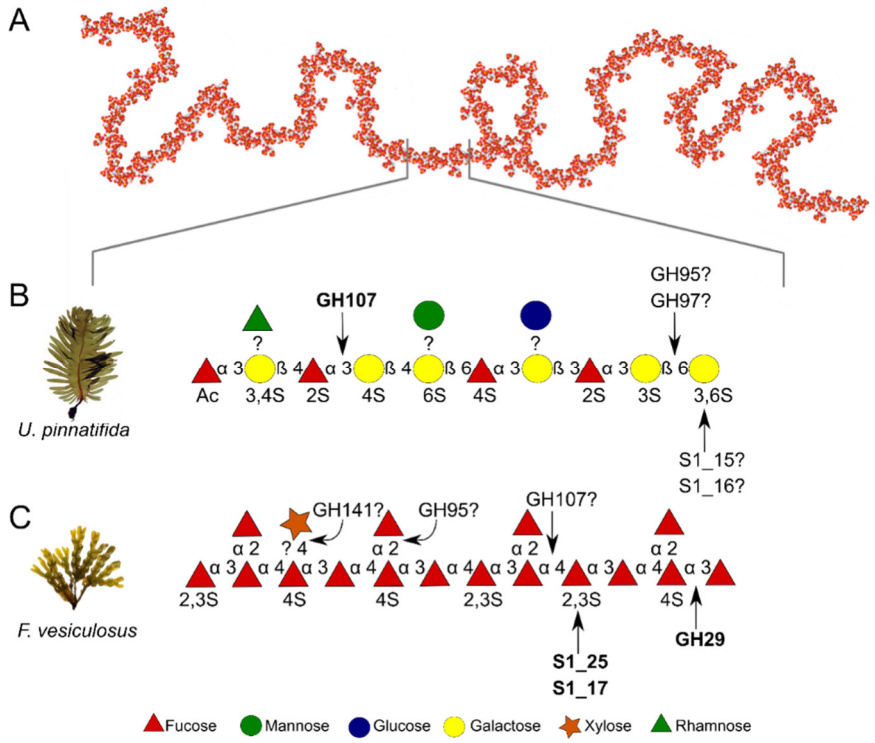


Figure 7. Fucoidan-degrading enzymes. (A) Tertiary structure of macroalgal fucoidan modified after (Kopplin et al. 2018). Structure of heterofucoidan from (B) *Undaria pinnatifida* and (C) homofucoidan from *Fucus vesiculosus*. CAZymes and sulfatase families with known function are shown in bold, whereas candidates and unknown sugar linkages are marked with a question mark.

CAZymes as quantitative tools in oceanography

In recent years, technological innovations have led to great improvements of existing techniques but also to new methods investigating carbohydrates. Elucidating the composition, structure and mixture of glycans is still a challenging task because many approaches have strict requirements of the samples (i.e. concentration, purity, sample matrix, interfering substances such as salts). Techniques like mass spectrometry (MS) or nuclear magnetic resonance (NMR) are often cost- and lab-intensive (Gray et al. 2019). Others provide insights into the monosaccharide composition but lack the information of 3D-structure, i.e. linkage type, branching, and substitutions (e.g. carboxyl, acetyl or sulfate groups), as it is the case with non-standardized acid hydrolysis of carbohydrates combined with chromatography (Panagiotopoulos and Sempéré 2005 and references therein). Enzyme-based methods however are less lab- and cost-intensive, faster and highly accurate (Garcia-Vaquero et al. 2017 and references therein). Such methods take advantage of substrate specific CAZymes and have been mainly applied to plant polysaccharides and carbohydrate detection in food industry (McCleary 1981; Gibson et al. 1992; McCleary et al. 1997; McCleary and Blakeney 1999; McCleary et al. 2007; Vidal-Melgosa et al. 2015; Mangan et al. 2019). Recently, this approach has been transferred to the marine environment to quantify the storage and cell wall polysaccharide laminarin of macro- and microalgae (Becker et al. 2017; Scheschonk et al. 2019). The usage of CAZymes overcomes the need of highly purified and/or concentrated samples and hence, allows the analyses of glycan mixtures from environmental samples. The enzymatic laminarin assay revealed spatial and temporal variations of its concentration in algal cells and its contribution to the marine POC pool (Becker et al. 2020). Developing enzyme-based methods also for other marine glycans can help to determine their abundance and turnover rates in the carbon cycle.

Aim of the study

Anionic polysaccharides are an important class of biological macromolecules in the marine environment. They are shaping the three-dimensional structure of flora, enzymatic feeding strategies of bacteria and potentially sequestering carbon in form of sinking particles. However, little is known about their turnover rates by bacteria and their exact role in the marine carbon cycle. Therefore, the anionic polysaccharides ulvan and fucoidan were chosen to shed light on different aspects on the bacterial degradation of algal polysaccharides and their fate in the marine environment. The objectives of the present dissertation are displayed as chapters, each presented by one of the model glycans.

1. Identification of fucoidan-degrading enzymes from '*Lentimonas'* sp. CC4

The fucoidan-degrading marine Verrucomicrobium '*Lenitmonas'* sp. CC4 has been studied recently by combining transcriptomic and proteomic analyses with classical microbial cultivation approaches. '*Lentimonas'* sp. CC4 possesses over two hundred putative fucoidan-degrading enzymes, which are primarily located on a 0.89 mbp plasmid. Strain CC4 has 113 sulfatases, 100 glycoside hydrolases and 17 carbohydrate esterases. The isolate exhibits a remarkable substrate specificity for fucoidans from different macroalgal species in combination with an enhanced repertoire of exo-acting CAZymes and sulfatases (Sichert et al. 2020). To shed light on its enzymatic pathways for fucoidan degradation as well as to investigate the hypothesis of an exo-acting growth strategy, I focused on exo-acting CAZymes and sulfatases that were upregulated in the presence of fucoidan. I performed intense biochemical characterizations of fucosidases of the families GH29 and GH95 as well as one sulfatase of family S1_15. Fucoidan activity was investigated using nine structurally different fucoidans originated of the three macroalgal orders Fucales, Laminariales and Ectocarpales. Our detailed mechanistic insights were combined with structural biology to determine the substrate and linkage specificity of these enzymes. The biochemical and structural characterization of these enzymes is the base for future studies studying the complete utilization of fucoidan degradation.

2. Structural analysis of ulvan-degrading enzymes from *Formosa agariphila* KMM3901^T

Formosa agariphila KMM3901^T is a Gram-negative, aerobic, and motile bacterium (Ivanova et al. 2004) originally isolated from Troitsa Bay, Gulf of Peter the Great, East Sea from the green alga *Acrosiphonia sonderi* (Nedashkovskaya et al. 2006). Genome analysis revealed that strain KMM3901 exhibits an algae-associated lifestyle, using many putative genes for the degradation of algal polysaccharides in at least 13 PULs. These PULS are characterized by dozens of sulfatases,

PLs, GHs, SusD-like proteins, transporter proteins, transcriptional regulators and other proteins (Mann et al. 2013). One of these PULs contains genes essential for the degradation of sulfated algal polysaccharide ulvan, including GH105 and ulvan lyase homologs which have been previously described as ulvanases (Kopel et al. 2016). To elucidate the ulvan degradation pathway from *F. agariphila*, a collaborative approach was needed combining different research fields. We contributed to this multidisciplinary study by investigating the structure of enzymes by X-ray crystallography. Primarily, I focused on endo-acting enzymes initiating the hydrolysis and genes with low sequence identity compared to their known structural homologues, indicating a novel activity and/or 3D structure.

Chapter 1 – Structural analysis of ulvan-utilizing enzymes

Study motivation in brief:

Have you ever wondered who cleans up this "mess"? Continue reading!

© Photo credit: Pixabay.

Chapter 1.1. A marine bacterial enzymatic cascade degrades the algal polysaccharide ulvan

Lukas Reisky,[1#] Aurélie Préchoux,[2#] Marie-Katherin Zühlke,[3,4#] Marcus Bäumgen,[1] Craig S. Robb,[5,6] **Nadine Gerlach**, [5,6] Thomas Roret,[7] Christian Stanetty,[8] Robert Larocque,[7] Gurvan Michel3, Tao Song,5,6 Stephanie Markert,[3,4] Frank Unfried, [3,4] Marko D. Mihovilovic,[8] Anke Trautwein-Schult,[9] Dörte Becher,[9] Thomas Schweder,[3,4*] Uwe T. Bornscheuer,[1*] Jan-Hendrik Hehemann[*5,6]

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

[2] Sorbonne Université, CNRS, Integrative Biology of Marine Models (LBI2M), Station Biologique de Roscoff (SBR), 29680 Roscoff, Bretagne, France

[3] Pharmaceutical Biotechnology, Institute of Pharmacy, University Greifswald, 17487 12 Greifswald, Germany

[4] Institute of Marine Biotechnology, 17489 Greifswald, Germany

[5] Max Planck-Institute for Marine Microbiology, 28359 Bremen, Germany

[6] University of Bremen, Center for Marine Environmental Sciences (MARUM), 28359 Bremen, Germany

[7] Sorbonne Université, CNRS, FR 2424, Station Biologique de Roscoff (SBR), 29680 Roscoff, Bretagne, France

[8] Institute of Applied Synthetic Chemistry, TU Wien, 1060 Vienna, Austria 9 Institute of Mi-crobiology, University Greifswald, 17487 Greifswald, Germany

[#] These authors contributed equally to this work.

* To whom correspondence should be addressed: schweder@uni-greifswald.de (TS); uwe.bornscheuer@uni-greifswald.de (UB); jhhehemann@marum.de, jheheman@mpi-bremen.de (JHH)

Published in Nature Chemical Biology

Received: 07th January 2019

Accepted: 21st May 2019

Published online: 8th July 2019

DOI : 10.1038/s41589-019-0311-9.

Abstract

Marine seaweeds increasingly grow into extensive algal blooms, which are detrimental to coastal ecosystems, tourism, and aquaculture. However, algal biomass is also emerging as sustainable raw material for the bioeconomy. The potential exploitation of algae is hindered by our limited knowledge of the microbial pathways – and hence the distinct biochemical functions of the enzymes involved – that convert algal polysaccharides into oligo- and monosaccharides. Understanding these processes would be essential, however, for applications like the fermentation of algal biomass into bioethanol or other value-added compounds. Here we describe the metabolic pathway that enables the marine flavobacterium *Formosa agariphila* to degrade ulvan, the major cell wall polysaccharide of bloom forming *Ulva* species. The pathway involves 12 biochemically characterized carbohydrate-active enzymes, including two polysaccharide lyases, three sulfatases and seven glycoside hydrolases that sequentially break down ulvan into fermentable monosaccharides. This way, the enzymes turn a previously unexploited renewable into a valuable and ecologically sustainable bioresource.

Introduction

Algal photosynthesis provides half of the global primary production (Field et al. 1998). Carbon dioxide is converted into carbohydrates, which are polymerized into polysaccharides to store energy, build cell walls, and perform other biological functions. Algae are furthermore considered as a promising renewable carbon source, due to their competitive growth rates and unique cell walls. Unlike plants that are rich in woody tissue, comprising the insoluble polysaccharides cellulose and the aromatic polymer lignin, which increases recalcitrance against enzymatic digestion, algal cell walls are rich in gel-forming polysaccharides that are highly hydrated (Kloareg and Quatrano 1988). Hydration and the absence of lignin make harsh chemical and physical pretreatment of cell walls unnecessary and allow for easy access of enzymes that can digest the polysaccharides into fermentable monosaccharides. Accordingly, recent studies showed that bioengineered microbes equipped with agarases and alginate lyases can efficiently digest and rapidly convert polysaccharides from brown and red algae into bioethanol (Wargacki et al. 2012b).

Sessile macroalgae, such as brown algae that form kelp forests, are ecologically valuable because they provide nutrition and habitats for fish and other organisms and, consequently, harvesting them would exacerbate pressure on natural populations. However, the planktonic macroalgae *Ulva armoricana*, *Ulva rotunda* and other *Ulva* spp. that thrive in eutrophic, nutrient-rich coastal waters, grow into expansive blooms that occur with increasing frequency in recent years. They pose ecological but also economical threats when they accumulate on beaches used for recreation (Ménesguen and Piriou 1995; Liu et al. 2013; Smetacek and Zingone 2013). Fertilized by nitrate from agriculture that is washed into the ocean by rivers, *Ulva* blooms during summer produce up to 50–100 000 tons of biomass every year, which must be removed at high expense from the northern and western coast of France (Ménesguen and Piriou

1995). Even larger blooms occur in China (Liu et al. 2013). Blooms of *Ulva* are thus a global phenomenon that is bound to increase with farming activities, rendering the polysaccharide ulvan, which accounts for up to 30 % of the algal dry weight (Lahaye and Robic 2007), an emerging yet untapped resource.

Ulvan is a branched polysaccharide composed of repeating disaccharide units, in which D-glucuronic acid (GlcA) is β-1,4-linked or L-iduronic acid (IdoA) is α-1,4-linked to L-rhamnose-3-sulfate (Rha3S), which is α-1,4-linked within the main chain. Some of the uronic acids are replaced by β-1,4-linked D-xylose (Xyl), which can be sulfated at position 2 (Xyl2S). Furthermore, Rha3S can be modified by β-1,2-linked GlcA side chains and the GlcA-Rha3S or IdoA-Rha3S pattern can be interrupted by consecutive GlcA residues (Lahaye et al. 1997; Lahaye 1998; Lahaye and Robic 2007). Increased interest in the enzymatic degradation of ulvan recently led to the description of several ulvan-active enzymes (Ulaganathan et al. 2017; Melcher et al. 2017; Ulaganathan et al. 2018b; Ulaganathan et al. 2018a; Reisky et al. 2018). So far, and to the best of our knowledge, only two types of enzymes from different carbohydrate-active enzyme (CAZyme) families showed activity on ulvan. Ulvan polysaccharide lyases of the families PL24, PL25 and PL28 catalyze the initial cleavage between Rha3S and GlcA or IdoA, resulting in the formation of unsaturated uronic acid residues at the end of the formed oligosaccharide. Unsaturated uronic acid residues are removed by glycoside hydrolases (GHs) from the family GH105 (Collén et al. 2014; Salinas and French 2017). Ulvan-specific degradation-related gene loci ('polysaccharide utilization loci', PULs) such as PUL H from *Formosa agariphila* encode PL28 and GH105 together with over 10 additional, putative enzymes, which were predicted to be involved in ulvan utilization. While PL28 and GH105 degrade ulvan, the other enzymes that were produced in *Escherichia coli* did not show activity (Salinas and French 2017). This result suggested that a complex cascade of sequential enzymatic reactions is required for complete ulvan degradation (Foran et al. 2017; Salinas and French 2017).

Here, we experimentally established the complex ulvan degradation pathway of *F. agariphila* KMM 3901^T, a marine flavobacterium, which was isolated from a green alga in the Sea of Japan (Mann et al. 2013). These degradation-related enzymes are encoded in an ulvan-specific PUL in the bacterial genome (Salinas and French 2017).

Results

Bacterial ulvan-specific PULs

To decipher the ulvan degradation pathway, we first searched microbial genomes hosted at NCBI for potential ulvan-specific PULs using the known ulvan lyase PL28 as query. We identified 12 putative ulvan PULs in 12 Bacteroidetes genomes (Fig. 1a), including the recently discovered PUL H of *F. agariphila* (Mann et al. 2013; Salinas and French 2017), a more than 75 kb long genomic region

consisting of 39 genes (Fig. 1b). We verified the boundaries of PUL H with a comparative global proteome analysis of *F. agariphila* cells fed with ulvan and with control substrates (rhamnose and fructose), respectively, as sole carbon source. Ulvan promoted bacterial growth (Supplementary Fig. 1) and elicited quantitative changes of most proteins that are encoded by PUL H (Fig. 1b and 2, and Supplementary Fig. 2). Besides ulvan, also the monosaccharide rhamnose induced, albeit less strongly, the expression of PUL H genes. For a few proteins (P2_SusD, P3_TBDR, P8_GH2) even higher protein amounts were detected with rhamnose, compared to ulvan. The increased abundance of enzymes involved in the degradation of ulvan-derived monosaccharides indicated a co-regulation of genes for the metabolization of ulvan and its corresponding monosaccharides (Fig 2, Supplementary Figs. 2 and 3, Supplementary Data Sets 1 and 2). PUL H includes 17 potential carbohydrate-active enzymes (CAZymes) from different GH and PL families (http://www.cazy.org/ (Lombard et al. 2014a)) and eight sulfatases from five S1 subfamilies (http://abims.sb-roscoff.fr/sulfatlas/ (Barbeyron et al. 2016a)). For most of these enzymes, their role in ulvan depolymerization remains unknown. A co-occurrence analysis of putative enzymes and associated genes within the set of 12 PULs from marine Bacteroidetes identified conserved CAZymes in the putative ulvan pathways (Fig. 1c). This analysis allowed us to focus our biochemical experiments on a smaller subset of CAZymes and sulfatases, whose involvement in ulvan utilization was suggested by our proteomic results (Fig. 1b). In addition to the two already known (Salinas and French 2017) ulvanolytic enzyme activities (ulvan lyase and unsaturated glucuronyl hydrolase, GH105) we uncovered eight so far unknown enzyme functions for the complete depolymerization of ulvan. Besides a novel PL family, we identified and characterized six GH families (GH2, GH3, GH39, GH43, GH78, GH88) and three sulfatases. Activity-based screenings of these enzymes were used to identify their function in the ulvan degradation pathway. The selection of putative CAZymes and sulfatases for cloning, heterologous expression and characterization was guided by the co-occurrence analysis of genes in the diverse ulvan PULs (Fig. 1c).

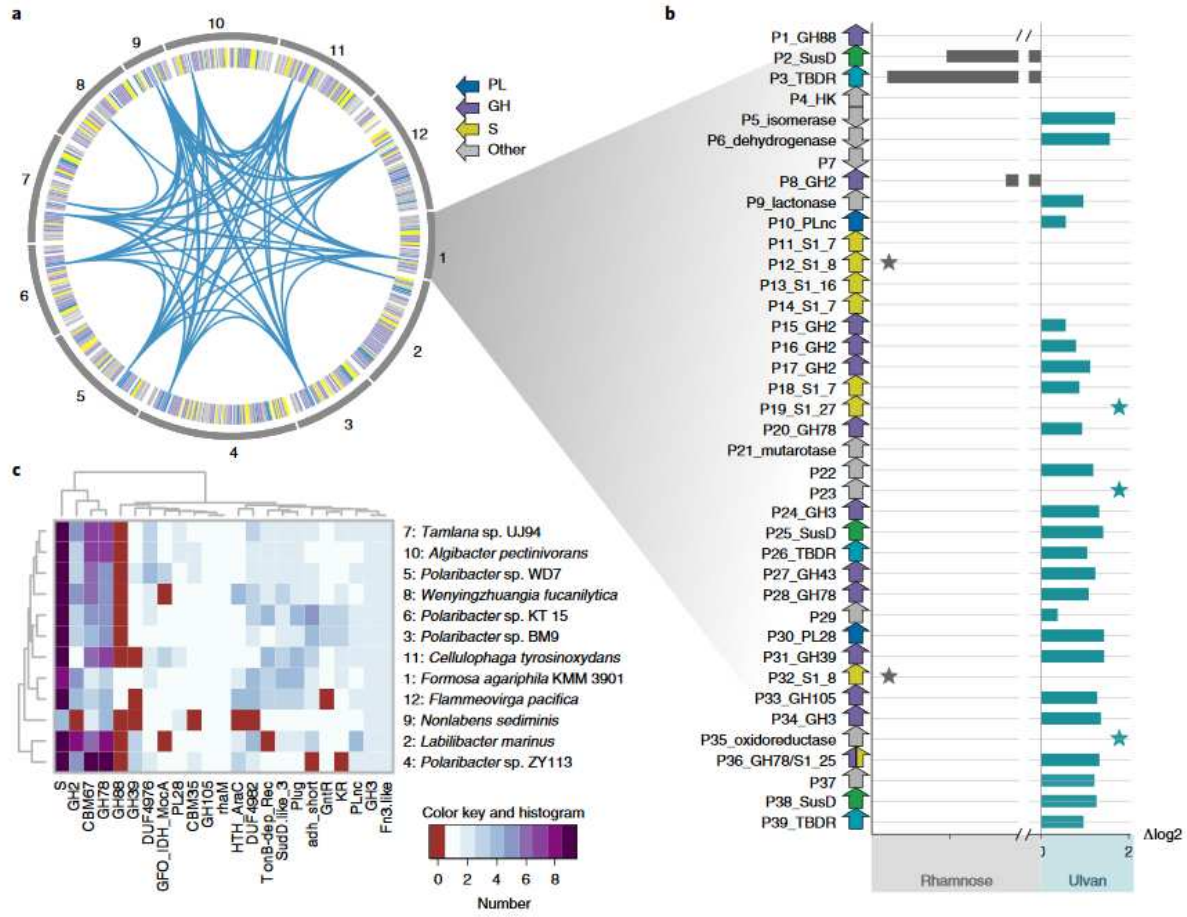


Fig. 1 Genomic overview of putative ulvan PUL in marine Bacteroidetes and the proteomic response of the *F. agariphila* PUL to ulvan and rhamnose. a, Comparative genomics of ulvan PUL that contain the known PL28 ulvan lyase (connected with blue lines when over 50% identical) revealed that the enzymes are encoded by conserved genes in diverse marine Bacteroidetes genomes, including the model organism of this study, *F. agariphila*, shown as no. 1; the complete list of all analyzed strains is provided in c. SusD and TBDR proteins are colored as 'other' in this panel. **b**, Ulvan and rhamnose as sole carbon sources elicit quantitative changes in proteins encoded in the putative ulvan PUL in *F. agariphila*. Bars indicate relative changes between both conditions. A positive $\Delta\log_2$ value corresponds to higher protein abundance with ulvan, whereas a negative value corresponds to higher protein abundance with rhamnose. Stars mark proteins that were exclusively quantified in either ulvan- or rhamnose-grown cells (Fig. 2, Supplementary Fig. 2 and Supplementary Dataset 1). Arrows refer to the orientation of genes that encode the respective proteins. Proteins encoded by the ulvan PUL were numbered (P1–P39) and protein function was indicated (Fig. 2). In the case of GH and sulfatases (S), families and subfamilies were specified (Lombard et al. 2014a; Barbeyron et al. 2016a); for example, GH2 (family) or S1_7 (family and subfamily). **c**, Co-occurrence analysis of genes in the predicted 12 putative bacteroidetal ulvan PUL highlights a conserved set of ulvan-degrading enzymes. The dendograms shown above and to the left of the similarity heatmap depict the pairwise similarities between rows and columns, respectively.

Sulfatases active on ulvan

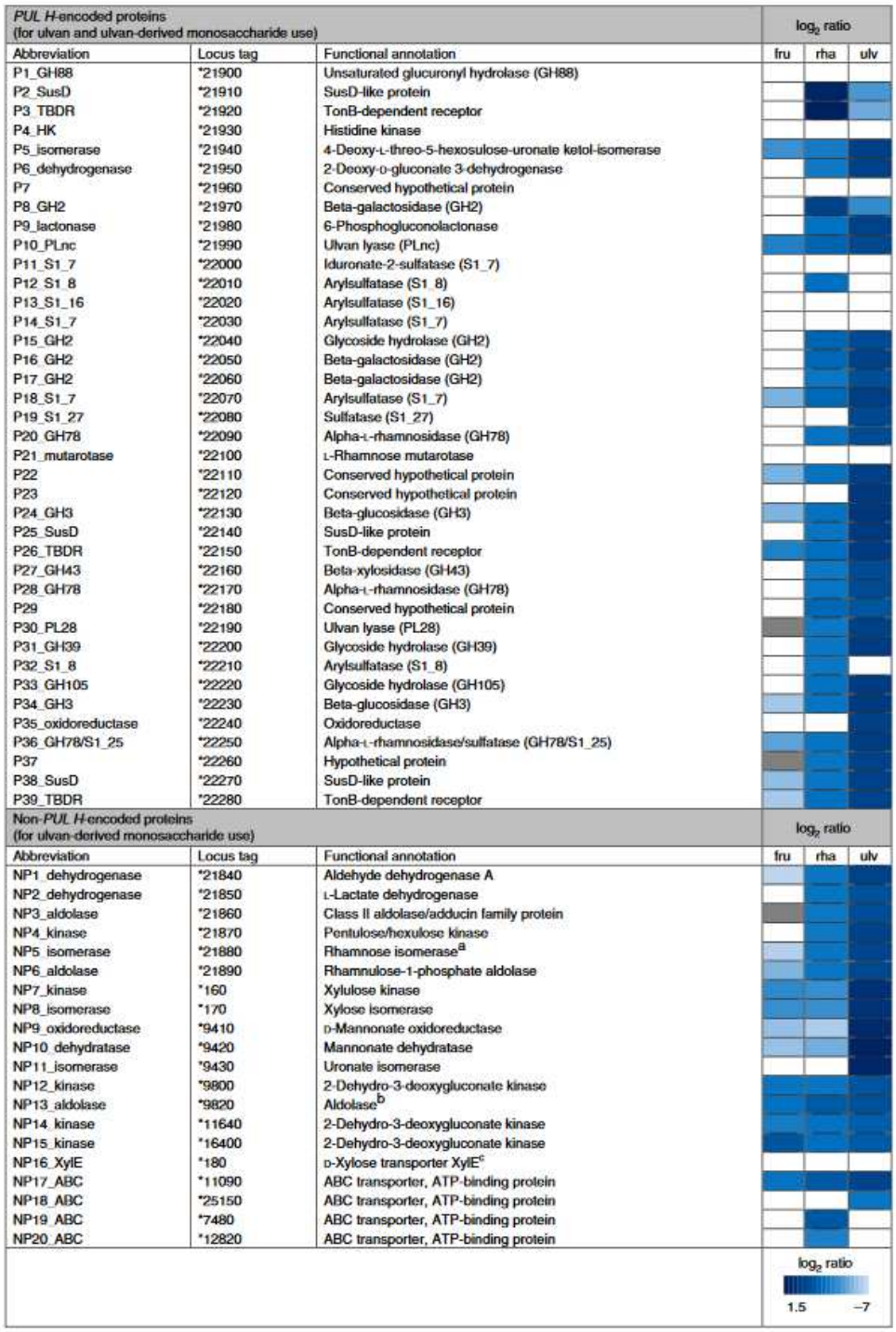
Ulvans feature a large structural variability, with substitution by sulfate esters at various positions. This chemical diversity is influenced by several factors such as the algal species, the environmental conditions or the seasons (Lahaye and Robic 2007). The studied PUL of *F. agariphila* encodes 8 formylglycine-dependent sulfatases belonging to 5 subfamilies of the SulfAtlas S1 family (Fig. 2): S1_7: 3 genes; S1_8: 2 genes; S1_16: 1 gene; S1_25: 1 gene; S1_27: 1 gene (http://abims.sbroscoff.fr/sulfatlas (Barbeyron et al. 2016a)). With such a diversity of S1 subfamilies, these sulfatases likely display significant differences in substrate recognition, even though they are all predicted to act on ulvans. We expressed 7 sulfatases in soluble form in *E. coli*. After purification, these recombinant sulfatases were incubated with ulvan polymers from three different sources (Agrival, Elicityl, and one extracted from an Altantic *Ulva* sp. collected in Roscoff, France). As shown by the HPAEC analyzes of released sulfate ions, 6 sulfatases are clearly active on ulvan polymers, although their activity varies depending on the polysaccharide sources (Supplementary Fig. 4). The sulfatase P18_S1_7 (for numbering/nomenclature see Fig. 2) was most active on ulvan polymers, particularly on the xylose-rich ulvan (Supplementary Figs. 4 and 5) and can desulfate oligosaccharides containing the motif Rha3S-Xyl2S-Rha3S. Thus, this sulfatase likely proceeds in an endolytic mode of action. This assumption is consistent with the "open groove" topology of the active site unraveled by the P18_S1_7 crystal structure (Fig. 2a and 2g). Interestingly, P14_S1_7 (predicted as exolytic, since this sulfatase is almost inactive on ulvan, Supplementary Fig. 4) and P18_S1_7 (predicted as endolytic) belong to the same subfamily (S1_7). Such dissimilar modes of action within the same (sub)family have been described in glycoside hydrolase and polysaccharide lyase families (Thomas et al. 2013; Mewis et al. 2016). In comparison to P18_S1_7, the S1_25 sulfatase module of P36 (referred to as P36_S1_25) presents moderate activities on polymers. On oligosaccharides, P36_S1_25 was the most active enzyme. This enzyme specifically desulfates L-rhamnose at the 3-position and can act on the motif Rha3S-Xyl-Rha3S in an exolytic mode of action.

Sequence analyses revealed that P18_S1_7 (485 residues) and the S1_25 sulfatase module of P36_S1_25 (443 residues) are only distantly related (25% identity) and thus belong to two different SulfAtlas S1 subfamilies[20], S1_7 and S1_25, respectively. We determined the crystal structure of these two sulfatases, with higher resolution for P18_S1_7 (1.23 Å) and lower resolution for the sulfatase module of P36_S1_25 (2.91 Å). P18_S1_7 and P36_S1_25 adopt a similar fold with two α/β-structural domains, an *N*-terminal catalytic domain SD1 (Ser25-Asp388; P18_S1_7) separated by a structured linker (Arg389-Val397) from a *C*-terminal domain SD2 (Ala398-Pro483). Nonetheless, the sulfatase module of P36_S1_25 is a smaller protein and lacks some secondary elements, which are present in P18_S1_7 (the β-strands β6 and β8, the α-helices α5, α7, α8 and several short 3:10 helices). Notably, the helix α7 and the loops connecting it to the main part of SD1 constitute a protruding extension, which overhangs the active site (Fig. 2d). The active site of P18_S1_7 is a large, open groove with a strong basic character (Fig. 2a and 2g). This type of active site topology is consistent with the endo-character and its efficiency on polymeric ulvan (Supplementary Fig. 4). In contrast, the active site of P36_S1_25

is a pocket (Fig. 2c), which is consistent with its activity on oligosaccharides (Supplementary Fig. 4). The most similar protein in the Protein Data Bank (PDB) is the human iduronate 2-sulfatase (IDS, 31% sequence identity; PDB: 5FQL; Fig. 3b,e) (Demydchuk et al. 2017). Interestingly, IDS also displays a pocket active site topology (Fig. 3b). Therefore, different active sites (and subsequently different modes of action) can exist within the same S1 subfamily. Such differences in topology likely explain the varying efficiencies at the polymer level observed for P11_S1_7, P14_S1_7 and P18_S1_7 although they all belong to the S1_7 subfamily (Supplementary Fig. 4).

The catalytic machinery of the S1 family sulfatases (Hanson et al. 2004) is well conserved in P18_S1_7 and P36_S1_25. We find the catalytic nucleophile (Cys74 and Cys58, respectively), residues involved in Ca^{2+} coordination (Asp35, Asp36, Asp312 and His313; Asp18, Asp19, Asp284 and Asn285), residues stabilizing the catalytic nucleophile (Arg78 and His128; Arg62 and Gly110), and residues of the sulfate-binding S subsite, as defined in the recent nomenclature for sulfatase-binding subsites (Hettle et al. 2018) (Lys125, His213 and Lys325; Lys108, His182 and Lys297) (Fig. 3g, Supplementary Figs. 6-8). His313 in P18_S1_7 is not the most frequent residue for the coordination of the calcium ion (usually an asparagine), but a histidine at this position is found in a minority of sulfatases and is part of the updated PROSITE signature "Calcium-binding site 2" (Barbeyron et al. 2016a). Most surprising is the replacement of His128 in P18_S1_7 by Gly100 in the sulfatase module of P36_S1_25. Indeed, a histidine at this position is supposed not only to stabilize the catalytic formylglycine, but also to abstract its Oγ2 proton at the end of the catalytic cycle to induce the sulfate elimination and the aldehyde regeneration (Hanson et al. 2004). Nonetheless, this glycine is strictly conserved in the closest homologs of the sulfatase module of P36_S1_25 (105 sequences with >50% identity; Supplementary Fig. 7), suggesting that the function of the histidine at this position may not be essential in this ulvan sulfatase subgroup.

While some sulfatases were not quantified in our metabolic labeling approach (Fig. 1b and 2), they were detected by subproteome analysis in the membrane-enriched fraction. In five cases, lipoprotein signal peptides were predicted and P18_S1_7 and P36_S1_25 were highly abundant in the intracellular soluble fraction (Supplementary Data Set 3). Taken together, these results indicate a periplasmic localization of sulfatases, with some of them putatively membrane-bound. Notably, the sulfatase P36_S1_25 activity is found in a multimodular enzyme that contains also a GH78 domain. Comparative genome analyses indicated multimodular enzyme structures in the ulvan PUL H of *F. agariphila* (Salinas and French 2017) and other putative ulvan-degrading Bacteroidetes strains (Supplementary Data Set 4).



PUL H-encoded proteins (for ulvan and ulvan-derived monosaccharide use)			log_2 ratio		
Abbreviation	Locus tag	Functional annotation	fru	rha	ulv
P1_GH88	*21900	Unsaturated glucuronyl hydrolase (GH88)			
P2_SusD	*21910	SusD-like protein			
P3_TBDR	*21920	TonB-dependent receptor			
P4_HK	*21930	Histidine kinase			
P5_isomerase	*21940	4-Deoxy-L-threo-5-hexosulose-uronate ketol-isomerase			
P6_dehydrogenase	*21950	2-Deoxy-D-gluconate 3-dehydrogenase			
P7	*21960	Conserved hypothetical protein			
P8_GH2	*21970	Beta-galactosidase (GH2)			
P9_lactonase	*21980	6-Phosphogluconolactonase			
P10_PLnc	*21990	Ulvan lyase (PLnc)			
P11_S1_7	*22000	Iduronate-2-sulfatase (S1_7)			
P12_S1_8	*22010	Arylsulfatase (S1_8)			
P13_S1_16	*22020	Arylsulfatase (S1_16)			
P14_S1_7	*22030	Arylsulfatase (S1_7)			
P15_GH2	*22040	Glycoside hydrolase (GH2)			
P16_GH2	*22050	Beta-galactosidase (GH2)			
P17_GH2	*22060	Beta-galactosidase (GH2)			
P18_S1_7	*22070	Arylsulfatase (S1_7)			
P19_S1_27	*22080	Sulfatase (S1_27)			
P20_GH78	*22090	Alpha-L-rhamnosidase (GH78)			
P21_mutarotase	*22100	L-Rhamnose mutarotase			
P22	*22110	Conserved hypothetical protein			
P23	*22120	Conserved hypothetical protein			
P24_GH3	*22130	Beta-glucosidase (GH3)			
P25_SusD	*22140	SusD-like protein			
P26_TBDR	*22150	TonB-dependent receptor			
P27_GH43	*22160	Beta-xylosidase (GH43)			
P28_GH78	*22170	Alpha-L-rhamnosidase (GH78)			
P29	*22180	Conserved hypothetical protein			
P30_PL28	*22190	Ulvan lyase (PL28)			
P31_GH39	*22200	Glycoside hydrolase (GH39)			
P32_S1_8	*22210	Arylsulfatase (S1_8)			
P33_GH105	*22220	Glycoside hydrolase (GH105)			
P34_GH3	*22230	Beta-glucosidase (GH3)			
P35_oxidoreductase	*22240	Oxidoreductase			
P36_GH78/S1_25	*22250	Alpha-L-rhamnosidase/sulfatase (GH78/S1_25)			
P37	*22260	Hypothetical protein			
P38_SusD	*22270	SusD-like protein			
P39_TBDR	*22280	TonB-dependent receptor			
Non-*PUL H*-encoded proteins (for ulvan-derived monosaccharide use)			log_2 ratio		
Abbreviation	Locus tag	Functional annotation	fru	rha	ulv
NP1_dehydrogenase	*21840	Aldehyde dehydrogenase A			
NP2_dehydrogenase	*21850	L-Lactate dehydrogenase			
NP3_aldolase	*21860	Class II aldolase/adducin family protein			
NP4_kinase	*21870	Pentulose/hexulose kinase			
NP5_isomerase	*21880	Rhamnose isomerase[a]			
NP6_aldolase	*21890	Rhamnulose-1-phosphate aldolase			
NP7_kinase	*160	Xylulose kinase			
NP8_isomerase	*170	Xylose isomerase			
NP9_oxidoreductase	*9410	D-Mannonate oxidoreductase			
NP10_dehydratase	*9420	Mannonate dehydratase			
NP11_isomerase	*9430	Uronate isomerase			
NP12_kinase	*9800	2-Dehydro-3-deoxygluconate kinase			
NP13_aldolase	*9820	Aldolase[b]			
NP14_kinase	*11640	2-Dehydro-3-deoxygluconate kinase			
NP15_kinase	*16400	2-Dehydro-3-deoxygluconate kinase			
NP16_XylE	*180	D-Xylose transporter XylE[c]			
NP17_ABC	*11090	ABC transporter, ATP-binding protein			
NP18_ABC	*25150	ABC transporter, ATP-binding protein			
NP19_ABC	*7480	ABC transporter, ATP-binding protein			
NP20_ABC	*12820	ABC transporter, ATP-binding protein			

Fig. 2 List of PUL H-encoded and relevant non-PUL H-encoded proteins, corresponding locus tags and functional annotation as well as their relative abundance (mean log_2 ratio) with the respective carbon source. Empty/white squares refer to non-quantified proteins while gray squares indicate OFF proteins that could not be quantified due to a lack of ^{14}N signals (see Methods). Proteins were numbered (P1–P39, PUL H-encoded proteins; NP1–NP20, non-PUL H-encoded proteins) and protein function was indicated. In the case of GH and S, families and subfamilies were specified (Lombard et al. 2014a; Barbeyron et al. 2016a), for example GH2 (family) or S1_7 (family and subfamily). *BN863_, for example *21900 refers to locus tag BN863_21900. [a] Identified by BLAST against the Uniprot database, previously annotated as xylose isomerase-like TIM barrel domain protein. [b] 4-hydroxy-2-oxoglutarate aldolase/2-dehydro-3-deoxyphosphogluconate aldolase. [c] Only captured by sub-proteome analysis of ulvan-grown cells (Supplementary Dataset 2). fru, fructose; rha, rhamnose; ulv, ulvan.

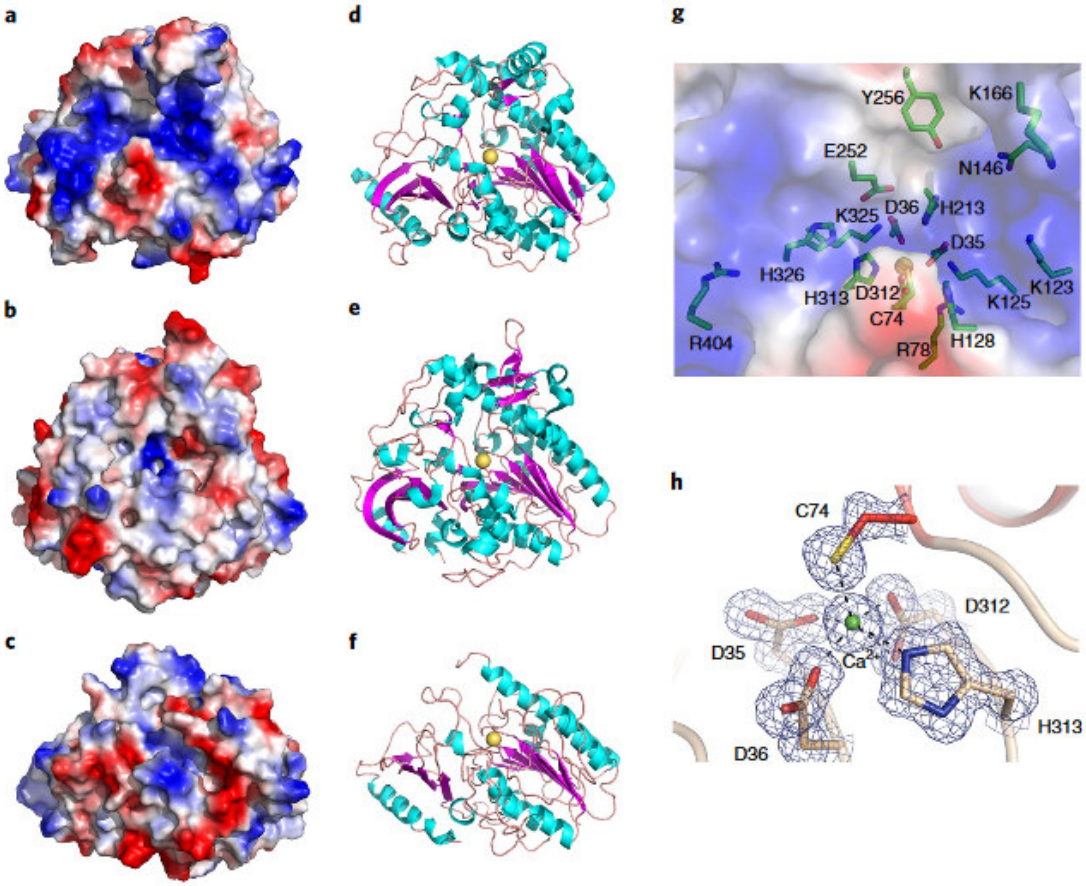


Fig. 3 structural analyses of ulvan-specific sulfatases. a-c, Molecular surface of P18_S1_7 (a) and the human iduronate 2-sulfatase (PDB: 5FQL) (**b**), both of which belong to the S1_7 subfamily, as well as that of the S1_25 sulfatase module of P36 (P36_S1_25) (**c**). These molecular surfaces are colored according to electrostatic potential ranging from deep blue, +, to red, −. **d–f**, Fold representation of P18_S1_7 (**d**), the human iduronate 2-sulfatase (**e**) and P36_S1_25 (**f**). The structures are shown in cartoon style. The α-helices and the β-strands are colored in cyan and magenta, respectively. **g**, Key conserved residues in the catalytic groove of P18_S1_7. The amino acids are presented as sticks. The calcium ion is shown as a yellow sphere. The molecular surface of P18_S1_7 is shown as semi-transparent background. **h**, Electron density around the catalytic calcium-binding site of P18_S1_7. The coordination residues (Asp35, Asp36, Asp312 and His313) and the catalytic residue Cys74 are shown as sticks. Interactions with the calcium are represented by black dashed lines. The map shown is σA-weighted $2mF_o - DF_c$ maps (with m the figure of merit and D the Sigma-A weighting factor) contoured at 1.2σ ($(0.07 \times 10)/Å^3$)

Enzymatic ulvan degradation

In brief, the distinct function of each enzyme was established by activity testing on ulvan and on defined enzymatically produced ulvan oligomers using photometric assays, fluorophore-assisted carbohydrate electrophoresis (FACE) and carbohydrate polyacrylamide gel electrophoresis (C-PAGE), high performance anionic exchange chromatography with pulsed amperometric detection (HPAEC-PAD) and mass-spectrometry. Detailed procedures of these steps are outlined in the Online Methods section.

Structures of all-important carbohydrate intermediates were confirmed by 1D and 2D nuclear magnetic resonance (NMR) spectroscopy together with mass-spectrometry analysis.

We performed an initial photometric screening, which detects the unsaturated uronic acid moiety (Δ) introduced by the lytic mechanism of lyases. We show that P10_PLnc and P30_PL28 are both endo-acting ulvan lyases generating the same product pattern, implying that they have a similar specificity (Supplementary Fig. 9). P30_PL28 accepts GlcA and IdoA at the cleavage site and generates the dimer Δ-Rha3S and the tetramer Δ-Rha3S-Xyl-Rha3S as main products (Reisky et al. 2018). Both ulvan lyases, P30_PL28 (Kopel et al. 2016) and P10_PLnc, appear to initiate ulvan depolymerization outside of the bacterial cell. P30_PL28 contains an additional ulvan-binding module (Melcher et al. 2017) and a type IX secretion system signal that drives secretion (Sato et al. 2010), corroborating the proteomic results (Supplementary Fig. 10, Supplementary Data Sets 2 and 3). P10_PLnc might be associated to the outer membrane (Supplementary Data Set 3), although a periplasmic localization is also possible (Supplementary Fig. 10, Supplementary Data Set 2).

Two variants of ulvan lyase with distinct localizations indicate synergistic functions: while P30_PL28 is an extracellular enzyme catalyzing rapid dissolution of insoluble ulvan, P10_PLnc most likely dissolves soluble ulvan oligomers at the cell surface, where uptake proceeds through the expressed TonB-dependent receptor system into the periplasm. Here, the unsaturated uronyl residue (Δ) at the non-reducing end of oligomers is removed by the exo-acting unsaturated glucuronyl hydrolases (outer membrane P1_GH88 and periplasmic P33_GH105) (Supplementary Figs. 11-13), thus forming 5-dehydro-4-deoxy-D-glucuronate. The resulting Rha3S was purified and the structure was confirmed by NMR (Supplementary Figs. 14 and 15, Supplementary Table 1). This monosaccharide is desulfated by the S1_25 sulfatase domain of P36_S1_25 yielding rhamnose, which can enter the cellular sugar metabolism (Fig. 3, Supplementary Fig. 16). Rha3S-Xyl-Rha3S was another major intermediate which was isolated (Supplementary Figs. 17 and 18, Supplementary Table 2). Rha3S-Xyl-Rha3S was desulfated by the sulfatase P36_S1_25 to yield Rha-Xyl-Rha3S, which was isolated to confirm the desulfation site at the non-reducing end (Supplementary Figs. 19-21, Supplementary Table 3). Next, Rha-Xyl-Rha3S is converted by the periplasmic P20_GH78 to Rha and Xyl-Rha3S (Fig. 4, Supplementary Fig. 22). The CBM67 domain of P20_GH78 likely elevates specificity for rhamnose and contributes to substrate recognition[28]. Finally, the dimer Xyl-Rha3S is further cleaved by P24_GH3 or P27_GH43 to yield Xyl and Rha3S, making these the first identified β-xylosidases that are active on ulvan oligosaccharides (Fig. 3, Supplementary Figs. 22 and 23). Notably, only the P24_GH3 was previously found to be active on 4-methylumbelliferyl-β-D-xylopyranoside (MUX) showing that the two enzymes have different substrate specificity at the aglycone site (Salinas and French 2017).

Besides ulvan lyases, the endo-active alpha-1,4-L-rhamnosidase GH39 cleaves rhamnose sections interspersed between xylose residues within the polymer. Such a function has, to the best of our knowledge, not been described in this family before. Accordingly, larger oligomers with consecutive

Xyl-Rha3S units that are resistant to the ulvan lyases P30_PL28 and P10_PLnc were efficiently degraded by P31_GH39 (Supplementary Fig. 24). The catalytic order of ulvan lyases and P31_GH39 was interchangeable as the larger degradation products of P31_GH39 were prime substrates for both ulvan lyases (Supplementary Fig. 25). The dimers Xyl-Rha3S and Xyl2S-Rha3S were isolated as the smallest products and the structure was elucidated by NMR, identifying GH39 as an α-rhamnosidase active on ulvan (Supplementary Figs. 26-29, Supplementary Tables 4 and 5). While Xyl-Rha3S is further degraded as described above, Xyl2S-Rha3S was resistant to P24_GH3 or P27_GH43 and needs to be desulfated by the P32_S1_8 sulfatase prior to enzymatic conversion by these enzymes (Supplementary Fig. 30). Desulfation of Xyl2S within the trimer Rha3S-Xyl2S-Rha3S, released by P30_PL28 and P33_105 digestion (Supplementary Figs. 31 and 32, Supplementary Table 6), was catalyzed by the P18_S1_7 sulfatase (Supplementary Fig. 33).

GlcA side chains present on some *O*2 residues of Rha3S[7] are removed by P17_GH2. When P17_GH2 was added to untreated ulvan, it produced a single band in FACE with the same mobility as a GlcA (Supplementary Fig. 34a) while not decreasing the overall molecular weight of the raw ulvan as seen by C-PAGE (Supplementary Fig. 9). To confirm this activity, defined oligomers with GlcA side chains were produced from ulvan with P30_PL28 and P31_GH39 with or without P33_GH105. The structure of Δ-Rha3S[2GlcA]-Xyl-Rha3S and Rha3S[2GlcA]-Xyl-Rha3S, was confirmed by NMR (Supplementary Figs. 35–38, Supplementary Tables 7 and 8) and these products were used as substrates for P17_GH2. This enzyme was also active on these smaller oligomers (Fig. 3, Supplementary Fig. 34b). This result indicates that the GlcA side chains were removed from polymeric ulvan or from smaller intermediates (Supplementary Fig. 34c), although in *F. agariphila* we predict P17_GH2 to be localized in the periplasm and thus to be active on oligomers, which also applies to P31_GH39.

GlcA side chains partially shielded the main chain against hydrolysis by P31_GH39. When the GlcA residues were removed by P17_GH2, a higher degree of degradation was observed with P31_GH39 (Supplementary Fig. 39). The newly determined crystal structure of P17_GH2 (Supplementary Fig. 40) contains a pair of *N*-terminal β-sandwich domains, a TIM-barrel with the active site, two more β-sandwich domains and a *C*-terminal putative carbohydrate-binding module connected by an extended flexible linker at the *C*-terminus that places the CBM over the active site (Supplementary Fig. 40a). The active site pocket is at the surface of the catalytic domain; its size provides just enough space to accommodate one GlcA residue. The catalytic site of this enzyme, obscured by the aforementioned CBM, further deviates from other members of the GH2 family. In most GH2 the nucleophile and acid/base catalytic residues are approximately 200 residues apart at the *C*-terminal ends of strands 4 and 7 of the conserved $(\alpha/\beta)_8$-TIM barrel fold. In P17_GH2, the nucleophile is conserved (Glu509) but the acid/base position has a tryptophan (Trp447) (Supplementary Fig. 40d). Two alternative possibilities exist for the acid/base of P17_GH2 Glu411 found on strand 3 and Asp908 from the C-terminal domain (CTD) are both approximately 6.8 Å from Glu509 and could contribute to catalyzing hydrolysis as acid/base residues (Supplementary Fig. 40c).

Monosaccharide metabolism

Ulvan degradation releases different monosaccharides to be further utilized by *F. agariphila*. Many of the enzymes involved in monosaccharide metabolism had significantly higher relative abundances with ulvan compared to fructose or rhamnose as substrate (Fig 2, Supplementary Fig. 3, Supplementary Data Set 1). Based on this result and on the MetaCyc database (Caspi et al. 2018), pathways for monosaccharide utilization were deduced, which are consistent with previously proposed pathways (Salinas and French 2017). Unlike the PUL H-encoded polysaccharide-degrading proteins, these monosaccharide-utilizing proteins are randomly distributed across the *F. agariphila* genome (Supplementary Fig. 3).

The spontaneous conversion of α- to β-anomer (mutarotation) of free α-L-rhamnose is a relatively slow process. This rate-limiting step affects growth of L-rhamnose-utilizing bacteria (Ryu et al. 2005; Richardson et al. 2008) because the first metabolic enzyme rhamnose isomerase (EC 5.3.1.14) is specific for the β-anomer (Korndörfer et al. 2000). Various bacteria, such as *E. coli* and *Rhizobium leguminosarum*, contain the L-rhamnose mutarotase, accelerating the rate of mutarotation of α- to β-L-rhamnose (Bateman 2002; Camacho et al. 2009). In contrast to the proteobacterial L-rhamnose mutarotase genes, which are part of small operons dedicated to the uptake and use of free L-rhamnose (Ryu et al. 2005; Richardson et al. 2008), the P21_mutarotase gene is localized in PUL H. We solved the crystal structure of the P21_mutarotase at 1.47 Å (Fig. 4, Supplementary Table 9) with one molecule in the asymmetric unit. P21_mutarotase adopts a ferredoxin-like fold with an antiparallel β-sheet of 4 β-strands flanked by a bundle of 3 α-helices. The P21_mutarotase structure superimposed with the characterized L-rhamnose mutarotases YiiL (PDB: 1x8d) and RhaU (PDB: 2qlw) with rmsd on C_α of 0.76 Å and 0.73 Å, respectively (Ryu et al. 2005; Richardson et al. 2008). Similar to these, the P21_mutarotase (Fig. 5a,b) formed a dimer with a large hydrophobic dimeric interface antiparallel β-sheets from each monomer (Fig. 5c). All key residues of the active site are well conserved in the P21_mutarotase (Fig. 5d,e).

F. agariphila further metabolizes the β-L-rhamnose via L-rhamnulose-1-phosphate, which is then cleaved by an aldolase (putatively NP3_ or/and NP6_aldolase, Fig. 2) into L-lactaldehyde and dihydroxyacetone phosphate (Fig. 6) (Salinas and French 2017). The corresponding genes are located directly upstream of PUL H (Supplementary Figs. 2 and 3). Glucuronic and unsaturated uronic acids are stepwise converted into KDG (2-dehydro-3-deoxy-D-gluconate), which enters the central metabolism via D-glyceraldehyde 3-phosphate and pyruvate (Fig. 6). Corresponding genes are encoded within PUL H, PUL A or elsewhere in the genome (Supplementary Figs. 3 and 41). NP8_isomerase and NP7_kinase convert D-xylose to D-xylulose-5P, which is an intermediate of the pentose phosphate pathway. In addition, putative monosaccharide transporters were identified (Fig. 6). A d-xylose transporter (NP16_XylE) was quantified in the mem- brane fraction in the subproteome experiments (Supplementary Dataset 2).

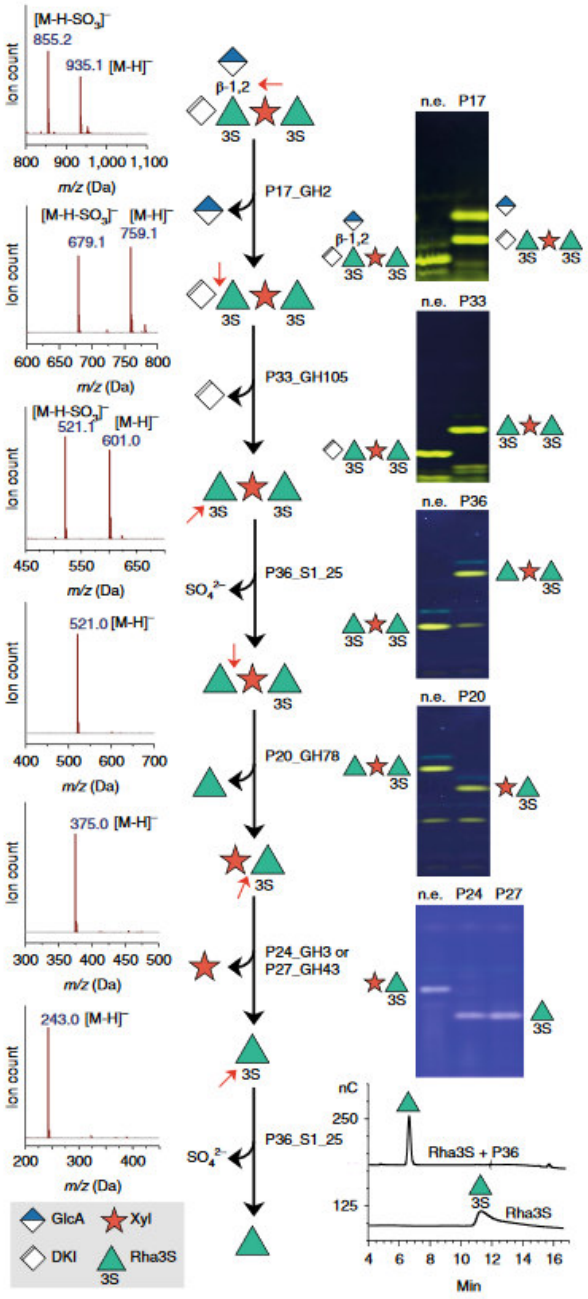


Fig. 4 Zooming into the degradation of ulvan fragments. The experimental procedure to uncover the order of enzymes for ulvan degradation is shown exemplarily for an ulvan pentamer. All other investigated enzyme activities are shown in the Supplementary Information. All intermediate products were purified and their structures were confirmed by NMR and mass spectrometry. Mass spectrometry spectra for individual oligomers are shown on the left next to the respective oligomer. Full spectra for all purified oligomers are shown in the Supplementary Information together with the corresponding NMR spectra. Red arrows indicate cleavage points of the following step. FACE gels for the analysis of the enzymatic interconversion steps are displayed on the right next to the respective enzyme. Full gel images including standards are shown in the Supplementary Information. The desulfation of Rha3S was detected by HPAEC–PAD and the full chromatograms are shown in the Supplementary Information. Numbers with 'S' attached to the sugar symbols indicate the position of sulfate groups. n.e., no additional enzyme.

Four ATP-binding proteins of ATP-binding cassette transporters were more abundant with ulvan or with rhamnose in the metabolic labeling experiments (Fig. 2 and Supplementary Dataset 1), indicating that ATP-binding cassette transporters are involved in monosaccharide uptake. Specific mono- or oligosaccharides generated by the above-described enzymatic steps were also verified by high-performance liquid chromatography–evaporative light scattering–electrospray ionization–mass spectrometry (HPLC–ELS–ESI–MS) (Supplementary Figs. 42–48).

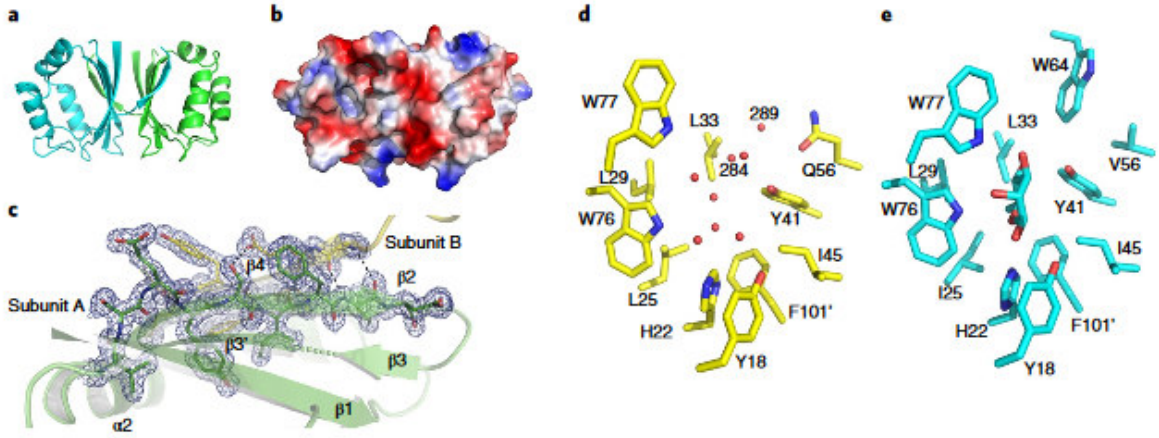


Fig. 5 Structure of the l-rhamnose mutarotase P21_mutarotase. a, P21_mutarotase dimer shown in cartoon style. **b**, Molecular surface of the P21_mutarotase dimer color coded according to electrostatic potential ranging from deep blue, +, to red, −. **c**, Electron density around the inter-subunit β-sheet in the mutarotase P21_mutarotase dimer. The β4-strand found at the C-terminal extremity of the subunit B is involved in β-sheet formation with the subunit A through hydrogen bonding with the β2-strand. Subunits A and B are green and yellow, respectively. Hydrogen bonds between β2 and β4 are shown as black dashed lines. The map shown is σA-weighted $2mF_o - DF_c$ maps contoured at 1.2σ ((0.12 × 10)/Å^3). **d,e**, Active site of P21_mutarotase (**d**) and of YiiL bound to an l-rhamnose (**e**). The amino acids are presented as sticks. The carbon atoms are colored in yellow and in cyan in P21_mutarotase and YiiL, respectively. The small red spheres are water molecules in the P21_mutarotase structure.

Discussion

Using the DNA sequence of the known ulvan polysaccharide lyase PL28 as query, 12 ulvan PULs were extracted from the NCBI-GenBank, including the biochemically characterized *F. agariphila* ulvan PUL. All PULs were from Bacteroidetes, indicating that our procedure was selective for this phylum since ulvan PULs also exist in Gammaproteobacteria (Koch et al. 2019b). Interestingly, although four ulvan PULs were from the genus *Polaribacter*, they did not cluster on the heatmap (Fig. 1c) indicating that ulvan PULs are diverse at the genus level. Also, within different ulvan PULs, PL28 or PLs from PLnc are over 50% identical at the pairwise amino acid sequence level. Conservation and invariable presence suggest that the first steps of the ulvan degradation cascade proceed through similar enzymes in these organisms. On the other hand, the GH88 enzyme was only present in ulvan PULs of *Flammeovirga*

pacifica and *F. agariphila*. GH88 is an exo-acting, unsaturated glucuronyl hydrolase. Its absence in other ulvan PULs could be compensated for by the presence of a GH105, which has the same function. Thus, the later steps in ulvan degradation proceed in dissimilar ways in bacteria.

As shown in the protein domain distribution analysis, the most abundant proteins are sulfatases, which catalyze the removal of sulfate from ulvan. Sulfatase copy numbers ranged from 4-12. At the same time, PLs or GHs such as GH2, GH78 and GH39 in the *F. agariphila* ulvan PUL were also abundant and have several copies in the other predicted ulvan PULs. Notably, some of the proteins of the ulvan PUL, such as the sulfatase P36_GH78/S1_25, are multimodular enzymes. Our analyses indicated similar domain structures of ulvan-degrading enzymes in other marine Bacteroidetes strains. However, the cursory inspection of gut *Bacteroides* genomes revealed no multimodular GH78 and sulfatase fusion proteins. This suggests that some gene fusions involved in polysaccharide degradation could be more abundant in the marine environment (Hehemann et al. 2017).

Our biochemical analyses demonstrated that six of the putative sulfatases (P11_S1_7, P12_S1_8, P18_S1_7, P19_S1_27, P32_S1_8 and P36_S1_25) are indeed ulvan-active sulfatases (Supplementary Fig. 4). However, the sulfatase P14_S1_7 was inactive on both, ulvan from Elicityl and a xylose-rich ulvan from Atlantic *Ulva* spp. and displayed only faint activity on an ulvan from Agrival. This apparent inactivity may be due to a strict exolytic character of P14_S1_7. Consequently, activity maxima are not the same for different types of ulvans. Substrate diversity may cause the variable enzyme content in Bacteroidetes (Fig. 1c). This diversity may reflect an adaptation to the different types of ulvans present in *Ulva* spp. Such fine scale adaptation points towards the exploration of PUL microdiversity as a promising avenue for enzyme discovery and for the biocatalytic elucidation of ulvan structures.

Our elucidation of the enzymatic ulvan degradation cascade and characterization of 12 of its enzymes has major implications. Firstly, the conservation of CAZyme- and sulfatase-encoding genes in ulvan PULs of different bacteria underlines their importance and provides a mean to reliably predict new ulvan degradation pathways for bioengineering. Secondly, the substantially extended knowledge of the specific substrate scope of each enzyme enables the targeted use of these enzymes for the production of a variety of novel defined, tailor-made ulvan oligomers, representing useful products, e.g., for pharmaceutical or cosmetic applications. Moreover, these enzymes provide a way to deconstruct ulvan cell walls, which may facilitate the extraction of marine poly- or oligosaccharides and other valuable molecules such as proteins from *Ulva* spp. Finally, the enzymatic cascade allows for the production of bulk monomeric sugars from the abundant, so far underexplored renewable, the green tide *Ulva*.

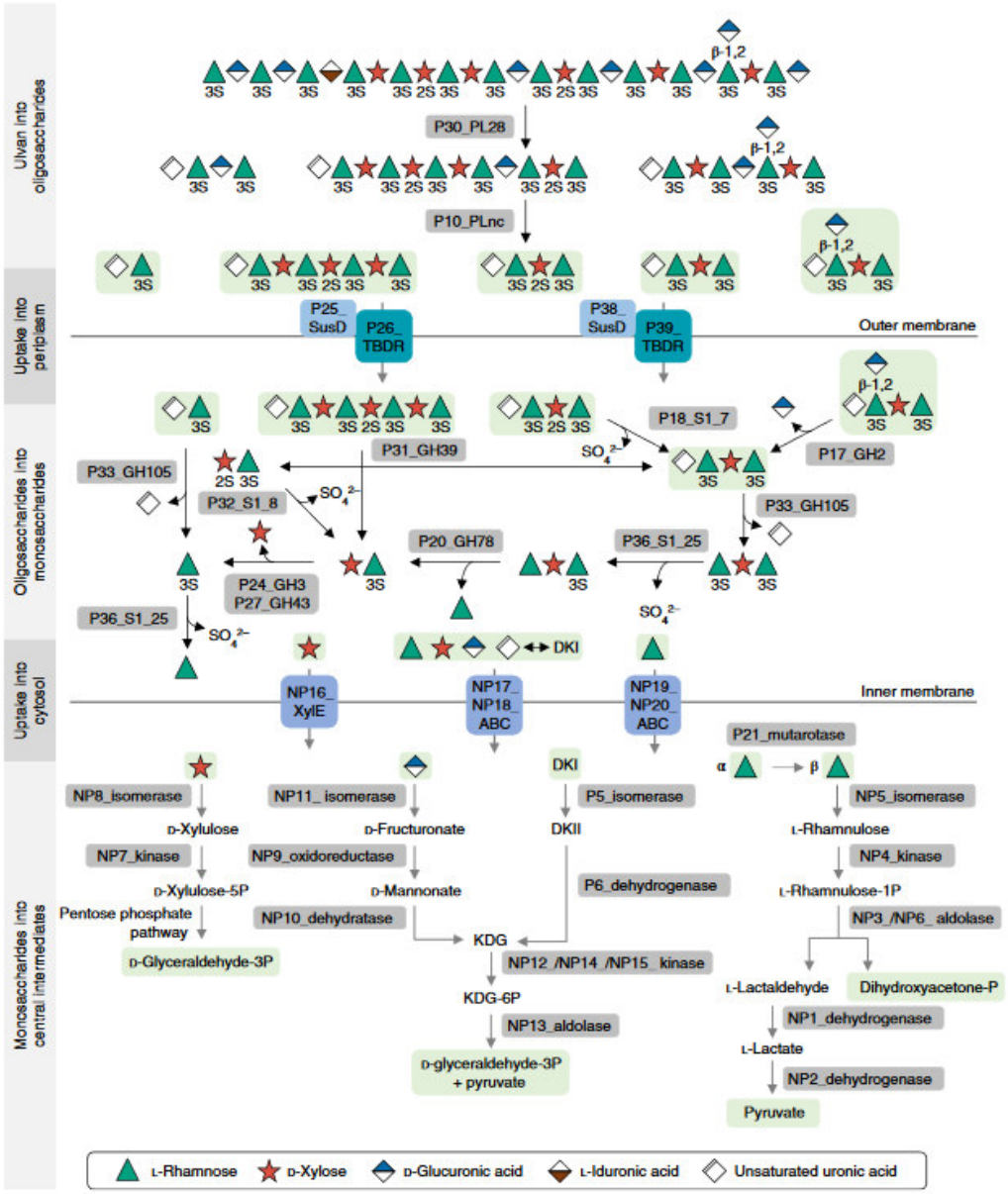


Fig. 6 Model of the ulvan degradation pathway in *F. agariphila* as suggested by the proteogenomic, biochemical and structural biological analyses in this study. Redundant pathways are omitted to maintain clarity. The ulvan molecule on top represents a part within the larger ulvan chain where rhamnose and iduronate are α- while xylose and glucuronate are β-configured. The formed products - at both ends of the initial ulvan molecule after cleavage with P30_PL28 - are not shown in the downstream degradation pathway. Activity of ulvan lyases P30_PL28 and P10_PLnc will form an unsaturated uronic acid residue from glucuronic acid or iduronic acid at the non-reducing end of the products. Numbers with 'S' attached to the sugar symbols indicate the position of sulfate groups. Black arrows indicate pathways elucidated by proteogenomic, biochemical and structural biological analyses, while gray arrows only refer to proteome analyses or additional structural analyses in the case of P21_mutarotase. For numbering and nomenclature see Fig. 2. For reasons of simplicity, the linkage of the TBDRs to the TonB-ExbBD complex or a putative membrane association of certain enzymes were not included. KDG, 2-dehydro-3-deoxy-d-gluconate; DKI, 5-dehydro-4-deoxy-d-glucuronate; DKII, 3-deoxy-d-glycero-2,5-hexodiulosonate.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41589-019-0311-9.

Author contributions

J.-H.H., T.S., G.M. and U.T.B. initiated the study and directed the project. L.R., A.P., R.L. and M.B. cloned the genes and expressed and purified the enzymes for the degradation reactions. M.B., J.-H.H. and L.R. isolated ulvan and purified oligomers. Metabolites were analyzed by C.S. via NMR and HPLC–ELS–MS for which M.D.M. provided resources. L.R. and M.B. performed biocatalyses for the analyses in gel-based assays. A.P. together with M.B. performed HPAEC–PAD analyses. M.-K.Z. with support from S.M., F.U. and A.T.-S. performed the proteome analyses for which D.B. provided the resources. N.G., C.S.R. and T.R. performed crystallographic experiments and solved the protein structures. G.M. analyzed the crystal structure of the l-rhamnose mutarotase and of the sulfatases. S.T. performed the computational analyses of PUL predictions. J.-H.H. and L.R. wrote the paper with input from U.T.B., G.M., S.M., M.-K.Z. and T.S. All authors read and approved the final manuscript.

Acknowledgements

We thank the German Research Foundation (DFG) for funding through the Research Unit FOR2406 ‘Proteogenomics of Marine Polysaccharide Utilization’ (POMPU) (by grant nos. BO 1862/17-1 to U.T.B., HE 7217/2-1 to J.-H.H. and SCHW 595/10-1 to T.S.). J.-H.H. acknowledges funding by the Emmy-Noether Program of the DFG (grant no. HE 7217/1-1). G.M. is grateful to the French National Research Agency (ANR) for its support with regards to the investment expenditure program IDEALG (grant no. ANR- 10-BTBR-04) and the Blue Enzymes project (reference ANR-14-CE19-0020-01). M.-K.Z. and F.U. were supported by scholarships from the Institute of Marine Biotechnology e.V. We thank C. Leroux for mass spectrometry analyses and M. Czjzek and A. Boraston for helpful discussions. We are indebted to the local contacts for their support during X-ray data collection at the PROXIMA-1 and PROXIMA-2 beamlines (SOLEIL Synchrotron, Saint Aubin, France) and the P11 beamline (DESY, Hamburg, Germany). We thank A. Otto, S. Junker and T. Sura for help with the metabolic labeling approach and T. Hinzke for support with analyses of the proteome data. We thank F. Lesourd (Agrival, Plouenan, France) for the gift of the ‘Agrival’ ulvan sample.

Competing interests

The authors declare no conflict of interest.

Methods

Prediction of ulvan PUL. A total of 118,981 bacterial genomes were downloaded from the NCBI-GenBank using an in-house script (updated in 9 October 2018). Hmmer v.3.0 was used to identify proteins with a PL28 or sulfatase domain, using a cut-off value of 1 × 10−10 (Finn et al. 2011). Hidden Markov models of PL28 and sulfatase were obtained from dbCAN2 and the pfam database, respectively (Bateman 2002; Zhang et al. 2018). Models for the new PLnc family have not been released, thus blastp was used to identify its homologs, using 1 × 10−50 and 30% sequence identity as cut-off values (Camacho et al. 2009). In each bacterial genome, if the adjacent 50 proteins to the aforementioned marker genes contained three marker genes (PL28, PLnc and sulfatase), this locus was considered as a potential ulvan PUL hit. To further determine PUL boundaries, 100 proteins surrounding the predicted ulvan PUL were collected and then locally annotated using pfam and dbCAN Hidden Marikov models (cut-off value 1 × 10−10). First, PL28 or PLnc families were set as boundaries, which were extended if adjacent genes are annotated as sugar utilization proteins, such as GH, PL, sugar transporter and transcription factors. In cases where five continuous genes were not related to sugar utilization or ulvan degradation, the last functionally relevant protein was taken as the putative ulvan PUL boundary. Protein sequences within putative ulvan PUL were collected for further analysis. Circos was used to visualize the different ulvan PUL (Connors et al. 2009). Blastp was used to calculate the identity between PL28 sequences from different ulvan PUL (cut-off value: 1 × 10−10, over 50% identity). To simplify and reduce non-conserved proteins, domains with less than 80% presence among the predicted ulvan PUL were excluded. Domain numbers in each PUL were counted, summarized and displayed in R studio.

Proteome analyses. Whole cell proteome—metabolic labeling. A $^{14}N/^{15}N$ relative quantification approach, based on metabolic labeling, was used for protein quantification as described previously (MacCoss et al. 2003). For this purpose, *F. agariphila* KMM 3901^T was cultivated in a synthetic seawater medium (MPM) (Schut et al. 1993) containing either ^{14}N- or ^{15}N-ammonium chloride, supplemented with 0.2% of the individual carbon source: ulvan, rhamnose or fructose. Cultivation (21 °C, 170 r.p.m.) comprised three steps: (1) 24 h of marine broth 2216-cultivation and subsequent (2) pre-cultures as well as (3) main cultures in the above-described minimal medium. At an optical density (OD600 nm) of 0.5, cells were collected from main cultures by centrifugation (30 min, 9,384g, 4 °C). Cell pellets were suspended in TE-buffer (10 mM Tris, 10 mM EDTA) and cells were disrupted by sonication (four cycles of 25 s). Cell debris and protein extract were separated by centrifugation (10 min, 21,460g, 4 °C). In case of the ^{15}N-labeled samples, protein extracts of all samples from all three carbon sources were combined to form the ^{15}N-labeled reference pool, which served as an internal

standard. Protein concentration was determined using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). Then, 12.5 µg of protein of each ^{14}N-sample was combined with 12.5 µg protein of the ^{15}N-labeled pool. These mixtures were separated by one-dimensional SDS–PAGE. Protein lanes were cut into ten equal-sized pieces, destained and proteins were in-gel-digested with a 1 µg ml^{-1} trypsin solution (Grube et al. 2015). Peptides were separated by reversed-phase chromatography and analyzed in an LTQ-Orbitrap Classic mass spectrometer equipped with a nanoelectrospray ion source (Otto et al. 2010). Data represent three independent experiments (n = 3).Mass spectrometry data were searched with Sorcerer SEQUEST v.27, rev.11 (Thermo-Finnigan, Thermo Fisher Scientific) against a target decoy database including all *F. agariphila* KMM 3901^T protein sequences, corresponding reversed sequences (decoys) as well as common laboratory contaminants (total 7,224 entries) as described previously (Otto et al. 2010), but using a false positive rate of 0.05. In brief, peak intensities of the ^{14}N-peptide ions of a protein versus its ^{15}N-peptide ions were compared to calculate a regression ratio. Only unique peptides and peptides with an R^2 above 0.7 were taken into account. Non-quantified peptides were manually checked. Average regression ratios were then exported. Proteins with at least two quantified peptides were considered for the following calculations: ratios were median-centered and log-transformed, termed as $\log_2$ ratios, per sample. If proteins were quantified in at least two of the three replicates, means and s.d. were calculated from these values. To identify relative changes between the different carbon sources, $\log_2$ ratios of fructose- or rhamnose-cultivated cells were subtracted from $\log_2$ ratios of ulvan-cultivated cells, termed $\Delta\log_2$. Fold-changes correspond to the exponentials of these $\Delta\log_2$ values. Statistical analyses were performed with Welch's two-sided t-test (permutation-based false discovery rate 0.01) using Perseus v.1.6.0.7 (Tyanova and Cox 2018), based on the $\log_2$ ratios. Putative ON/OFF proteins were marked with ^{15}N (OFF) or ^{14}N (ON) in Supplementary Dataset 1 but were not included in any of the calculations. Only if a protein was identified as an ON/OFF protein in all three replicates, it was assigned to a fixed value (10/−10), to highlight these proteins.

Subproteome fractionation. F. agariphila KMM 3901^T was cultivated as described above, except that no ^{15}N-labeling was performed and only ulvan was applied as a carbon source. For the surface proteome (trypsin-shaving approach), 1.5 ml of cell suspension was removed from the culture and centrifuged (5 min, 5,867g, 4 °C). Cells were washed with 50 mM triethylammoniumbicarbonate-buffer and finally resuspended in 45 µl triethylammoniumbicarbonate-buffer. To cleave proteins from the cell surface, 5 µl of a 1 µg ml^{-1} trypsin solution was added. The solution was transferred onto a 0.22 µm cellulose-acetate spin-column and incubated for 15 min at 900 r.p.m. and 37 °C. The flow-through was collected by centrifugation (10 min, 4,000g, 4 °C), another 1 µl of trypsin was added and the sample was incubated at 900 r.p.m. and 37 °C overnight. The peptide mixture was desalted using C18 StageTips. The following solutions were used: 0.1% (v/v) acetic acid in ultra-pure water (buffer A) and 0.1% (v/v) acetic acid in acetonitrile (buffer B). Before the sample was added, C18 material was rinsed and equilibrated with buffer A and washed with buffer B in between these steps. After the sample was added, buffer A was used for washing and buffer B for elution. In the case of cytosolic, membrane-associated and

extracellular protein fractions, cells from 100 ml of culture were collected by centrifugation (30 min, 9,384g, 4 °C). Cell pellets and supernatants were processed separately as previously described (Antelmann et al. 2001; Eymann et al. 2004). One-dimensional-SDS–PAGE, in-gel-digestion and liquid chromatography–tandem mass spectrometry analysis were performed as described above. Experiments were carried out in triplicates (n = 3). Database searches were done with Sorcerer SEQUEST v.27, rev. 11 (see above). Results were summarized and filtered using Scaffold v.4.4.1.1 (Proteome Software): protein and peptide false discovery rate was set to 0.01 and protein identification required two peptides minimum. For protein quantification, the normalized spectral abundance factor was calculated for each protein giving the percentage (%NSAF) of all proteins in the same sample (Zybailov et al. 2006). If proteins were identified in at least two of the three replicates, they were considered for further calculations.

Gene cloning and expression. Expression constructs were prepared using the FastCloning strategy (Li et al. 2011) with genomic DNA from *F. agariphila* KMM 3901^T (collection number DSM15362 at DSMZ) as template for the amplification of the inserts. Generally, the pET28 constructs were prepared as described previously (Reisky et al. 2018) with the gene primers shown in Supplementary Table 10. To clone the gene for the formylglycine-generating enzyme from *F. agariphila*, the vector backbone was amplified with the primers 5′-AATA GCGC CGTC GACC ATCA TCAT CATC ATCAT-3′ and 5′-CATG GTTA ATTC CTCC TGTT AGCC CAAA AA-3′ from pBAD/myc-his A. For the pFA, constructs were cloned and overexpressed as previously described (Groisillier et al. 2010). Briefly, genes were PCR-amplified using the NEB Q5 High-Fidelity DNA Polymerase system. PCR reactions were done with 30 cycles (denaturation: 95 °C; annealing: 60 °C; elongation: 72 °C) using 0.5 units of enzyme in a total reaction of 50 µl using the primers shown in Supplementary Table 10. Amplicons were cleaned up using the QIAquick PCR Purification Kit (Qiagen) and digested with the appropriate restriction endonucleases. All ligations were done in the linearized T7 system vector pFO4. Genes encoding the sulfatases P18_S1_7, P19_S1_27, P32_S1_8 and P36_S1_25 were ordered codon-optimized for E. coli and sub-cloned into pET28 with NheI and XhoI from Genscript. The optimized nucleotide sequences are shown in the Supplementary Information. *E. coli* BL21(DE3) was transformed with pET28-based plasmids. For the overexpression, 50 ml ZYP-5052 (Studier 2005) with 100 µg ml^{-1} kanamycin were inoculated from an overnight culture in LB medium containing 50 µg ml−1 kanamycin. The culture was grown at 30 °C and 180 r.p.m. until the OD600 nm reached 1.0 and was then cooled to 20 °C for 48 h. In the case of sulfatases, the formylglycine-generating enzyme from *F. agariphila* was co-expressed. LB medium with 100 µg ml^{-1} ampicillin and 50 µg ml−1 kanamycin was inoculated from an overnight culture in the same medium and incubated at 37 °C and 180 r.p.m. until the OD600 nm reached 0.3 to 0.5. After the addition of 1.5 mM l-arabinose and incubation for 90 min at 37 °C, the culture was cooled to 18 °C for 2 h before 0.5 mM isopropyl-β-d-1-thiogalactopyranoside (IPTG) was added and the culture was incubated overnight at 18 °C. Alternatively, sulfatases were expressed from the pFA constructs in E. coli BL21(DE3) cells grown in LB medium supplemented with 15 µg ml^{-1} ampicillin, at 37 °C, until

reaching an OD600 nm of 0.8. Expression was induced with 0.1 mM IPTG overnight at 18 °C. For crystallization screening, *E. coli* BL21(DE3) cells were transformed with the plasmids containing the gene fragment of interest, then grown in the autoinduction ZYP-5052 medium (200 µg ml−1 ampicillin, 20 °C, 72 h). Cells were collected by centrifugation (10,000g, 4 °C, 20 min) and the cell pellets were stored at −20 °C until further use. Samples from the cultivations equivalent to 1 ml of culture with an OD600 nm of 7 were taken before collection and the cells were collected by centrifugation (13,000g, 4 °C, 2 min). Pellets were resuspended in 500 µl 50 mM HEPES with 100 mM NaCl (pH 7.4). After chemical lysis with BugBuster (Merck), whole cell protein (W) samples were obtained before removal of the cell debris by centrifugation (13,000g, 4 °C, 10 min). Samples of the soluble protein fraction were taken from the respective supernatant.

Enzyme purification. Cell pellets were thawed and resuspended in 50 mM NaPi with 300 mM NaCl (pH 8.0) and lysed by three cycles of sonication (2.0 min, 30% pulse, 50% power). After centrifugation (10,000g, 4 °C, 20 min), the supernatant was filtered (0.45 µm) and loaded onto a 5 ml HisTrap FF crude column (GE Healthcare) equilibrated with lysis buffer. Alternatively, Rotigarose-His/Ni beads (Karl Roth) were used in gravity flow columns. After washing, the protein was eluted with 50 mM NaPi and 300 mM NaCl containing 300 mM imidazole (pH 8.0). Fractions containing the protein of interest were pooled and desalted using PD-10 columns (GE Healthcare) equilibrated with 50 mM NaPi pH 7.4. Proteins were analyzed by SDS–PAGE on 12.5% acrylamide gels. Then, 1% (v/v) 2,2,2-trichloroethanol was used for the visualization of proteins under ultraviolet light (Ladner et al. 2004). Alternatively, proteins were stained with Coomassie Blue (PhastGel Blue R). All enzymes were used undiluted, or in dilutions of 1:5, 1:10 or 1:20 with enzyme storage buffer (Supplementary Table 11). Alternatively, cells were subjected to mechanical lysis and cytoplasmic extracts were loaded onto an HisTrap column (5 ml, GE Healthcare) equilibrated with 50 mM Tris, 0.2 M NaCl, 20 mM imidazol, 1 mM $CaCl_2$ at pH 8.0. Recombinant proteins were eluted with around 250 mM imidazole and then loaded onto a Hiprep Desalting column (26/10, 53 ml, GE Healthcare) to eliminate the imidazole, which notably interfered with sulfatase activity. Purified enzymes were concentrated (Amicon Ultra Centrifugal Filter, 30 kDa) to a concentration of 1 mg ml^{-1} (Nanodrop).

Purification of ulvan. Green tide *Ulva* sp. was collected near Roscoff (France) and dried. Alternatively, dried Ulva biomass from the Atlantic coast in Spain was purchased as organic sea lettuce (Kulau). Ulvan was extracted according to the literature (Robic et al. 2009). The dialysis step was exchanged by precipitation with acetone (80% (v/v) final concentration). After washing, acetone-precipitated ulvan was dissolved in deionized water and freeze-dried. Alternatively, ulvan was obtained from Agrival or Elicityl. Enzyme assays. Generally, reactions were performed in 50 mM HEPES pH 7.4 with 100 mM NaCl or 35 mM Tris pH 8.0 with 50 mM NaCl. Initial degradation of ulvan into larger oligomers was monitored by C–PAGE, while smaller degradation products and the conversion of purified oligomers was analyzed by FACE. For lyases, the increase in absorbance at 235 nm was recorded over time. For unsaturated uronyl hydrolases (GH88 and GH105), the decrease in absorbance at 235 nm of ulvan lyase

products was monitored. For screening reactions, 10% (v/v) clarified lysate as used for the SDS–PAGE was added. Untreated ulvan was generally used at a concentration of 1 g l^{-1} while purified oligomers were used at 0.25 gl^{-1}. Incubation was performed overnight at room temperature.

Sulfatase activity assay on ulvan polymers. Activity assays were conducted on three different ulvan polymers from Ulva species: a commercial ulvan from Elicityl, an ulvan that was a gift from the company Agrival, and an ulvan extracted from Ulva sp. collected on the north cost of Brittany (Roscoff). Then, 10 µl of each ulvan solution (1% w/v in H_2O) was incubated with 10 µl of purified sulfatase (1 mg ml^{-1}) in a final volume of 80 µl of 25 mM Tris-HCl, 0.1 M NaCl, 0.5 mM $CaCl_2$, pH 8.0 buffer mix, for 18 h at 37 °C. For each reaction, a control sample was prepared using similar conditions but with an inactivated enzyme (100 °C, 10 min). Reaction mixtures and blanks were then filtered (10 kDa, Amicon Ultra, Millipore) to measure the amount of free sulfate in the filtrates. Ulvan-specific sulfatase activity was measured by HPAEC. Using an ICS5000 system (Thermo Scientific Dionex), anions from reaction mixture filtrates were injected (AS–AP Autosampler) and separated using an AG11-HC guard column (4 × 50 mm) mounted in series with an AS11-HC anion-exchange column (4 × 250 mm). Elutions were performed with isocratic 12 mM NaOH at a flow rate of 1 ml min^{-1} (Single Pump-5), and the detection of anions was leaded by an Analytical CD Conductivity Detector associated to a suppressor (ASRS 500, 4 mm) running at 50 mA. Using a standard curve of sulfate, concentration of sulfate released by the enzymatic reaction was calculated from the difference of the amount of sulfate between samples and the associated blanks.

Sulfatase activity assay on characterized ulvan oligosaccharides. Here, 10 µl of ulvan oligosaccharides (0.5–1% w/v in H_2O) were incubated with 15 µl of purified sulfatase (0.5 mg ml^{-1}) in a final volume of 75 µl of 5 mM Tris-HCl, 10 mM NaCl, 0.5 mM $CaCl_2$, pH 8.0 buffer, for 18 h at 37 °C. The recombinant enzymes P33_GH105 or P36_GH78 were added (2 µl, 3 mg ml^{-1}). Each reaction mixture was centrifuged (14,000g for 10 min) before injection. Oligosaccharide detection was realized by HPAEC analyzes on the same ICS 5000 system described for the sulfate quantification. Elutions were performed at a flow rate of 0.5 ml min^{-1} using a NaOH multistep gradient from 8 to 280 mM (45 min). Oligosaccharides were detected by conductivity mode under a current suppression of 50–300 mA.

Carbohydrate electrophoresis. FACE was performed with 2-aminoacridone as a fluorophore (Hehemann et al. 2010). For C–PAGE, samples were mixed with an equal volume of FACE loading buffer (Hehemann et al. 2010). Gels and running conditions were identical to FACE. Carbohydrates were visualized by staining with Stains-All solution (0.25 g l^{-1} in 1.7 mM Tris-HCl pH 7.5 with 25% (v/v) isopropanol). The gels were destained with 25% (v/v) isopropanol in deionized water.

Purification of oligomers and structure determination. Ulvan was digested with purified enzymes in Tris-HCl pH 8.5 at room temperature. Oligomers were separated on two XK 26/100 (GE Healthcare) in series filled with Bio-Gel P-2 (Rio- Rad) using 100 mM $(NH_4)_2CO_3$ as mobile phase at a flow rate of 1 ml min^{-1}. After lyophilization of the fractions containing the products, oligomers were dissolved in D_2O

and lyophilized three times before NMR spectra were recorded on a Bruker Avance III HD 600 (600 MHz) spectrometer (Bruker) in D_2O solutions. The structures were independently elucidated based on one- and two-dimensional (correlated spectrometry, heteronuclear single quantum coherence, heteronuclear multiple bond correlation and total correlated spectrometry) methods, and the assigned ^{1}H and ^{13}C-NMR signals were then compared with literature data, showing excellent consistency (Lahaye et al. 1997; Lahaye 1998). For samples containing uronic acid structures, it was required to neutralize the otherwise acidic NMR samples with Na_2HPO_4 to pH 7–8 (pH-electrode calibrated to H^+) to achieve fully resolved signals for the carboxylic acid and neighboring positions (^{13}C). HPLC–ELS–MS analysis was performed by injection of ~0.1% solutions (1–5 µl) on a Nexera ultra-HPLC system from Shimadzu (equipped with two binary LC-30AD pumps plus degassers, a CTO-20 column oven) and a LC–MS-2200 EV mass spectrometry detector and an additional evaporative light scattering detector (JASCO ELS-2041). Analysis was performed with mobile phase A = H_2O (0.1% HCOOH) and mobile phase B = CH3CN on a C18 column (XSelect CSH XP C18 2.5 µm 3 × 50 mm) at 40 °C. Flow rate was 1.3 ml min^{-1} (0–3 min) with 5% B from 0–0.15 min, 5–98% B from 0.15–2.2 min and 98–5% B from 2.2–2.5 min.

Crystallization of proteins and structure determination. Crystallization trials of P18_S1_7 (pFA13 construct) and of the family S1_25 sulfatase module of the bimodular GH78 l-rhamnosidase P36 (pET28 construct, referred to as P36_S1_25) were undertaken at room temperature using the vapor-diffusion method in sitting drops containing a 2:1 ratio of pure protein (12.9 and 13.0 mg ml^{-1}, respectively) and of precipitant solution. P18_S1_7 and P36_S1_25 were mixed with reservoir solution containing 100 mM MIB pH 5.0 and 25% PEG 1,500 and 100 mM MES pH 6.5 and 25% PEG 2,000 MME, respectively. Crystals of the l-rhamnose mutarotase P21_mutarotase (pFA16 construct, concentration: 14.9 mg ml^{-1}) were obtained by the hanging-drop vapor-diffusion method at room temperature and also at a 2:1 protein:precipitant ratio with a reservoir solution containing 100 mM sodium acetate pH 4.6 and 4.3 M sodium formate. Crystals of P18_S1_7, P21_mutarotase and P36_S1_25 were cryo-protected with 10, 14 and 14% glycerol, respectively, and flash-frozen in liquid nitrogen. X-ray diffraction experiments were carried out at 100 K at beamlines PROXIMA-1 (PX1) for P18_S1_7 and P21_ mutarotase and PROXIMA-2 for P36_S1_25 (SOLEIL Synchrotron). Diffraction data of P18_S1_7, P21_mutarotase and P36_S1_25 were obtained at 1.23, 1.47 and 2.91 Å, respectively, and were processed using XDS (Kabsch et al. 2010). Scaling and merging were performed using the program Aimless from the CCP4 package(Winn et al. 2011). The structure of P21_mutarotase (a dimer of 2 × 115 residues), P18_S1_7 (475 residues) and P36_S1_25 (467 residues) were solved by molecular replacement with the CCP4 suite program MolRep56 using the structures of the rhamnose mutarotase RhaU from *R. leguminosarum* (PDB: 2QLX) (Richardson et al. 2008), of the human iduronate-2-sulfatase (PDB: 5FQL) (Demydchuk et al. 2017) and of the putative sulfatase YidJ from Bacteroides fragilis (PDB: 2QZU) as starting models, respectively. Refinement and model building of P18_S1_7 and P21_mutarotase were undertaken using the PHENIX program suite57 and the Coot software

(Emsley et al. 2010). Initial refinement of the P36_S1_25 structure was performed with BUSTER (Smart et al. 2012) and PHENIX (Adams et al. 2010), and then manual examination and rebuilding of the refined coordinates were carried out in Coot (Emsley et al. 2010). Structural validation was undertaken using MOLPROBITY (Chen et al. 2010). SEC-purified P17_GH2 crystallized in a 1:1 ratio of 7 mg ml^{-1} protein in 20 mM Tris pH 8.0 and mother liquor in the JBScreen PACT^{++} HTS and JBScreen Classic HTS I (Jena Bioscience). A single crystal from the screen grown in 20 % PEG 3350, 0.1 M Bis-Tris pH 7.5, 0.2 M sodium bromide was cryo-protected in 30 % glycerol before X-ray crystallography. The diffraction data were collected at DESY P11 automatically integrated in XDS and scaled and merged in Aimless (Hehemann et al. 2010; Evans and Murshudov 2013). The structure of P17_GH2 was solved by molecular replacement using PDB: 5dmy as a search model in Phaser (McCoy et al. 2007). The structure was built automatically using buccaneer and manually in Coot building directly into the $2F_o - F_c$ maps (Cowtan 2006; Emsley et al. 2010). Structural validation was carried out using MOLPROBITY (Chen et al. 2010).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data that support the findings of this study are available from the corresponding authors upon reasonable request. The protein structures are deposited in the PDB under 6HHM, 6HHN, 6HPD and 6HR5. Mass spectrometry data were deposited to the ProteomeXchange Consortium via the PRIDE partner repository (Vizcaíno et al. 2016) with the dataset identifier PXD009299.

Chapter 1.2. Novel carbohydrate-binding module from *Formosa agariphila* KMM3901^T binds to macroalgal polysaccharides

Affiliations

Nadine Gerlach[1,2], Chiara Vanni[2,3], Jan-Hendrik Hehemann[1,2,*]

[1] MARUM – Center for Marine Environmental Sciences, University of Bremen, Germany

[2] Max Planck Institute for Marine Microbiology, Bremen, Germany

[3] Jacobs University Bremen gGmbH, Bremen, Germany

*Correspondence: jhhehemann@marum.de, +494212186775

To be submitted in Protein Science

Abstract

Microbial utilization of glycans involves an individual enzymatic cascade reflecting the structural complexity of the substrate. The enzymatic machinery includes carbohydrate-active enzymes (CAZymes), receptors, transporters and binding proteins. CAZymes like glycoside hydrolases and polysaccharide lyases often possess non-catalytic domains with auxiliary activities which enhance the enzyme function. Glycan recognition and binding is a crucial step in polysaccharide utilization and is intensively studied for microbes found in the intestinal of ruminants. In a diluted environment like the oceans, little is known how marine microbes interact with their substrates. Here, we describe the carbohydrate-binding module (CBM) of an exo-acting β-glucuronidase involved in the degradation of the sulfated polysaccharide ulvan by *Formosa agariphila* KMM3901^T. We combined biochemical analyses of more than 30 terrestrial and marine polysaccharides with X-ray crystallography to shed light on its mechanistic function. FaGH2CBM is located at the enzyme's C-terminus and characterized by a versatile binding motif including ulvan and other marine polysaccharides from macroalgae such as fucoidans. Our study further contributes to the enzymatic pathway of ulvan utilization by marine microbes.

Key words: algal polysaccharide, X-ray crystallography, CBM, ulvan, fucoidan, binding protein

Introduction

Marine macroalgae consist of up to 70 % of polysaccharides, which are mainly found in the cell walls maintaining the 3D structure (Kraan 2012). Algal glycans are remineralized by heterotrophic bacteria (Teeling et al. 2012; Teeling et al. 2016) and introduced into the marine carbon cycle. Ulvan is the main cell wall polysaccharide of the algae order Ulvales belonging to the globally distributed green algae (Chlorophyta). Ulvan is composed of repeating disaccharide units of D-glucuronic acid (GlcA) and L-iduronic acid (IdoA) linked to 3-sulfated rhamnose (R3S). These building blocks are termed ulvanobiuronic acid A or B, respectively. There are also disaccharides of RS3 linked to D-xylose (Xyl) found in lower amounts. The degree of sulfation varies, with a maximum of 30 % (Lahaye and Robic 2007). Ulvan is a water-soluble, anionic polysaccharide contributing up to 29 % of algal dry weight (Lahaye and Robic 2007) and therefore serving as an important food and energy source for marine heterotrophic bacteria. Recently, it has been shown that heterotrophic bacteria require a multistep utilization process which involves several sulfatases and carbohydrate-active enzymes (CAZymes) to degrade ulvan into fermentable monosaccharides (Reisky et al. 2019). *Formosa agariphila* KMM3901^T is a highly adapted flavobacterium possessing 13 polysaccharide utilization loci (PULs) to maintain an algae-associated lifestyle. PUL H consist of receptors, transporters, binding proteins, sulfatases and CAZymes for ulvan utilization (Mann et al. 2013). Out of these 39 enzymes, 12 enzymes of 11 families

act in a distinct order to perform the complete degradation of ulvan (Reisky et al. 2019). The structural analysis of a β-glucuronidase of glycoside hydrolase (GH) family 2 revealed a putative C-terminal carbohydrate binding module (CBM). P17_GH2 is an exo-acting enzyme which cleaves β-1,2-linked glucuronic acid from Rha3S[GlcA]-Xyl-Rha3S (Reisky et al. 2019). Many CAZymes are multimodular enzymes, which often possess non-catalytic domains to support the enzyme function (Ficko-Blean and Boraston 2012). These auxiliary activities are frequently involved in substrate recognition and binding (Boraston et al. 2004).

CBMs are defined as contiguous amino acid sequences with discrete folds within the modular structures of CAZYmes and cellulosomal scaffolding proteins with carbohydrate-binding activity. The initial discovery classified CBMs as cellulose-binding domains (CBDs) due to binding of several GH modules to cellulose (Tomme et al. 1996). Since then, CBDs were reclassified after other specifications were found. CBMs consist of 30 to 200 amino acids and can be located at both termini of a protein (Shoseyov et al. 2006). These modules can have ligand specificity of poly- and oligosaccharides because the carbohydrate-recognition site mirrors the solution confirmation of the targeted substrate and thereby minimizes the energetic costs upon binding. At present, CBMs are divided into 81 families (Lombard et al. 2014a) based on amino acid similarities (Tomme et al. 1996), fold or structural and functional similarities such as surface-binding, glycan-chain-binding, small-sugar-binding (Boraston et al. 2004).

CBMs play a pivotal role in polysaccharide-degrading enzymes (Boraston et al. 2004). For many CAZyme families putative CBMs are reported but overall, they are not well studied yet. To date, there is only one study reporting an ulvan-binding domain located at the C-terminus of an endo-acting ulvan lyase from *Nonlabens ulvanivorans* (Melcher et al. 2017). In the present work, we performed biochemical and structural analyses of a binding protein originated from an exo-acting β-glucuronidase of family GH2. Our biochemical and structural analyses add on to previous studies elucidating ulvan utilization pathway of marine microbes.

Results and discussion

Sequence analysis and global abundance.

Structural analysis of an exo-acting β-glucuronidase of family GH2 from *Formosa agariphila* KMM3901^T identified a putative carbohydrate binding domain located at the C-terminus (Reisky et al. 2019). Since no structural homologue is found in the PDB, we analyzed the sequence of FaGH2CBM in more detail. We performed a pBLAST in NCBI and five protein databases of microbial metagenomes (MAR). In total, 261 hits were analyzed varying from 98 % (100 % coverage) to 27 % (70 % coverage) sequence identity. Out of these hits 57 % were annotated as hypothetical protein, 16 % as GH2, 10 % as DUF4982, 4 % as beta-galactosidase and cytochrome C (Fig. S1A). FaGH2CBM-like sequences were mainly found in the superphylum Planctomycetes-Verrucomicrobia-Chlamydiae (PCV) and the phylum

Bacteroidetes together accounting for 73 % of all hits. Additionally, FaGH2CBM-like sequences were also detected in archaea and the phyla Proteobacteria as well as Chloroflexi, Actinobacteria and Firmicutes of the Terrabacteria group (Fig. S1B). Closest similarities were found in GH2s from *Tamlana* sp., *Jejuia pallidilutea* and *Algibacter pectinivorans* of the order Flavobacteriales. Comparing the hits from NCBI and MAR suggested that FaGH2CBM-like sequences from the marine environment are highly enriched in PVC (Fig. S1C). Various studies demonstrated the potential of marine PVC members as sulfated-glycan degrader (Reintjes et al. 2017; Reintjes et al. 2018; Wiegand et al. 2020). They are equipped with specialized sulfatases and CAZymes to utilize complex polysaccharides such as chondroitin or fucoidan (Wegner et al. 2013; Kim et al. 2016; van Vliet et al. 2019; Sichert et al. 2020).

FaGH2CBM was detected in 31 metagenomes from the environmental surveys of TARA Oceans (31 samples) and Malaspina (6 samples) accounting to 0.0002 - 0.0004 % and 0.0006 % of the total amount of ORFs, respectively. Overall, FaGH2CBM-like sequences were found in surface waters of the Atlantic, Pacific and Indian Ocean in temperate, subtropical and tropical regions (Fig. 1).

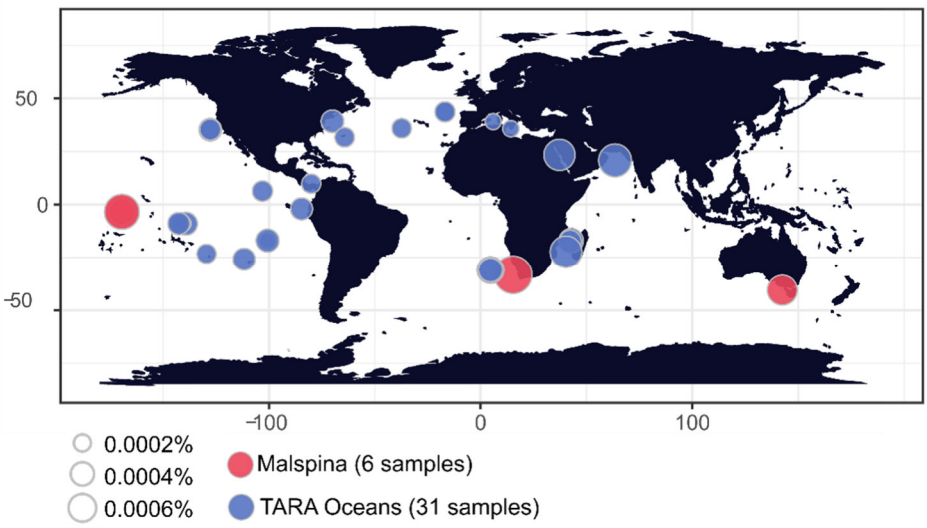


Figure 1. Distribution and abundance of FaGH2CBM in global metagenomes. Relative abundance is shown as % of the total metagenome of ORF's in each sample.

Structural insights.

FaGH2CBM consists of 143 amino acids with a theoretical isoeletric point (pI) of 5.5. We performed co-crystallization of FaGH2CBM with its putative ligand resulting from a multistep enzymatic hydrolysis of ulvan. Prior crystallization, the thermal stability and the oligomerization state were determined. FaGH2CBM is stable until 25 °C (Fig. 2A) and assembles as a monomer of about 16 kDa (Fig. 2B).

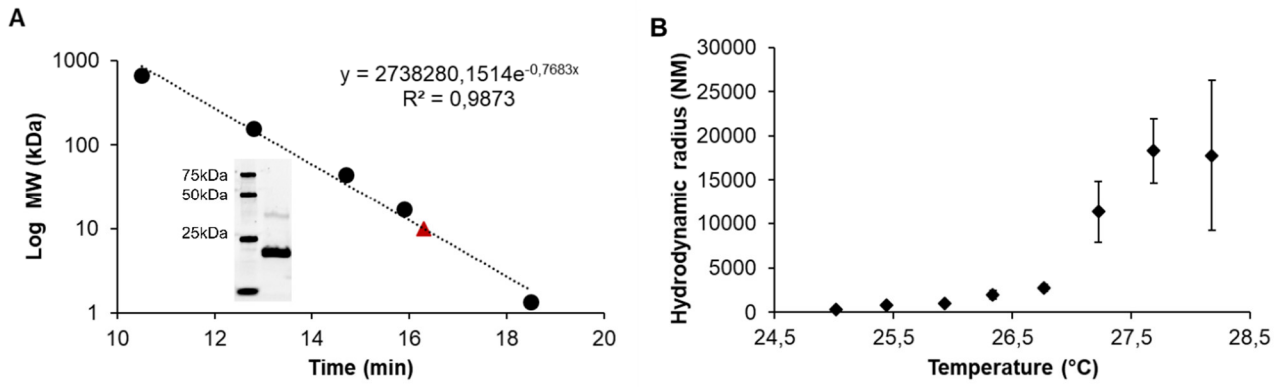


Figure 2. Determination of thermal stability and oligomerization state. (A) SDS-PAGE and analytical SEC of IMAC-purified protein (▲) and standards (●). MW: molecular mass. (B) Thermal stability is defined by a constant hydrodynamic radius (nm).

In total, six data sets were collected of co-crystallized FaGH2CBM. The structures were solved my molecular replacement ranging from 1.2 to 1.6 Å. FaGH2CBM is composed of three α-helices (Arg887-Leu889, Asn899-Asp901, Glu932-Asp937) and nine β-strands which are connected by 16 β-turns (Tab. S1). The β-strands are aligned in two antiparallel β-sheets of 6 and 3 strands (Fig. S2). Additionally, two beta hairpains (Met863-Arg865 and Glu939-Val947, Glu893-Arg896 and Ala950-Ala959) and four beta bulges were found (Tab.S2). The overall structure of FaGH2CBM is characterized by a unique β-topology, which is not found in any structural homologue and characterized by a low degree of conserved residues (Fig. 3A-B). FaGH2CBM is enriched in residues with charged and hydrophobic sidechains (Tab. S3, Fig. 3C-D) suggesting interaction with anionic polysaccharides like ulvan (Rahikainen et al. 2013). The binding of enzymes to glycans is driven by a favorable enthalpy, i.e. energy in form of heat is released. The binding is mediated by hydrogen bonds involving polar amino acid residues and water molecules as well as van der Waals' forces and hydrophobic stacking due to aromatic residues. Weaker interactions involve electrostatic forces and van der Waals' interactions (Holgersson et al. 2005). Since no electron density for bound substrate (Rha3S[GlcA]XylRha3S) was detected, we combined multiple sequence and structural alignment with web tools to predict putative residues involved in glycan binding. No putative binding site was identified using PLIP, PDBMotif and PDBe ProFun. PDBSum detected four significant clefts with a volume > 300 $Å^3$. The biggest cleft had a volume of 890 $Å^3$ containing 12 charged, three aromatic and two aliphatic residues (Fig. S3). Based on multiple sequence and structural alignment we identified two potential binding regions (Fig. S4). The first region consists of R865, K872 and R896 which are located 5.5, 6.9 and 9.9 Å apart. The second region contains residues T857, D908 and Q911 which are 3.8, 4 and 4.8 Å apart and are located above the active site of P17_GH2. The flexible loop (amino acids 840 - 847), which connects the catalytic domain with the putative binding domain, is likely involved in substrate binding. Hence, the second region is either exposed to putative substrates or re-directed above the active site in case FaGH2CBM is bound to a

substrate. Residues R865, K872, R896 and Q911 are highly conserved in GH2s but have not been described yet.

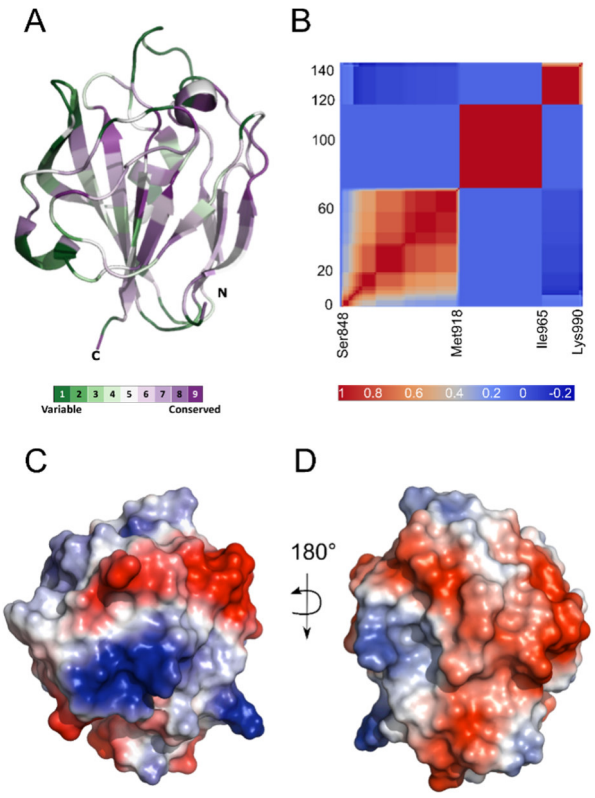


Figure 3. Structural insights into FaGH2CBM. Degree of conservation is shown as (A) cartoon representation of conserved amino acid residues and as (B) match correlation matrix: Heuristic PDB search compares query structure (x-axis) against those in the Protein Data Bank (y-axis). (C+D) Electrostatic potential shown is surface representation: cationic (blue) and anionic (red).

FaGH2CBM binds to macroalgal fucoidan

Besides co-crystallization experiments, we tested 30 marine and terrestrial polysaccharides using carbohydrate microarrays, ELISA and affinity gels to detect binding of FaGH2CBM.

No binding activity of FaGH2CBM was observed using carbohydrate microarray (data not shown). CBMs are known for their low binding affinity which usually varies in the millimolar range and allows a fast on/off binding rate of the ligand (Holgersson et al. 2005). Therefore, we performed affinity gel electrophoresis (AGE) using 0.05 – 0.15 % polysaccharide concentration. We observed no binding of the negative control BSA to any of the tested glycans. FaGH2CBM did not bind to glucomannan but retention in gels containing alginate, ulvan, laminarin, polygalacturonic acid and rhamnogalacturonan was observed (Fig. S5). CBMs are known for their promiscuity, i.e. binding of multiple different ligands, which is thought to enhance the efficiency for enzymes in environments where a variety of substrates are present (Boraston et al. 2004). To confirm the promiscuity of FaGH2CB, we performed ELISA using 0.5 µg ml^{-1} glycan concentration. Binding was only detected for anionic polysaccharides. The strongest

signal was observed for *Durvillaea potatorum* fucoidan, *Sargassum fusiforme* fucoidan, λ-carrageenan, *Macrocystis pyrifera* fucoidan, *Undaria pinnatifida* fucoidan, alginate as well as fucoidans from *Ecklonia maxima*, *Fucus serratus* and *Lessonia nigrescens*. No binding activity was found for ulvans or neutral and terrestrial polysaccharides. Non-productive protein-glycan-interaction due to electrostatic forces and hydrophobic interactions have been reported for CBMs binding to e.g. lignin (Rahikainen et al. 2013). Therefore, we repeated ELISA with 500 mM NaCl to reveal cationic interactions. The binding signal was diminished for alginate and λ-carrageenan and significantly reduced for fucoidans. The signal for *D. potatorum* fucoidan remained pronounced (Fig. S3).

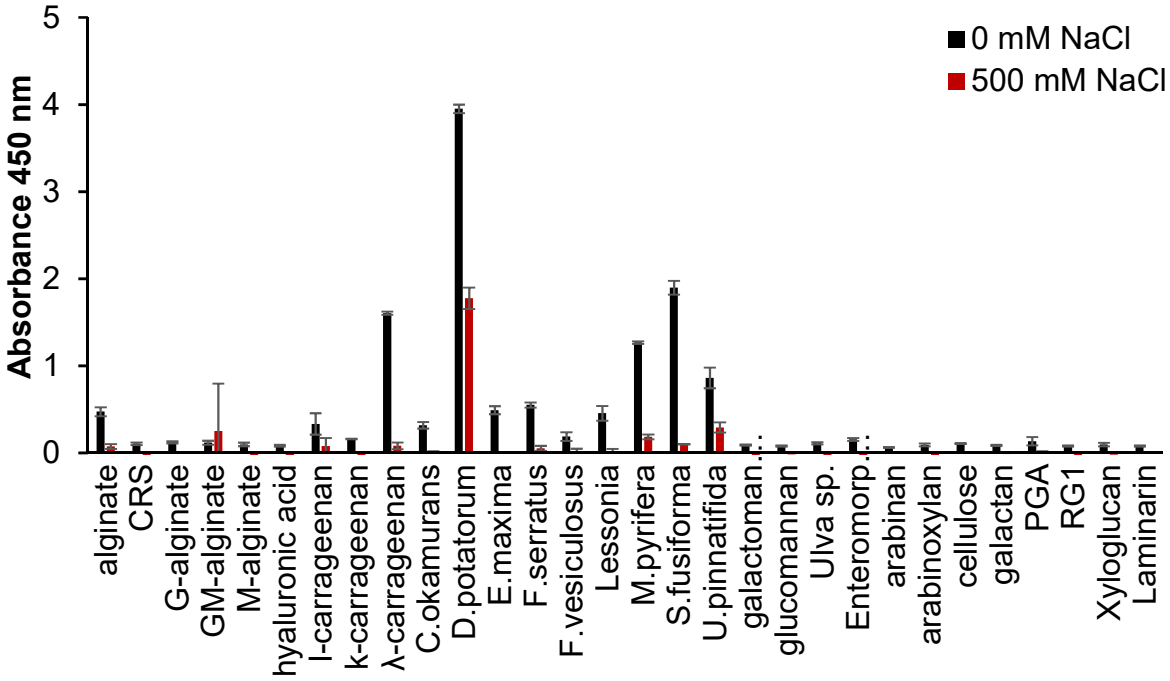


Figure 4. ELISA. Plates were coated with polysaccharides with and without NaCl and incubated with FaGH2CBM. Ani-His tag secondary antibody was used to detect recombinant plasmid bound to the polysaccharide. Binding activity was measured spectrophometrically at 450 nm. Error bars represent technical triplicates.

Conclusions

Similar to CAZymes, the activity of binding proteins depends on the core structure and confirmation of the substrate (Davies et al. 1997; Holgersson et al. 2005). Hence, it is not surprising that FaGH2CBM did not bind to ulvans in the carbohydrate microarray and ELISA but presumably to the ulvan tetrasaccharide which is cleaved by the exo-acting β-glucuronidase P17_GH2 (Reisky et al. 2019). Our binding assays indicates that FaGH2CBM is not monospecific to ulvan. Fucoidans from macroalgae are both structurally and compositionally versatile glycans. Depending on the algal species as well as

seasonal and spatial factors they contain ~ 20 – 50 % fucose and sulfate, respectively (Anastyuk et al. 2010; Fitton et al. 2015; Fletcher et al. 2017; Sichert et al. 2020). Certain species are enriched in other monosaccharides such as galactose(Shevchenko et al. 2015) or xylose (Bilan et al. 2014). The pronounced ELISA signal for fucoidan from *Durvillaea potatorum* suggests that it contains a similar building block of rhamnose, xylose and glucuronic acid which results from the degradation of ulvan (Fitton et al. 2015; Sichert et al. 2020). *Formosa agariphila* KMM309^T possesses the metabolic equipment to utilize pentoses like fucose, xylose and rhamnose as well as CAZymes which are organized in PULs targeting alginate, fucoidan and sulfated RG, mannan and laminarin (Berteau et al. 2002; Mann et al. 2013). To date, no fucoidan-binding protein has been reported and the only known ulvan-binding protein stems from the C-terminus of an endo-acting ulvan lyase from the marine flavobacterium *Nonlabens ulvanivorans* (Collen et al. 2011). This CBM bound to ulvan but not to alginate, heparin, dextran sulfate, iota carrageenan (Melcher et al. 2017). In contrast, FaGH2CBM and the binding domain of the ulvan lyase show no sequence similarities which together with their localization suggest different binding motifs (Ficko-Blean and Boraston 2012). Structural studies with enzyme-complexes are required to elucidate the binding activity with ulvan oligosaccharides and fucoidans.

Material and methods

Cloning and heterologous gene expression. The CBM from *Formosa agariphila* KM3901^T was PCR-amplified with a *Nhe*I restriction site at the forward primer and a *Xho*I restriction site at the reverse primer (Tab. S4) prior cloning into the expression plasmid pet28a(+) (69864, Merck Millipore) with a single 6x-His-Tag at the N-terminus (Hochuli et al. 1988). Recombinant purified plasmids were transformed into *Escherichia coli* BL21(DE3) according to the manufacturer's instructions (C2527H, New England Biolabs). Transformed cells were grown on lysogeny broth (LB) agar plates with 50 µg ml^{-1} kanamycin (Sigma Aldrich) at 37 °C overnight prior growing in 5 l Erlenmeyer flasks with 1 l of ZYP5052 auto-induction medium (Studier 2005) for four days at 20 °C and 150 rpm. Bacterial cultures were harvested at 4900 *g* and 4 °C for 15 minutes. For cell lysis, the pellet was resuspended in 20 ml buffer (25 % sucrose, 50 mM Tris pH 8), 30 mg of lysozyme (~ 7000 U mg^{-1}, Sigma Aldrich) was added and the mixture was incubated at room temperature (~ 25 °C) for 10 minutes while stirring. 40 ml of deoxycholate solution (20 mM Tris pH 8, 1 % w/v deoxycholate, 100 mM NaCl, 1 % w/v Triton X-100), $MgSO_4$ (5 mM final concentration) and 100 µl of 10 mg ml^{-1} DNase I (≥ 400 Kunitz mg^{-1}, Sigma Aldrich) were added. The mixture was incubated at room temperature (~ 25 °C) until it was no longer viscous and centrifuged at 30966 *g* and 4 °C for 45 minutes.

Protein purification. Immobilized metal affinity chromatography (IMAC) was carried out on a ÄKTA™ start chromatography system (29-0220-94, GE Healthcare Life Sciences) with a 5 ml cobalt column (HiTrap™ Talon® crude, 28953766, GE Healthcare Life Sciences). The supernatant from chemical lysis was applied to the column equilibrated in binding buffer (20 mM Tris pH 8, 500 mM

NaCl) at a flow rate of 5 ml min^{-1}. The column was subsequently washed with 40 ml binding buffer followed by elution with a linear gradient of 0 – 100 % elution buffer (20 mM Tris pH 8, 0.5 M NaCl, 0.5 M imidazole) over 25 ml. Size exclusion chromatography (SEC) was performed for concentrated IMAC fractions in NGC™ Chromatography System (Biorad) using a HiPrep 16/60 Sephacryl™ S-200 HR column (17116601, GE Healthcare Life Sciences) and 20 mM Tris pH 8 elution buffer.

IMAC and SEC fractions were verified for purity via sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) for soluble proteins (LAEMMLI 1970) according to the manufacturer's manual (Biorad). Purified protein fractions were concentrated in a stirred ultrafiltration unit (Amicon) using 50 psi nitrogen pressure and a 5 kDa membrane (Biomax®). The concentration was determined measuring the absorbance at 280 nm (BioSpectrophotometer basic, Eppendorf) using the calculated molar extinction coefficient (Tab. S4) computed by ExPASy ProtParam(Wilkins et al. 2005).

Thermal stability. Dynamic light scattering (DLS) was carried out to determine the thermal stability of the protein. Triplicates of 30 µl of SEC-purified protein were transferred into a microtiter plate (Aurora), centrifuged at 4500 *g* and 4 °C for 10 min to remove air bubbles, and analyzed in a DynaPro plate reader-II (Wyatt Technology). Thermal stability was monitored using a temperature gradient of 25 – 80 °C with an increase of 0.1 °C min^{-1} and 5 acquisitions per sample each measured for 5 s.

Analytical size exclusion chromatography. To confirm the molecular size of the protein and the state of oligomerization, analytical size exclusion chromatography (SEC) was performed using pooled and concentrated SEC fractions containing protein with low polydispersity. SEC was carried out in NGC™ Chromatography System (Bio Rad) using a high-resolution column ENrich™ SEC 650 10 x 300 (Bio Rad) and 1 ml min^{-1} flow rate with 20 mM Tris pH 8, 250 mM NaCl. The absorbance at 280 nm of 60 µl Gel Filtration Standard (1511901, Bio Rad) and protein sample was measured.

Protein crystallization and X-ray data collection. 10 sparse matrix and grid screens were tested to find conditions which allow the protein to crystallize. All screens were set up in 96-well sitting drop trials at 16 °C. Conditions which showed crystallized protein were repeated and optimized with slight modifications to the original recipe. Optimization was performed using hanging drop vapor diffusion method at 20 °C. 24-well hanging drop crystal trials were set up containing 500 µl mother liquor and three different dilutions (1:2, 2:1 and 1:1) of protein and crystallization solution with a total volume of 3 µl. Additionally, FaGH2CBM was co-crystallized (Hehemann et al. 2012) with 1 and 0.5 mg ml^{-1} Δ-Rha3S[2GlcA]-Xyl-Rha3S and Rha3S[2GlcA]-Xyl-Rha3S. Ulvan from *Ulva armoricana* was digested with ulvan lyase P30_PL28 and P33_GH105 from *Formosa agariphila* KMM3901^T and purified by SEC using a ~ 0.5 m column with P2 resin (Biorad) and 0.1 M ammonium carbonate buffer.

Protein crystals were soaked in mother liquor containing 25-30 % glycerol and flash-frozen to cryo-temperatures using liquid nitrogen. Data collection was done at beam line P11 (Meents et al. 2013) of PETRA III (Burkhardt et al. 2016) from the Deutsches Elektronen Synchrotron (DESY, Hamburg).

Structure determination. Data processing was done using XDS (Kabsch et al. 2010) and AIMLESS (Evans 2011) of the CCP4i Cloud (Krissinel et al. 2018). The structure was solved by molecular replacement using PHASER (McCoy et al. 2007) and the structural homologue PDB: 6HPD. The model was built using CCP4Build (Winn et al. 2011) and BUCCANER (Cowtan 2006). Refinement was done based on a Maximum Likelihood method using REFMAC V5.5 (Murshudov et al. 2011), manually finalized in COOT V0.8.7.1 (Emsley et al. 2010) and visualized in PyMOL Version 2.0 (Schrödinger, LLC). Statistics of data collection and structure determination are given in Tab. S5 in the supplements. ConSurf (Glaser et al. 2003) and DALI(Holm and Elofsson 2019) were used to visualize the degree of conserved residues in the 3D structure. The protein structure was analyzed with PDBSum (Laskowski et al. 2018). β-glucuronidases containing a CBM57 (PDB: 5UJ6) or a DUF4982 (PDB: 6MFV, 6ECA) originated from *B. unifroms*, *Pyrococcus horikoshii* and *Lactobacillus rhamnosus*, respectively, were used for structural alignment.

The atomic coordinates and structure factors have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

Binding studies. The binding properties of FaGH2CBM were determined using commercially glycans as well as self-made mannuronate-enriched alginate, guluronate-enriched alginate oligosaccharides and mixed-linked alginate oligosaccharides (Haug et al. 1967) as well as purified fucoidans (Sichert et al. 2020) (Tab. S6).

Carbohydrate microarray. Carbohydrate binding arrays were set up as previously described (Vidal-Melgosa et al. 2015; Mystkowska et al. 2018). In short, polysaccharides were dissolved in 55.2% glycerol, 44 % water, 0.8 % Triton X-100 and printed in triplicates onto nitrocellulose membrane with a pore size of 0.45 µm (Whatman, Sigma-Aldrich) using a microarray robot (Sprint, Arrayjet). Each of the polysaccharides was represented by 8 spots in the array: 2 spots with a concentration of 2.5 mg ml^{-1}, 2 x 0.5 mg ml^{-1}, 2 x 0.1 mg ml^{-1} and 2x 0.02 mg ml^{-1}. After blocking, the microarrays were incubated with 50 µg ml^{-1} FaGH2CBM, washed and incubated with anti-His tag for recombinant proteins produced in *E. coli*. Developed arrays were scanned at 2400 dots $inch^{-1}$ and analyzed using Array Pro Analyzer 6.3 (Media Cybernetics).

Enzyme linked immunosorbent assay (ELISA). ELISA was set up in triplicates as previously described (Cornuault et al. 2014). In short, microtiter plates (NUNC Maxisorp, Thermo Fisher Scientific) were coated with 0.5 µg ml^{-1} polysaccharides dissolved in MilliQ at 4 °C overnight, washed, blocked with 5% w/v low fat milk powder (Sigma Aldrich) in 1x PBS (MPBS) and washed.

Polysaccharide-coated plates were incubated with ~ 30 µg ml^{-1} FaGH2CBM in 5% MPBS for 1.5 hours, washed, incubated with anti-His-tag secondary antibody conjugated to horseradish peroxidase (A7058-1VL, Sigma Aldrich) at 5-11 µg ml^{-1} in in 5% MPBS and washed prior developing with tetramethylbenzidine (TMB) for 10 min. the reaction was stopped with an equal volume of 1 M HCl and analyzed at 405 nm using FLUOstar Omega microplate reader (BMG Labtech GmbH). The control set up contained only the enzyme buffer of FaGH2BM and was subtracted from the readout with enzyme to exclude unspecific binding signal from the anti-His-tag antibody.

Affinity gel electrophoresis (AGE). AGE was carried out as previously described (Mystkowska et al. 2018). In short, 12 % native gels containing 0.1 to 1 % glycans were electrophoresed for 2 h at constant 100 V on ice. Bovine serum albumin (BSA, Sigma Aldrich) was run as a negative control. The gels were stained using Coomassie blue and visualized in Gel Doc™ EZ Gel Imager (Biorad).

Sequence analyses. Relative sequences were retrieved from NCBI BLASTp, CAZy (Lombard et al. 2014b), TARA Ocean (Karsenti et al. 2011), Malaspina and MAR (Klemetsen et al. 2018). Sequences were aligned using MUSCLE (Edgar 2004). MSA was processed in ESPript 3.0(Robert and Gouet 2014) in order to identify conserved residues.

Acknowledgments and funding

We acknowledge DESY (Hamburg, Germany) for the provision of experimental facilities. Parts of this research were carried out at PETRAIII and we would like to thank Dr. Anja Burkhardt for assistance in using the beamline P11. We thank Silvia Vidal-Melagosa for microarray analysis, Craig S. Robb for input of experimental design and Lukas Reisky for providing purified ulvan oligos. J.H.H. acknowledges funding from the Max-Planck-Gesellschaft, the Emmy Noether Program of the DFG (grant number HE 7217/1-1) and the framework of the research unit FOR 2406 ‘Proteogenomics of Marine Polysaccharide Utilization’ (POMPU) of the DFG.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

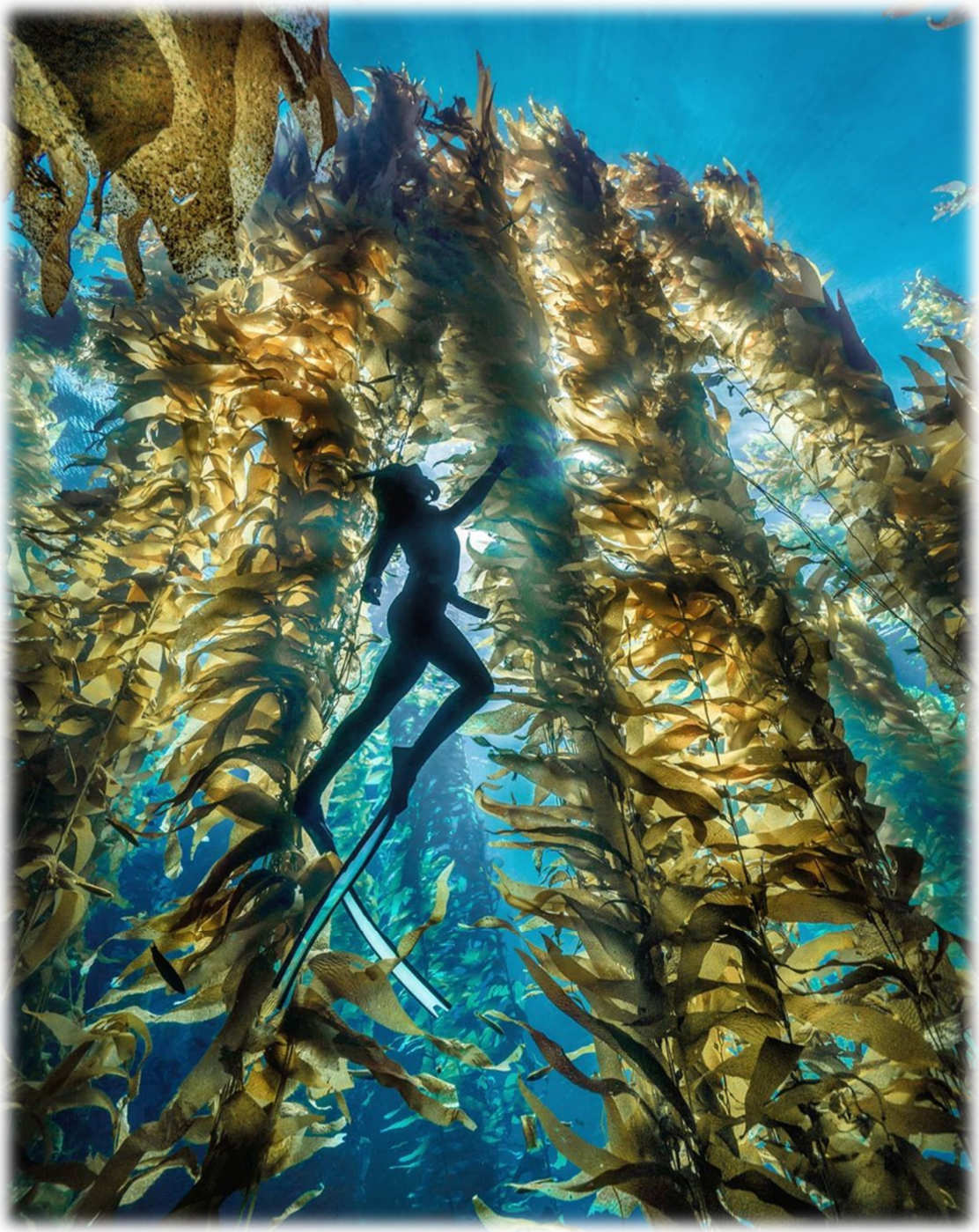
JHH and NG designed the study. NG collected X-ray diffraction data, performed binding studies and structural analyses. NG and CV performed metagenomic analyses. NG and JHH wrote the manuscript. All authors contributed to editing of the manuscript.

Supplementary material

In appendix.

Chapter 2 – Mechanistic insights into fucoidan degradation

Study motivation in brief:



Let us dive into another universe, shall we?

© Photo credit: John Kowitz.

Chapter 2.1. Crystal structure of a putative marine bacterial sulfatase of family S1_15

Nadine Gerlach[1,2], Andreas Sichert[1,2], Jan-Hendrik Hehemann[1,2,*]

Affiliations

[1] MARUM – Center for Marine Environmental Sciences, University of Bremen, Germany

[2] Max Planck Institute for Marine Microbiology, Bremen, Germany

*Correspondence: jhhehemann@marum.de, +4942121865775

To be submitted in Protein Science.

Abstract

The remineralization of algal polysaccharides by heterotrophic bacteria is a crucial step in the marine carbon cycle. Many marine glycans are characterized by a structural and compositional variety. Bacteria capable to degrade such complex polysaccharides require an individual multistep degradation pathway of several enzymes acting in a distinct order. Key degraders possess up to several hundreds of sulfatases indicating the importance of sulfated glycans as carbon source. Yet, the role of marine sulfatases in the utilization pathways of algal glycans remains poorly understood. Here, we identified a putative fucoidan sulfatase of subfamily S1_15 from the marine Verrucomicrobium '*Lentimonas'* sp. CC4. Previous transcript- and proteomic analyses revealed that 21_S1_15 is involved in the utilization of sulfated polysaccharides found in the cell walls of brown macroalgae. We solved the structure of 21_S1_15 to a resolution of 1.7 Å by X-ray crystallography. 21_S1_15 is a C-FGE type sulfatase, which assembles as monomer of 53 kDa. The highly conserved active site has a pocket-like topology surrounded by cationic amino acids, supporting the proteomic result that this sulfatase is involved in the degradation of sulfated fucans.

Key words: sulfatase family S1, fucoidan, marine bacteria, algal polysaccharides

Introduction

Sulfatases cleave sulfate ester groups from various biomolecules including carbohydrates, proteins and steroids, and thereby serve an important role in numerous physiological processes as well as biogeochemical cycles (Hanson et al. 2004). Similar to carbohydrate-active enzymes (CAZymes), sulfatases are specific for the targeted carbohydrate backbone (Hettle et al. 2018) and grouped into homologous enzyme families based on sequence similarities (Lombard et al. 2014a; Barbeyron et al. 2016a). More than 41000 entries for sulfatases from viruses and all kingdom of lives are grouped into four families. About two thirds are arylsulfatases belonging to family S1 (formylglycine-dependent sulfohydrolases, FGly-sulfatases, EC 3.1.6. and 3.10.1), which consist of 73 substrate-specific subfamilies as well as several non-classified members (Barbeyron et al. 2016a). S1 sulfatases have been reported to be highly specific, acting on various substrates including chondroitin (Wang et al. 2019b), glycosaminoglycans (Cartmell et al. 2017; Ndeh et al. 2020), ulvan (Reisky et al. 2019) and carrageenan (Hettle et al. 2019). Structural studies have revealed that S1 is highly conserved consisting of a small C-terminal and a large N-terminal domain harboring the active site. The active site is composed of ten conserved residues: five basic and four acidic/polar amino acids as well as a formylglycine (FGly), which derives from post-translational oxidation of a Cys or Ser. Both types of FGly residues (C- or S-FGE) are crucial for the sulfate ester hydrolysis

where first a sulfate diester is formed in the nucleophilic attack of the oxygen atom of the sulfate group with the aldehyde group of the FGly prior the release of the alcohol conjugated by an activated water (nucleophile) (Schmidt et al. 1995; Lukatela et al. 1998; Dierks et al. 1998). Furthermore, family S1 are metal-dependent enzymes. Divalent cations such as Mg^{2+} or Ca^{2+} are coordinated by negatively charged amino acids, whereas positively charged amino acids are responsible for substrate binding (Hanson et al. 2004). Despite the fact that family S1 is mechanistically and structurally very conserved, the subfamilies have been reported to be very substrate specific (Hanson et al. 2004; Barbeyron et al. 2016a), which raises the questions what feature of the enzyme structure is determining substrate specificity.

In this study we identified a putative fucoidan sulfatase from subfamily S1_15 (EC 3.1.6.4), which is upregulated in the presence of macroalgal fucoidan from *Fucus vesiculosus*. Fucoidans are a structurally highly diverse class of polysaccharides consisting of several monosaccharides which vary in the degree of sulfation (Ale and Meyer 2013; Skriptsova 2015; Deniaud-Bouët et al. 2017). Bacteria capable to utilize these complex glycans belong to the phyla Bacteroidetes and Gammaproteobacteria and the superphylum Planctomycetes–Verrucomicrobia–Chlamydiae (PVC) (Sakai et al. 2003; Descamps et al. 2006; Barbeyron et al. 2008; Silchenko et al. 2013; Chen et al. 2016; van Vliet et al. 2019; Sichert et al. 2020). These species upregulate the expression of substrate specific enzymes of various families including endo-acting fucoidanases of family GH107 and exo-acting fucosidases of family GH29. However, the most striking genetic feature of fucoidan-degraders is their possession of several hundreds of sulfatases (Thrash et al. 2010; Wegner et al. 2013; van Vliet et al. 2019; Sichert et al. 2020), of which none characterized member has been shown to act on fucoidan yet. The only known sulfatases to act on fucoidan originate from the marine mollusk *Pecten maximus* and the marine bacterium *Wenyingzhuangia fucanilytica* CZ1127^T, which belong to the subfamilies S1_17 and S1_25. These fucoidan sulfatases have been reported to be highly substrate specific exo-acting enzymes, exclusively cleaving 2,3-di-O-sulfate of sulfated fucoidan-oligosaccharides and *Ascophyllum nodosum* fucoidan. Other brown algae species such as *Fucus vesiculosus, Fucus evanescens* or *Sargassum horneri* do not get degraded by these enzymes (Daniel et al. 2001; Berteau et al. 2002; Silchenko et al. 2018). Such findings show that fucoidan sulfatases require enzymatic partners to be active. To date, there are no 3D structures of fucoidan sulfatases available which could contribute to a better understanding of their mechanistically motif.

To shed light on the role of subfamily S1_15 in fucoidan degradation, we combined biochemical analyses with X-ray crystallography. We solved the structure of 21_S1_15 to 1.7 Å and identified striking structural differences with other members of this subfamily. The enzyme was inactive on chromogenic substrates as well as various sulfated marine oligo- and polysaccharides from various marine algae species. The absence of activity suggest that the utilization of the structurally diverse

and complex polysaccharide fucoidan may require an individual multistep degradation pathway of several enzyme families in a distinct order as seen for RGII (Ndeh et al. 2017), ulvan (Reisky et al. 2019) and carrageenan (Ficko-Blean et al. 2017).

Results and discussion

Identification of a putative fucoidan sulfatase

Recently, the capacity of the marine bacterium '*Lentimonas'* sp. CC4 to grown on and utilize sulfated marine polysaccharides like fucoidan and iota-carrageenan has been reported (Sichert et al. 2020). Transcriptomic and proteomic analyses revealed that genes involved in glycan utilization are regulated in substrate-specific operons similar to polysaccharide-utilization loci (PULs) in Bacteroidetes. We re-inspected the data and found a fucosidase of family GH29 (22_GH29) and a sulfatase of subfamily S1_15 (21_S1_15) reached up to 0.5% and 0.25% of the total proteome of '*Lentimonas'* grown on fucoidan from *F. vesiculosus*. Both genes are potentially regulated by one promotor, since we identified an AT-rich -10 and -35 region upstream of the 22_GH29 gene but not upstream of 21_S1_15. Both genes are only separated by a three bp intergenic distance and read mapping of the RNA-seq data showed that both genes and the intergenic region have similar expression values. This indicates that both genes are transcribed as polycistronic mRNA (Fig. S1) and this transcriptional wiring suggests that both enzymes act together on fucoidan and therefore, we chose 21S1_15 for further characterization.

Domain organization and clustering of subfamily S1_15

Subfamily S1_15 is mainly found in marine bacteria of the phylum Firmicutes as well as the superphylum Planctomycetes-Verrucomicrobia-Chlamydia (PVC) as well as in soil or gut microbes belonging to the classes Bacteroidia, Flavobacteriia and Cytophagia of the phylum Bacteroidetes (Fig. S2). Marine members of the PCV have been shown to be highly specialized sulfated glycan degraders (Kim et al. 2016; Reintjes et al. 2017; Reintjes et al. 2018; Wiegand et al. 2020), suggesting that this subfamily might play an essential role in the degradation pathways of macroalgal fucoidans. BLAST analysis revealed that 21_S1_15 shares above 77 % identity (99 % coverage) and 64 % identity (98 % coverage) with sulfatases S1_15 from *Kiritimatiellaeota* (WP_136078751) and *R. sallentina* (EMI56448.1), respectively.

Subfamily clusters into 19 clades, displaying structural differences. The comparison of the sequence alignment among all 2000 available sequences indicates gene transfer among clades, involved with structural differences in the catalytic domain and, presumably impacting substrate specificity within this subfamily. Overall, marine S1_15 are characterized by a greater structural diversity than host associated/terrestrial Bacteroidetes (Fig. S2).

Sulfatase 21_S1_15 has a molecular mass of 53.96 kDa (Fig. 1A) and a predicted isoelectric point of pH 6.24. The functional domain consists of a N-terminal signal peptide of 0-19 amino acids, the catalytic domain and a C-terminal DUF4994 domain. Members missing the DUF4994 domain or having additional domains such as a carbohydrate-binding domain are rarely found (Fig. S2).

Activity screening

So far, S1_15 have been reported to be exo-acting enzymes, cleaving 6O sulfate of N-acetylgalactosamine from sulfated glycosaminglycans6 and galactose from keratan sulfate (Tohru et al. 1982). To shed light on the role of subfamily S1_15 in fucoidan utilization, we performed heterologous gene expression. Since family S1 requires the posttranslational modification to an FGly residue (Hettle et al. 2018), we have co-expressed 21_S1_15 with and without the maturing enzyme FGE to ensure the post-translational maturation of the cysteine. Prior activity assays, the thermal stability of 21_S1_15 was determined to be below ~ 37 °C by DLS (Fig. 1B). Hence, digests were set up at 30 °C. In total, we have tested five sulfated oligosaccharides and 16 sulfated polysaccharides originated from marine organisms including micro- and macroalgae.

Firstly, 21_S1_15 was tested for activity using the chromogenic substrate pNP sulfate which is a commonly used to measure sulfatase activity (Hanson et al. 2004). Neither 21_S1_15 nor the enzyme with co-expressed FGE were active on pNP sulfate. It has been shown that many sulfatases are not active on pNP sulfate but on their natural substrate as they require binding to the carbohydrate backbone (Hettle et al. 2018). Therefore, we tested variety of sulfated substrates from marine macroalgae including ulvan and fucoidan, sulfate rich exopolymers of microalgae and carrageenan oligosaccharides. As expected, 21_S1_15 was not active on chondroitin sulfate, the two ulvans from *Ulva* sp. and *Enteromorpha* sp., nor the five carrageenan oligosaccharides (Fig. S3). However, we could also not observe sulfate release from any of the nine structurally different fucoidans nor the exopolymers of four microalgae species. Since it was previously reported that sulfatases increase the activity of exo-fucosidases (Berteau et al. 2002), we tested if the adjacent sulfatase 21_S1_15 increases the yield of fucose release by 22_GH29. Therefore, both enzymes were incubated with two structurally different fucoidans from *F. vesiculosus* and *C. okamurans*. Neither sulfate was detected nor an increase of fucose could be observed by ion exchange chromatography.

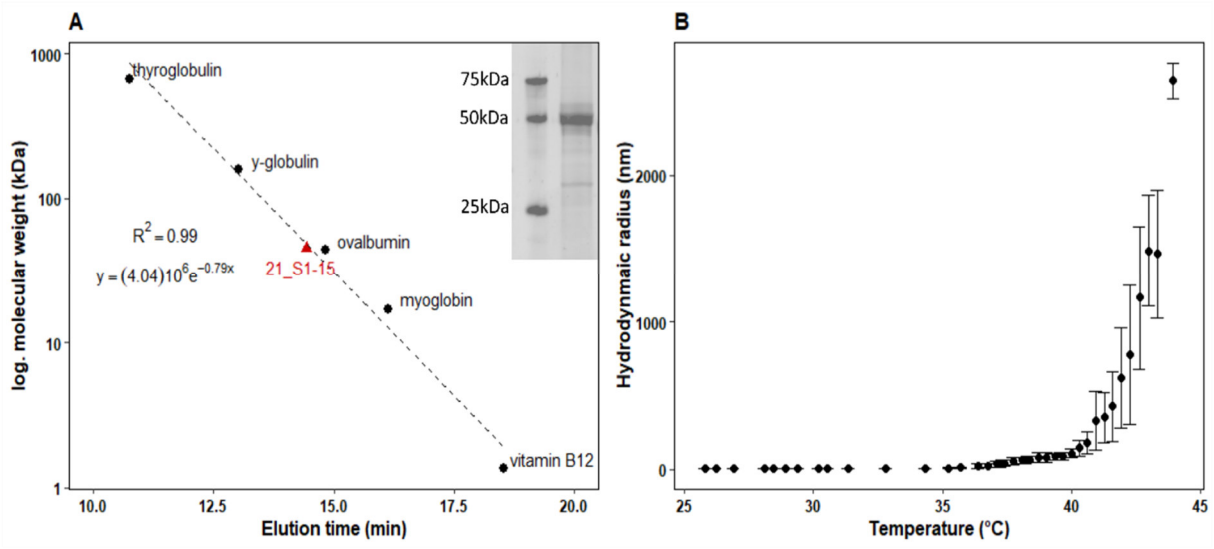


Figure 1. Molecular size and thermal stability of 21_S1_15. (A) Determination of molecular mass by analytical SEC and SDS-PAGE and (B) thermal protein stability indicated by a constant hydrodynamic radius (nm) until ~ 37 °C.

Overall structure of 21_S1_15

To gain a better understanding for the mechanically motif of 21_S1_15 we performed X-ray crystallography to gain detailed structural insights into this subfamily.

21_S1_15 crystallized at 4, 9 and 12 mg ml^{-1} in various conditions covering a wide range of concentration (4 – 20 %) and molecular weight of PEGs (2000 MME, 3350, 4000, 6000, 8000, 10000, 20000). The conditions contained different buffer system of 0.1 M pH 5.6 – 9, as well as 0.1 – 0.2 M cations (MgCl2, CaCl2, NaCl) or sodium iodide, ammonium sulfate, lithium sulfate, calcium acetate or sodium acetate. In total, 11 data sets from 1.3 – 2.5 Å of different conditions were collected of needle-like and cubic crystals. The holo-structure of 21_S1_15 was solved to 1.7 Å resolution of a $P2_1$ crystal (Tab. S1). The structure was built from 17 – 494 amino acid with one molecule in the asymmetric unit at 44 % solvent and a Matthew Coefficient of 2.32 at 481995.91 $Å^3$. 21_S1_15 assembles as a monomer (confirmed by analytical SEC, Fig. 1A) with overall dimensions of about 60x60x47 Å (Fig. 2A).

In total, 21_S1_15 consists of seven α-helices and 18 β-strands which are connected by six gamma turns and 46 β-turns. Additionally, one psi loop (F337-S340 and T349-S351), four β- bulges and seven helix-helix interactions were found. 21_S1_15 is characterized by three tunnels of 16.6, 20.2 and 50.1 Å length and four significant clefts, which have a volume > 900 $Å^3$ and enriched in aliphatic and charged residues.

Structural comparison of subfamily S1_15

21_S1_15 displays the same overall heart-like fold as the only structural homologue PDB:6S20 within this subfamily. The N-terminal region is composed of one β-sheet of 10 parallel strands which are surround by two times three α-helices. It harbors the active site containing additional electron density for a calcium ion, which interacts with C73, D31, D32, D289 and N299. The C-terminus consists of a β-sheet of four parallel strands and one α-helix (Fig. 2A). The degree of conserved amino acids is highest in the N-terminal domain (Fig. 2B). A correlation matrix of matched PDB structures (Holm and Elofsson 2019) along 21_S1_15 showed that the DUF4994 domain at the C-terminus is the least conserved part of this subfamily (Fig. S4). This region has been proposed to be the candidate for a substrate-discerning region in family S1 since it is the most structural diverse region (Hanson et al. 2004). However, 21_S1_15 and its structural homologue PDB:6S20 do not only share a similar N- but also C-terminal domain structure at 33 % sequence identity (93 % coverage). Structural differences can be found superimposing the structures of both enzymes. PDB:6S20 has two extended loop regions surrounding the active, of which the first is not present in 21_S1_15 and the second region is displayed as a short α-helix (302-315 aa). These structural differences lead a small pocket architecture typically found in exo-acting enzymes (Hettle et al. 2018) for PDB:6S20, but to a one-side open active site in 21_S1_15 (Fig. 2C, D). This structural difference could leave room for the backbone of the natural substrate of which only a terminal sulfate group can face towards the sulfate binding subsite (S-subsite) of the active site. Conceivably, these regions are involved in substrate specificity within this subfamily rather than the structurally diverse C-terminal domain comparing all subfamilies within family S1 (Hanson et al. 2004) by influencing the recognition of the leaving group component, i.e. the glycan backbone (Hettle et al. 2018).

Insights into the active site of subfamily S1_15

Like glycoside hydrolases, the active site of sulfatases is described in sugar-binding subsites for a better understanding of the mechanically motif of endo- and exo-acting enzymes (Davies et al. 1997; Hettle et al. 2018). The numbering of the sugar units within the ligand starts from the point of cleavage and increases towards the reducing (+ -subsites) and non-reducing ends (– -subsites), respectively (Ficko-Blean et al. 2017). In sulfatases, this concept has been expanded with the 0-subsite, which accommodates the monosaccharide and its sulfate group bound to the S-subsite[2]. Here, the point of cleavage is between the 0- and S- subsite.

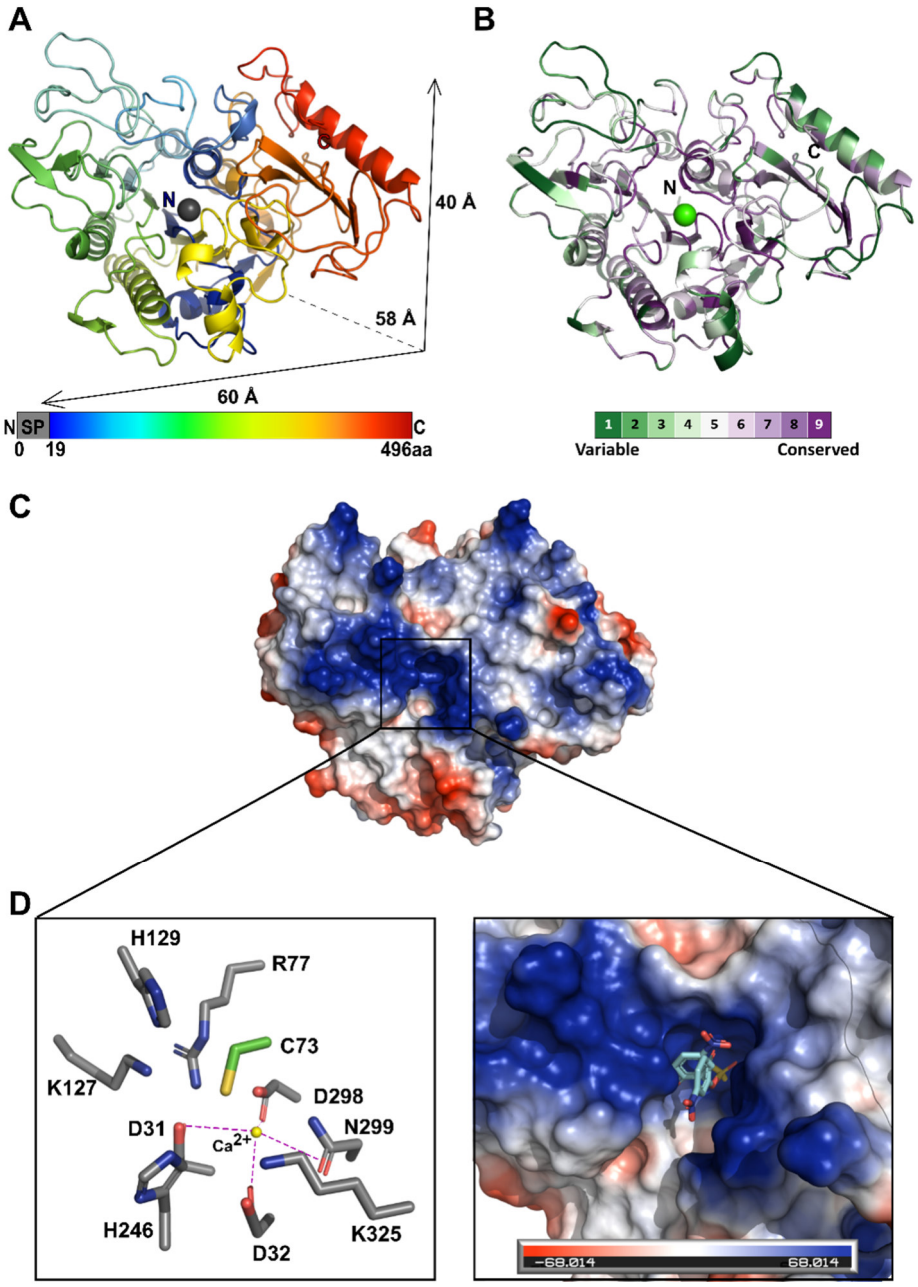


Figure 2. Overall structure of 21_S1_15. (A) Cartoon view and overall dimensions (Å): heart-like overall fold consisting of 12 helices and 12 strands. (B) surface view colour coded according to the degree of amino acid residue conservation predicted by ConSurf. (C) Surface view colour coded according to electrostatic potential calculate on PyMOL ranging from positive (blue) to negative (red) and (D) zoom into active site of 21_S1_15: highly conserved residues, catalytic nucleophile Cys colored in green, calcium ion which binds and activates the sulfate group of the substrate. Superimposition of *p*-nitrophenyl sulfate (colored in cyan and yellow) from PDB: 1E2S (Von Bülow et al. 2001) into 21_S1_15.

In family S1, the active site or S-subsite is composed of 10 highly conserved residues (Fig. 2D), which are located in six regions (Barbeyron et al. 2016a). In 21_S1_1_15 (Fig. S5), the first region, N-I/X-L/X-L/X-I/X-X-A/X-D-D/X, starts with a highly conserved N24 and ends with two D31/32 located in the active which are, together with D289 and N299 of the fifth region, N/X-T-L/X-I/X-I/X-F/X-T/X-S/X-D-NX-G, responsible for metal binding. Sulfatases of family S1 acquire divalent cations such as Mg^{2+} or Ca^{2+}, which are coordinated by negatively charged amino acids, whereas positively charged amino acids are responsible for substrate binding (Hanson et al. 2004). The second region, C/S-X-P-S/X-R-X-X-X-L/X-T/X-G/X-R/ X, starts with the C73 residue which is post-translationally modified to the catalytically active residue FGly (Schmidt et al. 1995; Lukatela et al. 1998; Dierks et al. 1998). With this cysteine, 21_S1_15 belongs to the C-FGE type, whereas its structural homologue PDB:6S20 belongs to the S-FGE type sulfatase. Moreover, this region contains R77, which is together with the other positively charged residues (K127/325 and H129/246 of the third, G-Y/V-X- S/T-X-X-X-G-K-X-X-H, fourth, P-F/X-F/X-L/X-Y/X-X-X-X-X-X-P/X-H-X-P/X, and six, K-X-X-X) positioning and binding the tetrahedral sulfate for the nucleophilic attack by FGly (Von Bülow et al. 2001).

The active site of exo-sulfatases have a small pocket architecture compared to open clefts for endo-sulfatases (Hettle et al. 2018; Ndeh et al. 2020). The 0- and S-subsite are characterized by positively charged residues which bind to sulfated substrates. 21_S1_15 is not only characterized by a cationic active site but also with a highly cationic surface surrounding the active, indicating anionic glycans such as fucoidans as natural substrate (Reisky et al. 2019). Compared to endo- and exo-acting sulfatases, 21_S1_15 is neither possessing a typically open cleft for endo-acting sulfatases nor a small active site architecture for reducing end exo-acting sulfatases and so the mode of action of this enzyme remains unclear (Fig. 3). The structural homologue of 21_S_15 is characterized by a typical small pocket active site architecture and is exclusively active on GalNAc6S (Ndeh et al. 2020). GalNac6S is not a known compound of fucoidans (Ale and Meyer 2013; Skriptsova 2015; Deniaud-Bouët et al. 2017), but part of other sulfated marine polysaccharides like chondroitin sulfate found in shark cartilage (Sim et al. 2007). The rather wide exposure of the active site in 21_S1_15 might be an indication for the motif of a non-reducing end exo-acting sulfatase. Moreover, the active site of 21_S1_15 is surrounded by several cationic amino acids (His201, Arg202), which are facing towards the putative plus subsite (Fig 3). These amino acids cannot be found in the structural homologue PDB:6S20. It is likely, that these residues coordinate the negatively charged, sulfated backbone of a substrate in the active site.

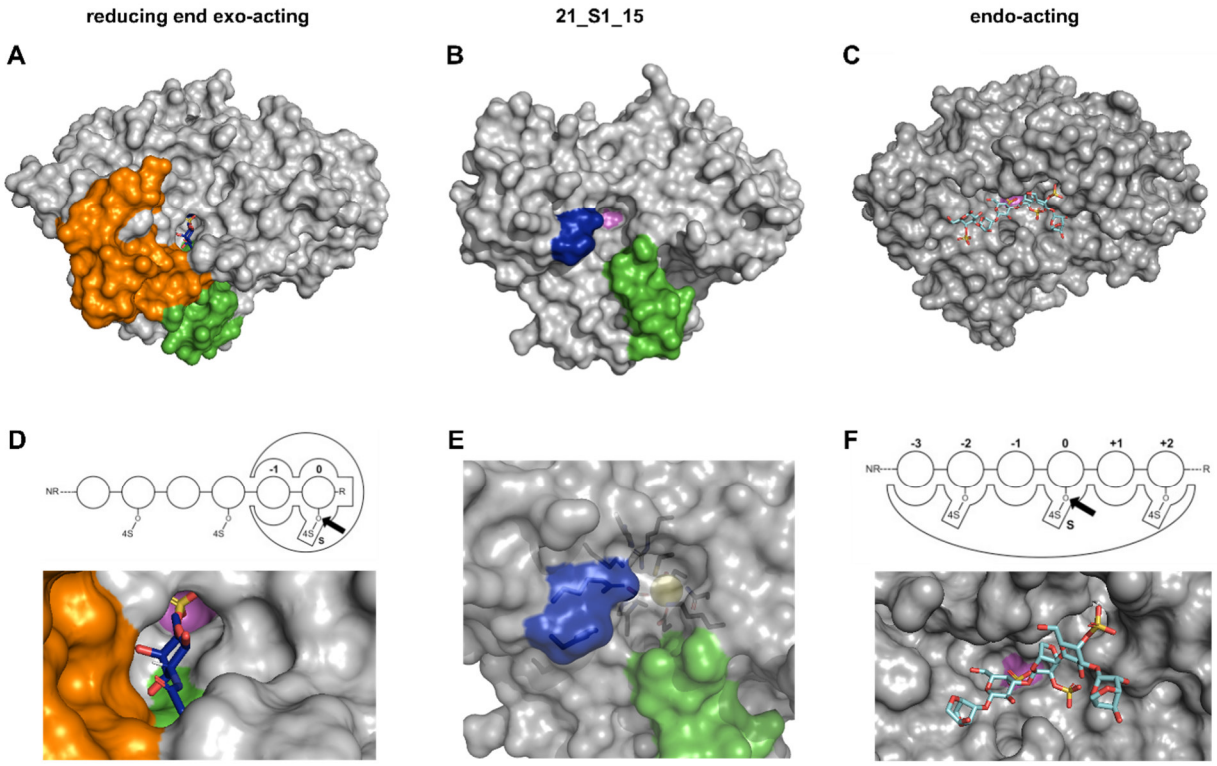


Figure 3. Comparison of the active site architecture. Surface view and zoom into the active site of (A, D) reducing end exo-acting sulfatase from subfamily S1_15 (PDB: 6S20), (B, E) and 21_S1_15 (C, F) endo-acting sulfatase of subfamily S1_19 (PDB: 6B1V). Structural differences of both members from subfamily S1_15 are highlighted in green and orange. Additional cationic residues of 21_S1_15 are shown in blue. Calcium ion is shown as yellow sphere, whereas the position of the catalytic FGly is shown in pink. The reducing end exo-acting and the endo-acting sulfatases are shown with their substrates in the active site

Conclusion

We have identified a sulfatase of subfamily S1_15 from the fucoidan-degrading marine Verrucomicrobium '*Lentimonas'* sp. CC4. 21_S1_15 is located, together with other CAZymes, in an operon structure upregulated in the presence of macroalgal fucoidan from *F. vesiculosus*. We combined X-ray crystallography and intense activity screening to gain insights into the S1_15 subfamily and its role in fucoidan utilization. The absence of activity towards any of the tested substrates can be explained by several factors: (i) The amount of cleaved sulfate is below the detection limit or (ii) the co-expression with the maturing enzyme did not activate 21_S1_15. More likely, 21_S1_15 might possess a specific substrate specificity like other known fucoidan sulfatases (Daniel et al. 2001; Berteau et al. 2002; Silchenko et al. 2018). In this case, the substrate specificity is not only influenced by (iii) the fucoidan source (i.e. algae species), but also by (iv) the 3D structure of the substrate. The active site of 21_S1_15 has a pocket-like topology known for exo-

acting enzymes. Therefore, it is likely that 21_S1_15 requires enzymatic partners to cleave fucoidan into smaller fragments to be active. Thereby, the accommodation of the building blocks of the substrate backbone determines the recognition of the substrate and ultimately the mode of action (i.e. exo- vs. endo-acting sulfatases) (Hettle et al. 2018). This hypothesis would be in line with hydrolytic pathways found in bacteria utilizing polysaccharides like ulvan (Reisky et al. 2019) and carrageenan (Hettle et al. 2019). To answer the remaining uncertainty regarding the substrate of 21_S1_15, natural substrates are. Generating defined oligos for complex glycans like fucoidans remains a challenging task, but in combination with structural biology provide detailed insights in the mechanistic motif of fucoidan sulfatases.

Material and methods

Cloning

Recombinant plasmid of 21_S1_15 was produced and purified as described previously (Reisky et al. 2019). In short, 21_S1_15 was amplified by PCR (M0492, New England Biolabs) and constructed with a 6x-His Tag (Hochuli et al. 1988) at the N-terminus and ~30 bp overhang (Tab. S2) to be ligated into the restriction sites *NheI* and *XhoI* of expression vector pET28a(+) (69864, Merck Millipore).

Recombinant plasmid of 21_S1_15 was transformed into *E. coli* BL21(DE3) (C2987I+C2527H, New England Biolabs) for protein overexpression. Bacterial cultures were grown in 5 L Erlenmeyer flasks with 1 L of ZYP5052 auto-induction medium (Studier 2005) and 100 µg mL^{-1} kanamycin (Sigma Aldrich) for four days at 20 °C and 150 rpm. Additionally, 21_S1_15 was co-expressed with pBAD vector containing FGE from *M. tuberculosis* (Hettle et al. 2018) to ensure maturation of the catalytic cysteine. Therefore, 21_S1_15 was grown in 5 L Erlenmeyer flasks with 1 L of LB, 2 µM $CuSO_4$, 50 µg mL^{-1} kanamycin and 50 µg mL^{-1} chloramphenicol (Sigma Aldrich) at 37 °C until OD_{600nm} ~ 0.5. The expression of FGE was induced by 0.1 % L-arabinose (Sigma Aldrich) for 2-5 h at 20 °C, followed by the induction of the sulfatase using 1 mM isopropyl-β-D-thiogalactopyranosid (IPTG, Sigma Aldrich) at 16 °C overnight.

Heterologous gene expression and purification

Bacterial cells were lysed chemically and immobilized metal affinity chromatography (IMAC) was carried out in ÄKTA™ start chromatography system (29-0220-94, GE Healthcare Life Sciences) with a 5 mL cobalt column (HiTrap™ Talon® crude, 28953766, GE Healthcare Life Sciences). Size exclusion chromatography (SEC) was performed for concentrated IMAC fractions in NGC™ Chromatography System (Biorad) using a HiPrep 16/60 Sephacryl™ S-200 HR column (17116601, GE Healthcare Life Sciences) and 20 mM Tris pH 8 250 mM elution buffer.

IMAC and SEC fractions were verified for purity via sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) for soluble proteins (LAEMMLI 1970). Protein fractions purified via IMAC and SEC were concentrated in a stirred ultrafiltration unit (Amicon) using 70 psi nitrogen pressure and a 10 kDa membrane (Biomax®). The protein concentration was determined measuring the absorbance at 280 nm (BioSpectrophotometer basic, Eppendorf) using the molar extinction coefficient (Tab. S2) computed by ExPASy ProtParam (Wilkins et al. 2005).

Analytical size exclusion chromatography

To determine the molecular size and the state of oligomerization of the protein, analytical size exclusion chromatography (SEC) was performed using pooled and concentrated SEC fractions. SEC was carried out in NGC™ Chromatography System (Bio Rad) using a high-resolution column ENrich™ SEC 650 10 x 300 (Bio Rad) and 1 mL min^{-1} flow rate with 20 mM Tris pH 8, 250 mM NaCl. The absorbance at 280 nm of 60 µL Gel Filtration Standard (1511901, Bio Rad) and protein sample was measured.

Thermal stability

Prior enzymatic assays, dynamic light scattering (DLS) was performed to determine the thermal stability of the protein. Triplicates of 30 µL of SEC-purified protein were transferred into a microtiter plate (Aurora), centrifuged at 4500 *g* and 4 °C for 10 min and analyzed in a DynaPro plate reader-II (Wyatt Technology). Thermal stability was monitored using a temperature gradient of 25 – 80 °C with an increase of 0.1 °C min^{-1} and 5 acquisitions per sample each measured for 5 seconds.

Enzyme activity assays

The activity of 21_S1_15 was tested photometrically (FLUOstar Omega microplate reader, BMG Labtech GmbH) at 405 nm using *4*-nitrophenyl sulfate potassium salt (pNP-S, Sigma Aldrich). The activity towards sulfated marine oligo- and polysaccharides was observed by and modified thin layer chromatography (TLC) protocol (Vreeburg et al. 2014) and anion chromatography monitoring the release of sulfate and fucose (Engel and Händel 2011). In short, HPAEC-PAD was used to detect the release of fucose using a Dionex ICS5000+ system (Dionex) equipped with a CarboPac™ PA10 (2 mm x 250 mm, Dinoex) analytical column and a CarboPac™ PA10 (2 mm x 50 mm, Dinoex) guard column. Sulfate release was observed using a Metrosep Trap 1 100/4 column operated on an IC930 Flex (Metrohm) and quantified with 1 – 100 mM MgSO4 as internal standard. TLC was performed using 2:1:1 butanol:MilliQ:acetic acid as mobile phase and silica gel TLC plates size 4x8cm (Merck) as stationary phase. Digest products were labelled with 2-aminoacridone (AMAC) for 2 h at 37 °C.

Enzyme assays were set up in triplicates at 30 °C and 300 rpm overnight with 0.1 mg ml^{-1} 21_S1_15, 0.5 mg ml^{-1} substrate and 200 mM NaCl and 20 mM Na-citrate pH 5.8. The following substrates were used for activity screening: ulvan from *Ulva* sp. (Elicityl) and *Enteromorpha* sp. (Elicityl), ion exchange chromatography-purified fucoidans (Sichert et al. 2020) from the macroalgae species *Fucus serratus* (Carbosynth), *Fucus vesiculosus* (Carbosynth), *Undaria pinnatifida* (Carbosynth), *Macrocystis pyrifera* (Carbosynth), *Ecklonia maxima* (Carbosynth), *Sargassum fusiforme* (Carbosynth), *Cladosiphon okamurans* (Carbosynth) *Lessonia nigrescens* (Carbosynth), and *Durvillaea potatorum* (Carbosynth) as well as carrageenan oligosaccharides: neocarrabiose-4-O-sulfate sodium salt (Dextra), neocarrabiose-4-O-sulfate sodium salt (Dextra Neocarratetraose-41,43-di-O-sulfate sodium salt (Dextra), Neocarrahexaose-41,3,5-tri-O-sulfate sodium salt (Dextra), Neocarrahexaose-24,41,3,5-tetra-O-sulfate sodium salt (Dextra), D-Galactose-6-O-sulfate sodium salt (Carbosynth). Additionally, we have used concentrated culture supernatant from the following microalgae species: *Thalassiosira weissflogii*, *Thalassiosira psuedonana*, *Thalassiosira rotula*, and *Chaetoceros socialis*.

Phylogenetic and cluster analyses

The SulfAtlas (Barbeyron et al. 2016a) database was accessed on 22.05.2020 and total number of 1884 UniProtIDs (Bateman 2019) of S1_15 sulfatases were retrieved. 1685 identifiers could be directly matched to the UniprotKB database and notably, the remaining 199 UniProtIDs were obsolete (deleted or redundant) and only 165 of them could be accessed via UniParc. Alongside the protein sequences, we retrieved the TaxID of the corresponding organism. Additionally, 45 S1_15 sequences from '*Lentimonas*' sp. CC4 (Sichert et al. 2020) and 1P49 S1_3 sulfatase from *H. sapiens* (as outgroup) were manually added to the collection resulting in a total of 1896 sequences. To infer the habitat of microbes encoding S1_15 sulfatases, TaxIDs were queried in PATRIC (Wattam et al. 2017) database to retrieve the habitat, salinity or isolation source. Ambiguities were manually resolved by looking up the habitat of isolates in culture collections.

Next, we analyzed the domain architecture of sulfatases with hmmscan (Eddy 2011) and Pfam-A hmm modules. The domain architecture was analyzed with DAMA (Bernardes et al. 2016). Sulfatase domains (PF00884) were then blasted against the SulfAtlas database and assigned to S1 subfamilies by their respective highest scoring blast hit. Along S1_15, S1_3 and S1_39 sulfatase domains, the different following domains were identified: Sulfatase_C, Alpha_L_fucos, Big_9, CBM_4_9, Cadherin_5, DHH, DHHA1, DUF1080, DUF229, DUF4976, DUF4994, Lipase_GDSL_2, Phosphodiest, TAT_signal and Yai. One protein had two S1_15 domains and other fusion proteins are listed together with additional metadata in supplementary Table S2.

For phylogenetic analyses S1_15 domains were extracted and aligned with mafft-linsi ('mafft --localpair --maxiterate 1000') (Nakamura et al. 2018). A phylogenetic tree was calculated using the

ete3 toolkit v3.1.1 (Huerta-cepas 2016) and the ete3 toolchain function 'none-trimal_gappyout-none-fasttree_full'. The tree was rooted using 1P49 as outgroup and clades were defined by sequences with a phylogenetic distance of less than 0.7. Then, we calculated the relative abundance of the different phyla or habitat per clades. For simplicity, habitats or phyla with less than 10 occurrences were categorized as 'other'.

Sequence analyses

Putative signal peptide was identified using the SignalP4 (Nielsen 2017) and neglected in the cloning process. Closely related sulfatases were found by NCBI pBLAST using the PDB data base(Berman 2000). Protein sequences of structurally characterized sulfatases were retrieved from SulfAtlas (Barbeyron et al. 2016a). Conserved amino acids were identified by multiple alignment using a MUSCLE alignment (Edgar 2004) analyzed in WebLogo (Crooks et al. 2004).

Protein crystallization and X-ray data collection

Sitting-drop vapor diffusion was set up with 96-well sparse matrix and grid screens at 16 °C to find conditions, which allow the protein to crystallize. Conditions that showed crystallized protein were repeated and, optimized with slight modifications to the original recipe using hanging-drop vapor diffusion at 16 °C. 24-well hanging drop crystal trials were set up containing 500 µL mother liquor and three different dilutions (1:2, 2:1 and 1:1) of protein and crystallization solution with a total volume of 3 µL. Protein crystals were soaked in 25-30 % glycerol in the mother liquor and flash-frozen to cryo-temperatures using liquid nitrogen. Data collection was done at the PETRAIII (Burkhardt et al. 2016) beam line P11 (Meents et al. 2013) from Deutsches Elektronen Synchrotron (DESY, Hamburg).

Structure solution and refinement

Diffraction data were processed using XDS (Kabsch et al. 2010) and AIMLESS (Evans 2011) in CCP4i Cloud (Krissinel et al. 2018). The structures were solved by molecular replacement with PHASER (McCoy et al. 2007) and the structural homologue PDB:1E2S (Von Bülow et al. 2001). The model was built using CCP4Build (Winn et al. 2011), BUCCANER (Cowtan 2006) and PARROT (Cowtan 2010). Refinement was done based on a Maximum Likelihood method using REFMAC V5.5 (Murshudov et al. 2011), manually finalized in COOT V0.8.7.1 (Emsley et al. 2010) and visualized in PyMOL Version 2.0 (Schrödinger, LLC). Data collection and refinement statistics are displayed in Tab. S2. The protein structure was analyzed using DALI (Holm and Elofsson 2019), ConSurf (Glaser et al. 2003) and PDBSum (Laskowski et al. 2018). The atomic coordinates and structure factors have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

Acknowledgments and funding

We acknowledge DESY (Hamburg, Germany) for the provision of experimental facilities. Parts of this research were carried out at PETRAIII and we would like to thank Eva Corsa for assistance in using the beamline P11. We thank Stefan Becker and Craig S. Robb for data collection at P11, A. Boraston for providing pBAD/myc-his A vector containing FGE and Guoyin Huang for providing microalgae exudates. J.H.H. acknowledges funding from the Max-Planck-Gesellschaft and from the Emmy Noether Program of the DFG (grant number HE 7217/1-1).

Conflict of interest

The authors declare no conflict of interest.

Author's contributions

NG, AS and JHH designed the study. AS designed primer and performed phylogenetic analysis. NG cloned, expressed, purified and crystallized the enzyme. NG collected X-ray diffraction data, conducted biochemical and structural analyses. NG and AS visualized the data. NG, AS and JHH wrote the manuscript.

Supplementary information

In appendix.

Chapter 2.2. Marine bacterial fucosidase of family GH95 degrades macroalgal fucoidans

Nadine Gerlach[1,2], Jan-Hendrik Hehemann[1,2,*]

Affiliations

[1] MARUM – Center for Marine Environmental Sciences, University of Bremen, Germany

[2] Max Planck Institute for Marine Microbiology, Bremen, Germany

*Correspondence: jhhehemann@marum.de, +494212186775

ORCID:

[3] 0000-0003-2849-4455

[4] 0000-0002-8700-2564

Manuscript submitted to Applied Microbiology and Biotechnology.

Abstract

Fucoidans are fucose-containing sulfated polysaccharides which are characterized by a great structural and compositional diversity. This class of glycans has been known for decades, but our knowledge on how marine organisms deal with such a complex carbon source is still limited. The backbone of fucoidans is mainly composed of 2,3-sulfated fucose, which is degraded by fucoidanases, sulfatases and fucosidases of the carbohydrate-active enzyme (CAZymes) families S1_17, S_25, GH29 and GH107. Here we describe a fucoidan-degrading CAZyme of family GH95. In total, we have tested 17 fucose- and galactose containing oligosaccharides, exudates of four microalgal species as well as structurally different fucoidans from nine macroalgal species. Chromatography of reaction products suggests that 30_GH95 is an exo-acting α-1,2/3-L-fucosidase with additional minor activity towards β-D-galactose. Fucosidase activity was confirmed using macroalgal fucoidans as substrates. 30_GH95 has a strong substrate specificity for fucoidan from *Fucus vesiculosus*, but is also active on *Durvillaea potatorum*, *Sargassum fusiforme* and *Ecklonia maxima* of the orders Fucales and Laminariales. Fucosidase activity was not observed for structurally different fucoidans from the same orders (e.g. *Undaraia pinnatifida* and *Fucus serratus*) or the order Ecotocarpales. Our study contributes to a better understanding of the enzymatic cascade of marine microbes to degrade fucoidans.

Keywords: CAZyme, GH95, fucoidan, fucosidase, galactosidase

Introduction

Fucose-containing sulfated polysaccharides (FCSPs) or fucoidans are a structurally diverse class of glycans which can be found in seagrasses (Kannan et al. 2013), marine invertebrates (Ribeiro et al. 1994; Mulloy et al. 1994) and the cell walls of brown macroalgae (Deniaud-Bouët et al. 2017). FCSPs can be grouped into homoglycans, which mainly consist of sulfated fucose, or heteroglycans whose backbone is enriched in other monosaccharides. The backbone of the brown algae order Fucales is dominated by L-fucose linked by α-1,3 and to a lesser extend α-L-1,2 and α-L-1,4 (Bilan et al. 2006; Holtkamp et al. 2009). Members of the orders Laminariales and Ectocarpales haven been described as galacto- and xylofucans as it is the case for *U. pinnatifida* (Lee et al. 2004; Hemmingson et al. 2006; Skriptsova et al. 2010; Synytsya et al. 2010; Zhao et al. 2018) as well as *P. plantagenia* (Bilan et al. 2014), respectively. However, there are also species in both orders which are considered as homofucans such as *L. hyperborea* (Kopplin et al. 2018) and *C. okamurans* (Lim et al. 2019). Differences in composition, linkage and branching derive not only from different algal species, but also due to seasonal and geographical parameters (Anastyuk et al. 2010; Fletcher

et al. 2017). This structural diversity requires an enzymatic counterpart to be utilized by marine microbes. Fucoidan-degrading bacteria belong to the superphylum Planctomycetes-Verrucomicrobia-Chlamydiae (PVC) (Sakai et al. 2003; Wegner et al. 2013; van Vliet et al. 2019), Proteobacteria (Bakunina et al. 2002) and Bacteroidetes (Descamps et al. 2006; Barbeyron et al. 2008; Chen et al. 2016). Recently, it has been shown that the marine Verrucomicrobium '*Lentimonas'* sp. CC4 performs a multistep enzymatic pathway which combines desulfation with the hydrolysis of the polysaccharide (Sichert et al. 2020). This enzymatic cascade involves dozens of enzymes from various sulfatase and carbohydrate-active enzyme (CAZyme) families of which only a limited number has been described yet. To date, only one esterase (Nagao et al. 2017), three exo-acting sulfatases of families S1_17 and S1_25 (Daniel et al. 2001; Silchenko et al. 2018), two exo-acting fucosidases (Tanaka and Sorai 1970; Berteau et al. 2002) and several endo-fucoidanases mainly belonging to family GH107 (Cao et al. 2018; Schultz-Johansen et al. 2018) have been found to act on FCSPs. These enzymes are characterized by a pronounced substrate specificity and presumably act in synergy. For example, fucosidase GH29 from the marine mollusk *Pecten maximus* only cleaves fucose from *Ascophyllum nodosum* fucoidan and the addition of a sulfatase increases its hydrolysis yield by 75 % (Berteau et al. 2002). However, these studies represent only snapshots of the degradation pathway. We still require more detailed biochemical and structural investigations to elucidate the complete utilization process.

So far, glycoside hydrolase (GH) family GH95 (E.C.3.2.1-, .51, .63) has not been shown to act on FCSPs yet represents a promising candidate. Similar to family GH29, GH95s are exo-acting enzymes that cleave terminal α-L-fucose and galactose (Katayama et al. 2004; Nagae et al. 2007; Guillotin et al. 2014; Rogowski et al. 2015). These families are crucial to initiate glycan degradation in the human gastrointestinal tract (Hobbs et al. 2019), but the role of GH95 in the marine environment remains greatly unknown. Fucosidases and galactosidases of families GH29 and GH95 are among the dominating genes in heterotrophic bacteria associated with particles (Sichert et al. 2020), macroalgae (Kim et al. 2016) and microalgae blooms (Teeling et al. 2012; Kappelmann et al. 2019), suggesting that algal glycans are potential substrates for marine GH95s. To shed light on the role of family GH95 in fucoidan degradation, we performed a biochemical characterization of a GH95 which stems from a marine Verrucomicrobium. '*Lentimonas'* sp. CC4 grows on various macroalgal fucoidans and regulates glycan utilization in several specific polysaccharide utilization loci (PULs), including several GH95, which are mainly located on a plasmid (Sichert et al. 2020). We found 30_GH95 is an α-1,2/3-L-fucosidase with minor activity towards β-1,4-D-galactose and a pronounced substrate specificity for the fucoidan from the macroalgal species *Fucus vesiculosus.*

Material and methods

Materials. All chemical and reagents were purchased from Sigma Aldrich (Germany) if not stated otherwise.

Production of recombinant plasmid. The N-terminal signal peptide was predicted using SingalP V4.1 sever (Nielsen 2017) with default options. The forward and reverse primers (Tab. S1) were designed with a single N-terminal 6-His-Tag (Hochuli et al. 1988) and a ~30 bp overhang and complementary to the pet28a(+) expression plasmid (69864, Merck Millipore). Genes were amplified with Q5 High-Fidelity DNA Polymerase (M0492, New England Biolabs) from genomic DNA of '*Lentimonas'* sp. CC4. Recombinant plasmids were assembled using Gibson Assembly (E2611, New England Biolabs). Purified plasmids were transformed into *Escherichia coli* NEB5α for subcloning and into *Escherichia coli* BL21(DE3) for overexpression according to the manufacturer's instructions (C2987I+C2527H, New England Biolabs). Transformed cells were grown on lysogeny broth (LB) agar plates with 50 µg mL^{-1} kanamycin at 37 °C overnight.

Protein expression and purification. Bacterial cultures were grown in 5 L Erlenmeyer flasks with 1 L of ZYP5052 auto-induction medium (Studier 2005) and 100 µg mL^{-1} kanamycin for four days at 20 °C and 150 rpm or with 1 L LB and 100 µg mL^{-1} kanamycin at 37 °C until OD >0.6 and induced with 1 mM IPTG at 16 °C at 150 rpm overnight. Bacterial cultures were harvested at 4900 *g* and for 30 minutes and chemically lysed. The cell pellet was resuspended in 15 mL buffer (25 % sucrose, 50 mM Tris pH 8), 30 mg of lysozyme (~ 7000 U mg^{-1}) was added and the mixture was incubated at room temperature (~ 25 °C) for 15 minutes while stirring. 30 mL of deoxycholate solution (20 mM Tris pH 8, 1 % w/v deoxycholate, 100 mM NaCl, 1 % w/v Triton X-100), $MgSO_4$ (5 mM final concentration) and 100 µL of 10 mg ml^{-1} DNase I (≥ 400 Kunitz mg^{-1}) were added. The mixture was incubated at room temperature (~ 25 °C) until it was no longer viscous and centrifuged at 30966 *g* and 4 °C for 30 minutes.

Protein purification was done via immobilized metal affinity chromatography (IMAC) carried out in ÄKTA™ start chromatography system (29-0220-94, GE Healthcare Life Sciences) with a 5 mL cobalt column (HiTrap™ Talon® crude, 28953766, GE Healthcare Life Sciences). The supernatant from chemical lysis was applied to the column equilibrated in binding buffer (20 mM Tris pH 8, 500 mM NaCl) at a flow rate of 5 mL min^{-1}. The column was subsequently washed with 40 mL binding buffer followed by elution with a linear gradient of 0 – 100 % elution buffer (20 mM Tris pH 8, 0.5 M NaCl, 0.5 M imidazole) over 25 mL. IMAC fractions were verified for purity via sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) for soluble proteins (LAEMMLI 1970). The 1 mm polyacrylamide gel was run at constant 200 V for 40 minutes in 1x Tris-Glycine-SDS (TGS) buffer. Purified protein fractions were concentrated in a stirred ultrafiltration unit (Amicon) using 70 psi nitrogen pressure and a 10 kDa membrane (Biomax®).

The concentration was determined measuring the absorbance at 280 nm (BioSpectrophotometer basic, Eppendorf) using the molar extinction coefficient of 151525 M^{-1} cm^{-1} computed by ExPASy ProtParam (Wilkins et al. 2005).

Thermal stability. Thermal stability was determined via dynamic light scattering (DLS) using a DynaPro Plate reader and the according Dynamics Software (Wyatt). Triplicates of 30 µl IMAC purified protein in 250 mM NaCl, 20 mM TRIS pH 8 were transferred into a microtiter plate (Aurora), centrifuged at 4500 *g* and 4 °C for 10 min and covered with an oil droplet (Zeiss). Thermal stability was monitored using a temperature gradient of 25 – 80 °C with an increase of 0.1 °C min^{-1} and 5 acquisitions per sample each measured for 5 seconds.

Analytical size exclusion chromatography. Molecular size and the state of oligomerization were determined via analytical size exclusion chromatography (SEC) using protein gel filtration standards (#151-1901, Bio-Rad). SEC was carried out at 1 ml min^{-1} using the ENrich™ SEC 650 10 x 300 Column (#7801650, Bio-Rad) connected to a NGC™ chromatography system (Bio-Rad) and operated at room temperature. IMAC purified protein was centrifuged for 30 min at 13000 *g* and 4 °C and 200 µl protein in 250 mM NaCl, 20 mM TRIS pH 8 were analysed. Protein elution of standards and sample was monitored at 280 nm.

Enzyme activity assay. The physiochemical properties of 30_GH95 were determined spectrophotometric by measuring enzyme activity at 405 nm in FLUOstar Omega microplate reader (BMG Labtech GmbH). The following substrates were prepared in 10 mM stock solution in Milli-Q incubated with 0.05 mg enzyme ml^{-1}: cNP-α-L-Fuc (Carbosynth), pNP-β-L-Fuc, pNP-β-D-Fuc, pNP-β-D-Gal and pNP-α-D-Gal. The influence of pH, NaCl and 1 mM divalent cations were investigated in triplicates in a total reaction volume of 100 µL and 1 mM cNP-α-L-Fuc.

Substrate specificity via HPAEC-PAD. Substrate specificity and linkage specificity was analyzed via high-performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) as previously described (Engel and Händel 2011). In short, ~ 1.2 mg ml^{-1} enzyme in 250 mM NaCl and 20 mM TRIS pH 8 were incubated with 9 $mgml^{-1}$ fucoidan or 0.5 mg ml^{-1} oligosaccharides, 0.1 M NaCl and 20 mM Na-citrate pH 5. at 30 °C and 200 rpm overnight (18-22 h) and the reaction was terminated at 99 °C for 10 min. HPAEC-PAD was carried out on a Dionex ICS5000+ system (Dionex) using a CarboPac™ PA10 (2 mm x 250 mm, Dinoex) analytical column and a CarboPac™ PA10 (2 mm x 50 mm, Dinoex) guard column. The following substrates were tested: 1,3/4/6-β-D-galactobiose (Megazyme), blood group H disaccharide (BGH, Carbosynth), blood group B tetraose type 5 (BGB, Elicityl), $Lewis_y$ tetraose (Le_y, Elicityl), $Lewis_b$ tetraose (Le_b, Elicityl), $Lewis_x$ triaose (Le_x, Elicityl), $Lewis_a$ triaose (Le_a, Carbosynth), 3-fucosyllactose (3-FL, Elicityl), reduced XFG xyloglucan (XFGol, Elicityl), 2'-fucosyllactose (2'-FL, Carbosynth) as well as ion exchange chromatography-purified macroalgal fucoidans (Tab. S2). Additionally, we have

used concentrated culture supernatant from the following microalgae species: *Thalassiosira weissflogii, Thalassiosira psuedonana*, *Thalassiosira rotula*, and *Chaetoceros socialis*. L-Fucose and D-galactose were used as internal calibration standards. The yield of fucosidase activity (%) was calculated based on the fucose content reported for macroalgal fucoidans elsewhere (Sichert et al. 2020).

Multiple sequence alignment and homology modelling. Closely related GH95 were found by NCBI pBLAST using the PDB data base (Berman 2000). Structural homologues were retrieved from CAZy (Lombard et al. 2014b). Conserved residues of the active site were identified by multiple sequence alignment using MUSCLE (Edgar 2004) in ESPript V3.0 (Robert and Gouet 2014). 30_GH95 was modelled using Phyre2 (Kelley et al. 2015) and visualized in PyMOL version 2.2.3 (Schrödinger, LCC). Protein crystallization was set up at 16 °C with IMAC purified protein as described elsewhere (Reisky et al. 2019).

Results

Identification as putative fucoidan CAZyme and domain architecture

30_GH95 is located on the plasmid of '*Lentimonas'* sp. CC4 in an operon, which is upregulated in the presence of *Fucus vesiculosus* fucoidan as sole carbon source (Sichert et al. 2020). Neighboring CAZymes and associated enzymes of 30_GH95 include a putative endo-fucoidanase of family GH107, putative galactosidases of families GH97 and GH117 as well as a sulfatase of subfamily S1_28.

NCBI pBLAST of 30_GH95 revealed 69 - 67 % sequence identity (99 - 97 % coverage) with GH95s from fucoidan-degrading bacteria *Kiritimatiellales bacterium* (WP_136078781.1), *Planctomycetes bacterium* (WP_146599266.1) and *Rhodopirellula* sp. (EMI42024.1) of the PVC superphylum. In contrast, 30_GH95 shared only 30 and 35 % sequence identity with structurally and biochemically characterized galactosidases (Rogowski et al. 2015) and fucosidases (Nagae et al. 2007), respectively.

The functional domain architecture of 30_GH95 indicated a N-terminal signal peptide of 0-19 amino acids and the catalytic domain of 20-819 amino acids. The signal peptide, which suggest a periplasmatic or extracellular location of the enzyme in '*Lentimonas'* sp. CC4, was removed by cloning to produce the recombinant protein in *E. coli*. 30_GH95 has a theoretical molecular weight of 91.3 kDa and a predicted isoelectric point of pH 6.59. The molecular weight was confirmed by SDS-PAGE of IMAC-purified enzyme showing a single band of > 80 kDa. Analytical SEC indicated a monomeric oligomerization state of this protein in solution (Fig. 1A).

30_GH95 has dual fucosidase and galactosidase activity

Family GH95 are exo-acting CAZymes that hydrolase the glyosidic bond of α-L-fucose or α-L-galactose via the inverting hydrolysis mechanism (Katayama et al. 2004). To determine the substrate and linkage specificity of 30_GH95 we tested 17 fucose and galactose containing substrates including chromogenic glycoconjugates, human milk- and blood oligosaccharides. Prior activity assays, the thermal stability of 30_GH95 was determined to ~ 33 °C (Fig. 1B). Therefore, enzymatic digests were set up at 30 °C.

First, we investigate the linkage and substrate specificity of 30_GH95 using chromogenic substrates. Spectrophotometric analysis revealed that 30_GH95 is active on cNP-α-L-fucose and to a lesser extent on pNP-β-D-galactose. β-D- and α-D-linked fucose as well as α-D-linked galactose were not cleaved by 30_GH95 (Fig. 1C).

Next, we determined the physiochemical properties of 30_GH95 using cNP-α-L-fucose which were used in the following experiments. 30_GH95 has a pH optimum of 5.7 in Na-citrate buffer (Fig. 1D) and the activity is boosted in presence of > 0.3 M NaCl (Fig. 1E), which suggests an extracellular localization exposed to the high salt concentration of seawater (3.5 % w/v ≈ 600 mM NaCl). The addition of inorganic salts or mono- and divalent cations prevalent in seawater had neither a weakening nor stimulating effect on enzyme activity (Fig. 1F). Similar to GH29, biochemical studies have reported acidic pH optima (Léonard et al. 2008) as well as the ability and disability (Fan et al. 2016) to utilize pNP substrates for GH95s.

To expand the linkage specificity, we measured fucose and galactose release from nine human milk- and blood oligosaccharides monitored by HPAEC-PAD. 30_GH95 was most active on 1,2-linked fucose from 2'-fucosyllactose, blood group H and XFG xyloglucan and to a limited extend also on 1,3-linked L-fucose of $Lewis_x$ and on 1,4-linked galactose of 2'-fucosyllactose. 1,3/4/6-β-D-galactobiose was not cleaved by 30_GH95 as well as 3-fucosyllactose, $Lewis_a$ DP3, $Lewis_y$, $Lewis_b$ and blood group B (Tab. S3).

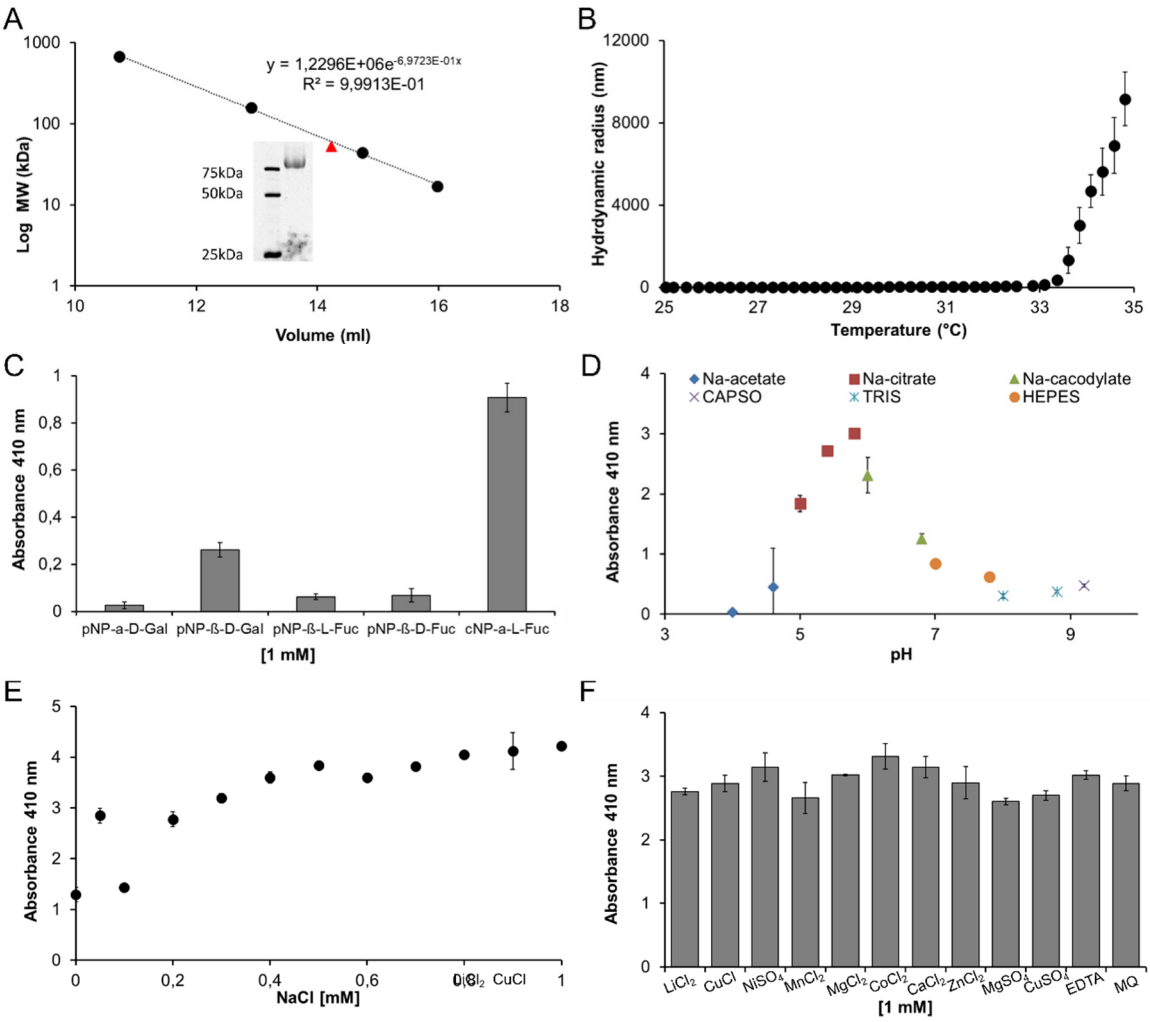


Fig. 1. Activity screening of 30_GH95. (A) SDS-PAGE and analytical size exclusion of standards (●) and 30_GH95 (▲). (B) Thermal stability via DLS. (C) Substrate and linkage specificity of chromogenic substrates. Physiochemical parameters using cNP-α-L-fucose: Influence of (D) pH using different buffers at 20 mM, (E) NaCl using 20 mM Na-citrate pH 5.7 and (F) cofactors using 20 mM Na-citrate pH 5.7 and 200 mM NaCl.

Fucosidase activity on fucoidans

Besides macroalgal fucoidans, fucose and galactose are dominant compounds of microalgae and their exudates (Gügi et al. 2015; Gaignard et al. 2019). To identify putative natural substrates of 30_GH95, we tested nine structurally different fucoidans and the culture supernatant of four microalgal species. No activity was observed for the exudates of the microalgal species *T. weissflogii*, *T. psuedonana*, *T. rotula* and *C. socialis*.

Fucosidase activity was confirmed for macroalgal fucoidans of the order Fucales and Laminariales. 30_GH95 cleaved ~ 0.3, 0.22, 0.03 and 0.018 % of the total fucose content from *F. vesiculosus, D. potatorum*, S. *fusiforme* and *E. maxima,* respectively. No fucose release was observed for *F.*

serreatus or other members of the order Laminariales (*U. pinnatifida, M. pyrifera, L. nigrescens*) and Ectocarpales (*C. okamurans*). Galactosidase activity on macroalgal fucoidans was not detected (Fig. 2).

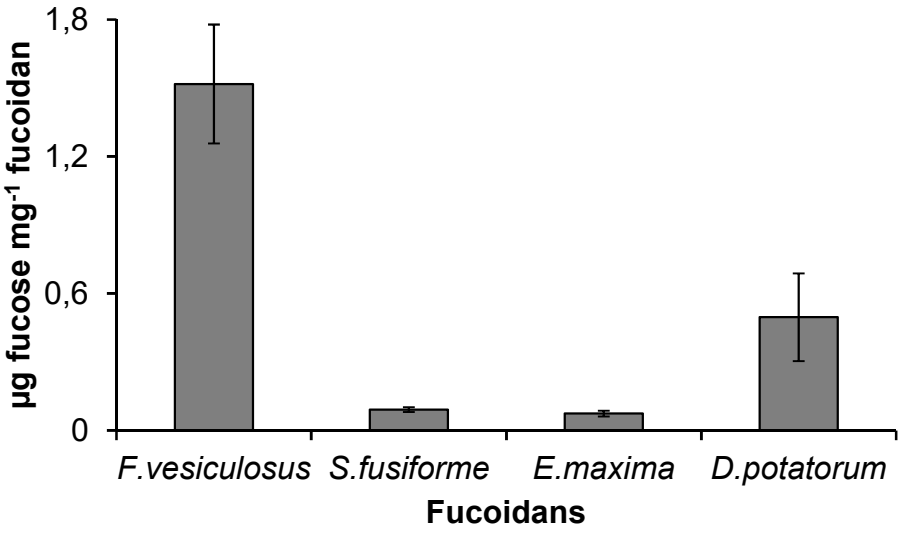


Figure 2. Fucosidase activity against macroalgal fucoidans. 30_GH95 was incubated with fucoidans overnight and fucose release was measured via HPAEC-PAD. Fucosidase activity is shown as released µg fucose per mg fucoidan. Error bars represent standard deviation of technical triplicates.

Sequence analysis indicates putative novel active site residue

The active site of family GH95 is highly conserved, consisting of 11 charged and hydrophobic amino acids L-W-H/T-N-N-E-R-H-W-H-D (Fig. 3). In 30_GH95, the first conserved Asn, which orientates the side chain of the putative catalytic residue Glu522, is replaced by Asp398 (Fig. 3), which has not been reported so far.

The catalytic residues include glutamic acid (Glu) and asparagine (Asn) that act as acid / base pair as well as aspartic acid (Asp), which activates the base Asn and a second Asn orienting the side chain of Glu (Nagae et al. 2007). Recently, it has been proposed that the His/Thr polymorphism in GH95s determines substrate specificity within this family (Rogowski et al. 2015). In 30_GH95 and closely related fucosidases this residue is a histidine (His), which is replaced by threonine (Thr) in the galactosidase from *B. ovatus* (Fig. S1). Since 30_GH95 is characterized by a dual fucosidase and galactosidase activity, the His/Thr polymorphism in GH95 might be involved in substrate binding and recognition. To shed light on the substrate specificity within this family, we planned to perform X-ray crystallography but 30_GH95 did not crystallize in any of the > 760 conditions tested yet. Therefore, we combined multiple sequence alignment with a structural homology model. The structural model of 30_GH95 (94 % coverage modelled at > 90 % accuracy) indicated a small

pocket architecture of the active site with a negative electrostatic potential (Fig. S2), suggesting that only desulfated terminal fucose or galactose fit into the small pocket. Exo-acting enzymes are known for their small pocket-like active site (Davies et al. 1997; Hettle et al. 2018). CAZymes involved in the degradation of anionic polysaccharides are often characterized by a cationic surface charge surrounding the active site (Reisky et al. 2019). These local patches are oppositely charged to the anionic glycan and thus, establish electrostatic protein-carbohydrate interactions characterized by a neutral charge (de Kruif et al. 2004).

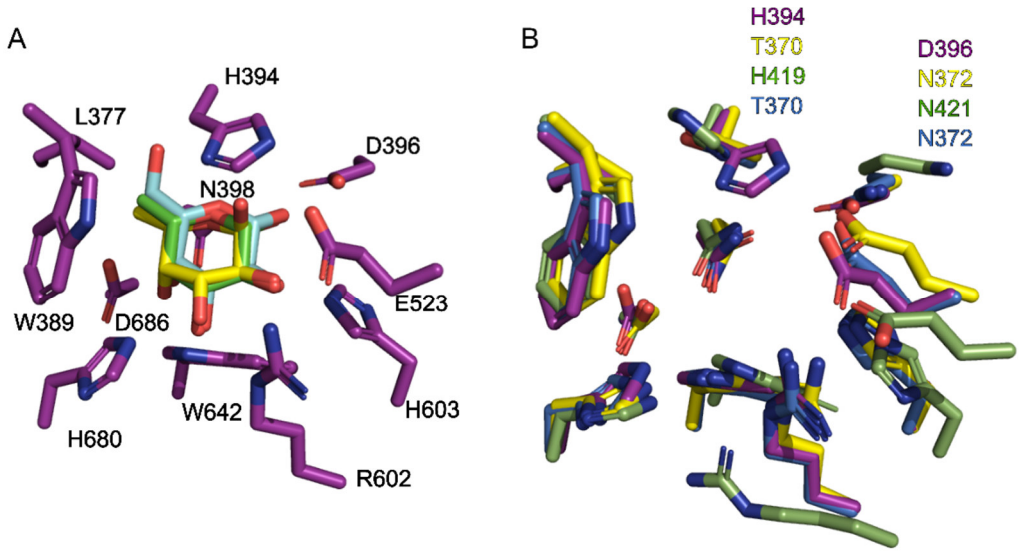


Fig. 3. Novel active site residue D396 in 30_GH95. (A) Active site residues of 30_GH95 (magenta) in superimposition with α-fucose (yellow), β-fucose (green), and β-galactose (turquois). The residues were identified with multiple sequence and structural alignment. The model was generated by Phyre2. Confidence in the model: 769 residues (94 % coverage) modelled at > 90 % accuracy. (B) Structural alignment of active site residues from galactosidase (PDB: 4UFC, yellow), fucosidase (PDB: 2EAB, green) and uncharacterized GH95 (PDB: 2RDY, blue). Modified after (Rogowski et al. 2015). Structural differences are labeled with amino acid number.

Discussion

Fucoidan-degrading bacteria perform a multistep degradation pathway to utilize FCSPs. The enzymatic cascade depends on the source and type of fucoidan but involves several esterases, fucosidases, sulfatases and fucoidanases (Sichert et al. 2020). To shed light on the role of family GH95 in FCSP degradation, we performed a biochemical characterization of a GH95 originating from the fucoidan-degrading Verrucomicrobium '*Lentimonas'* sp. CC4. 30_GH95 is an exo-acting α-1,2-L-fucosidase with minor activity towards α-1,3-L linked fucose and β-1,4-D-galactose.

We have tested fucoidans from nine different macroalgae species which differ in composition, linkage and the degree of sulfation. We observed a pronounced fucose release from *F. vesiculosus*, which is in line with proteomic analysis. We could detect minor fucosidase activity on *D. potatorum, S. fusofirme* and *E. maxima*. All species belong to the order Fucales and Laminariales and possess different structural features. While *F. vesiculosus* is characterized by an alternating α-1,3/4-L-fucose backbone (Chevolot et al. 2001; Bilan et al. 2004; Bilan et al. 2006), α-1,2-L-linked fucose is found in *C. okamurans*, *E. maxima* and *F. serratus* (Bilan et al. 2006; Sichert et al. 2020). Based on our results using fucosylated oligosaccharide from human milk and blood, it is likely that 30_GH95 cleaves α-1,2/3-linked fucose from *F. vesiculosus, D. potatorum, S. fusiforme* and *E. maxima*. The differences in the yield as well as no activity towards the other five fucoidans presumably reflects the amount of accessible fucose present in these four fucoidans. The highest yield was observed for *F. vesiculosus*, which has been considered a homofucan (Holtkamp et al. 2009). Exo-acting enzymes can only cleave fucose from the terminus of the glycan backbone or branches attached to it. Fucose decorated with sulfate or acetyl groups are only accessible in the combination with enzymatic partners (Berteau et al. 2002). The type and degree of such decorations vary among fucoidan species. While fucoidans are sulfated between 20 – 40 % (Fitton et al. 2015; Sichert et al. 2020), only few species have been reported to contain O-acetylation. Fucoidan from *C. okamurans* for instance has a α-1,3-linked fucose backbone with branched of a α-1,2-fucose. The absence of activity of 30_GH95 might be explained by sulfated and acetylated fucose (Nagaoka et al. 1999; Lim et al. 2019; Sichert et al. 2020).

Beside fucosidase activity, we also observed minor activity for β-D-galactopyranosidase and β-1,4-galactose in 2'fucosyllactose. The absence of activity on β-1,3/4/6-galactobiose is most likely explained by the active site architecture of 30_GH95. The catalytic activity of glycoside hydrolases is affected by their substrate-binding sites. These subsites recognize and subsequently accommodate each monosaccharide of the substrate (Davies et al. 1997). Only if these criteria are fulfilled hydrolysis is taking place. In the case of galactobiose it could be that it is not recognized by the subsites of 30_GH95 and hence not cleaved.

Galactose is a common monosaccharide found in many fucoidans (Fitton et al. 2015; Sichert et al. 2020). *S. fusiforme* and *E. maxima* possess α-1,4 and α-1,3/4/6-linked galactose, respectively (Hu et al. 2014; Hu et al. 2016; Sichert et al. 2020). So far, β-linkages were only reported for *U. pinnatifida* containing 1,3/4/6-linked galactose. Galactose is heavily sulfated at C3, C3,4 and C3,6 and by that presumably inaccessible for 30_GH95 (Lee et al. 2004; Hemmingson et al. 2006). Since we did not observe galactose release from macroalgal fucoidans, we were wondering for alternative natural substrates. Fucose and galactose are dominant compounds in microalgae cells and their exudates (Gügi et al. 2015; Gaignard et al. 2019) but little is known on their linkages. Therefore, we have tested the supernatant of four microalgae species. However, no fucose nor galactose release

was observed. This observation could be a result of the 3D structure of FCSPs. α-1,2/3 linked L-fucose and β-1,4-D-galactose might either (i) not be present in the tested exudates or (ii) they remain shielded by any type of decoration, or (iii) are not accessible prior enzymatic hydrolysis of other CAZymes such as endo-acting GH107s or sulfatases generating smaller fragments.

In conclusion, we report the biochemical characterization of a marine, bacterial GH95. 30_GH95 is an exo-acting α-1,2-L-fucosidase with minor activity towards α-1,3-L-fucose and β-1,4-galactose. Family GH95 has been previously described as fucosidases or galactosidases (Katayama et al. 2004; Nagae et al. 2007; Rogowski et al. 2015; Hobbs et al. 2019), but a dual activity of this family has not been reported yet. To the best of our knowledge this is the first study demonstrating activity on macroalgal fucoidan for family GH95. Our results contribute to the ongoing investigation on fucoidan utilization by marine microbes, but also lead the way for future biotechnological applications of FCSPs. In recent years, enzyme-assisted methods for the detection, extraction and quantification of glycans have become of raising interest (McCleary 1981; Vidal-Melgosa et al. 2015; Becker et al. 2017). These methods take advantage of highly specific enzymes and are often less time-, lab- and cost-intensive than other techniques. Fucoidans are heavily studied for their medical and therapeutic properties (Ale et al. 2011; Fitton et al. 2015) and as additives for food (Vo and Kim 2013; Zeuner et al. 2019) and cosmetics (Ale and Meyer 2013). Substrate specific enzymes like 30_GH95 can not only be used to investigate the structure of FCSPs (Ale and Meyer 2013; Nguyen et al. 2020), but also to modify the glycan, which can be then used in industry and pharmacy (Holtkamp et al. 2009).

Funding information

J.H.H. acknowledges funding from the Max-Planck-Gesellschaft and from the Emmy Noether Program of the DFG (grant number HE 7217/1-1).

Acknowledgements

We thank Alek Bolte for assistance with HPEAC-PAD, Guoyin Huang and Andres Sichert for providing microalgal supernatants and purified fucoidans.

Compliance with ethical standards

Conflict of interest

The authors declare no conflict of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Author contributions

NG: conducted experiments, data analyses and visualization. NG & JHH: wrote the paper.

Supplementary information

In appendix.

Chapter 2.3. Comparative biochemical and structural analyses of marine fucosidases of family GH29

Nadine Gerlach[1,2], Andreas Sichert[1,2], Ronan M. Keegan[3], Atharva Patharkar[4], Tatjana von Rosen[1,2,5], Craig S. Robb[1,2,6], Jan-Hendrik Hehemann[1,2,*]

Affiliations

[1] MARUM – Center for Marine Environmental Sciences, University of Bremen, Germany

[2] Max Planck Institute for Marine Microbiology, Bremen, Germany

[3] STFC, Rutherford Appleton Laboratory, Research Complex at Harwell, Didcot OX11 0FA, England

[4] University of Bremen, Germany

[5] Current address: Department of Biology, ETH Zurich, Wolfgang-Pauli-Strasse 27, 8093 Zurich, Switzerland

[6] Current address: Department

*Correspondence: jhhehemann@marum.de, +494212186775

ORCID:

NG: 0000-0003-2849-4455

AS: 0000-0002-1386-5535

JHH:0000-0002-8700-2564

Manuscript in preparation for Journal of Biological Chemistry. Data analyses is still in process.

Abstract

Fucoidans are a diverse class of fucose-containing sulfated polysaccharides. Owing to their structural diversity, fucoidans can only be degraded by highly specialized bacteria such as the marine Verrucomicrobium '*Lentimonas'* sp. CC4 possessing over 100 enzymes. Since these enzymes are interlocked in a complex degradation cascade, each enzyme only yields minuscule amounts of product challenging the characterization of fucoidan active enzymes without their cognate partner. To circumvent this, we have biochemically characterized 15 fucosidases of family GH29 from '*Lentimonas'* sp. CC4 of which four were also analyzed by X-ray crystallography. The analyses of reactions products of in total 14 chromogenic substrates and oligosaccharides by chromatography and spectrophotometry showed distinct substrate and linkage specificity. Moreover, we investigated the putative natural substrates by using structurally different fucoidans originating from nine macroalgal species. The pronounced substrate specificity of each enzyme suggest that they are part of several enzymatic pathways specialized on a specific fucoidan. Our findings not only gain further insights into the microbial remineralization of fucoidan, but also into the role of '*Lentimonas'* sp. CC4 as resource for novel fucoidan active enzymes.

Key words: GH29, fucosidase, fucoidan, X-ray crystallography, galactosidase, marine

Introduction

Fucoidans are diverse class of sulfated cell wall polysaccharide from brown algae with complex degradation pathway and a high environmental importance. Fucoidans, which make up 23% of algal dry weight, are highly branched and sulfated and their structure varies depending on algal species of origin and growth season (Skriptsova et al. 2010; Deniaud-Bouët et al. 2014; Deniaud-Bouët et al. 2017). For example, fucoidan from *Fucus vesiculosus* has a backbone of alternating α-1,3/α-1,4 linked L-fucose with sulfate esters on O-2 or O-3 with additional mannose and xylose residues (Nishino et al. 1994; Chevolot et al. 2001). Due to this structural complexity, fucoidans can only be degraded by highly specialized microbes such as the marine Verrucomicrobium '*Lentimonas'* sp. CC4 employing over 100 enzymes (Sichert et al. 2020). Consequently, fucoidans accumulate e.g. over a diatom spring bloom and have a high potential to sequester carbon in the ocean (Vidal et al. in preparation). Although previous studies identified potential pathway for fucoidans, we lack the biochemical characterization of fucoidan active enzymes hampering our understand why fucoidans are difficult to turn-over in the environment.

The enzymatic degradation of fucoidan is a multi-step reaction with several interlocked enzymes only active with their cognate enzyme partner. The activity of characterized fucoidanases is

typically low, yielding only a few percent degradation, but the yield can increase by the adding enzymes (Vickers et al. 2018; Schultz-Johansen et al. 2018). For example, the carbohydrate esterase (CE) increases the activity of an endo-enzyme and sulfatase the activity of a GH29 exo-fucosidases (Berteau et al. 2004; Nagao et al. 2017). This mutual dependency of enzymes complicates the biochemical analyses of a single fucoidanase without its cognate partner. Enzymes that act together on a polysaccharide are regulated in polysaccharide utilization loci (PUL) and therefore, PULs provide an opportunity to identity cognate enzymatic partners. For example, the marine *Bacteroidetes Wenyingzhuangia fucanilytica* CZ1127^T has a PUL encoding a S1_17 and S1_25 exo-sulfatases that stepwise degrade 2,3-di-O-sulfated fucooligosacchrides (Silchenko et al. 2018). The Verrucomicrobium '*Lentimonas'* sp. CC4 organizes its enzymes in operons, and therefore, these operons represent a resource to identify enzymes that act in synergy on fucoidan.

The degradation of fucoidan or fucoidan-derived oligosaccharides is catalyzed by exo-acting fucosidases and sulfatases. Proteomics on the marine Verrucomicrobium '*Lentimonas'* sp. CC4 revealed 35 homologs of GH29 fucosidases suggesting those enzymes are key to degrade fucoidans (Sichert et al. 2020). GH29 (EC 3.2.1.51 and .111) hydrolyze terminal α-1,2/3/4 or /6-linked fucose using a retaining mechanism (White Jr. et al. 1987; Eneyskaya et al. 2001b; Cobucci-Ponzano et al. 2003b). While the mechanistic motif of terrestrial fucosidases involved in glycan degradation is well studied (Hobbs et al. 2019), our knowledge on marine fucosidases is very limited. It appears that fucosidases, similar to other fucoidan-degrading enzymes, are characterized by a high substrate specificity. For instance, GH29 from the marine mollusk *Pecten maximus* is exclusively active on *Ascophyllum nodosum* fucoidan (Berteau et al. 2002), while the fuscosidase from the marine bacterium *Wenyingzhuangia fucanilytica* is only able to act on partially degraded fucoidan but not on the polysaccharide (Dong et al. 2017). These results suggest that the ability of exo-acting fucoisidases is limited to the amount of accessbile fucose. Therefore, it is likely that fucosidases require enzymatic partners to act synergistically such as sulfatases (Berteau et al. 2002; van Vliet et al. 2019) and/or endo-acting fucoidanases which generate fucoidan oligosaccharides. Interestingly, GH29s share the overall fold with endo-acting fucoidanases of family GH107 (Vickers et al. 2018). The catalytic active site is highly conserved in a $(\beta/\alpha)_8$ triosephosphate isomerase (TIM) barrel domain (Sulzenbacher et al. 2004; Sela et al. 2012; Summers et al. 2016). Typically, the catalytic residue Asp acts as nucleophile and Glu as Bronsted acid/base (Sulzenbacher et al. 2004). Since the catalytic are located on different loops up to 50 amino acids apart, they are difficult to identify, and these insights can only be derived from protein structures. Only recently, the first substrate-complex of GH29 with $Lewis_X$, $Lewis_Y$ and $Lewis_A$ oligosaccharides revealed unusual substrate binding pockets with a selectivity for Gal in the +2 subsite, GlcNAc in the +1 subsite and flexible cleavage of α1,3 and α1,4 L-Fucose in the -1 subsite

(Hobbs et al. 2019). Notably, no marine GH29 was crystallized with a native substrate, illustrating our limited understanding of marine fucosidases.

In this study, we have structurally and biochemically characterized 15 GH29s from the marine Verrucomicrobium '*Lenitmonas'* sp. CC4. This is the first in-depth study within this family. Our findings lead the way for a reclassification of family GH29 into subfamilies reflecting their substrate and linkage specificity.

Results and discussion

Domain architecture and expression of marine GH29s

'*Lenitmonas'* sp. CC4 possess 35 GH29s of which 13 are no regulated / expressed, while the remaining ones are regulated in four operons specialized on fucoidans from brown macroalgae (Sichert et al. 2020). GH29s from '*Lenitmonas'* sp. CC4 are characterized by a similar domain architecture consisting of a putative N-terminal signal peptide and a linker connecting the catalytic domain harboring the active site with a C-terminal binding domain of families 6, 32 or 35.

We produced 24 GH29s as recombinant plasmids of which 15 were soluble expressed in *E. coli* BL21, purified by IMAC and further analyzed to reveal mechanistically and structurally differences among marine GH29s. Besides LCC4_2_71, all GH29s possess an N-terminal signal peptide, suggesting a periplasmic or extracellular location. All recombinant proteins possess a theoretical molecular mass of 53 – 68 kDa (Tab. S1), which was confirmed by SDS-PAGE as well as analytical SEC. LCC4_2_293 and LCC4_2_169 are characterized by a dimeric and trimeric oligomerization state, respectively. All other GH29s are monomers except LCC4_2_171, which partially exhibits a dimeric oligomerization state presumably due to protein aggregation.

Marine GH29s are fucosidases and galactosidases

GH29s are exo-acting α-L-fucosidases with a broad substrate and linkage specificity. While some GH29s exclusively cleave α-1,3 or α-1,4-linked fucose, others are capable to hydrolyze both linkages. Interestingly, the GH29 from the gut bacterium *B. thetaiotamicron* exhibits weak activity against α-L-fucose, while pNP-β-D-galactopyranose is efficiently hydrolyzed. Compared to fucosidases, GH29 from B. *thetaiotamicron* possess a second pocket to bind β-D-galactose (Guillotin et al. 2014). As fucose and galactose are abundant compounds of macroalgal fucoidans, we tested 14 fucose- and galactose containing substrates to shed light on the mechanistic motif of marine GH29.

Prior activity screening, the thermal stability of each enzymes was determined varying from 33 – 46 °C (Tab. S1). All enzyme digests were performed at 30 °C. None of the 15 GH29s from '*Lentimonas'* sp. CC4 exhibited detectable activity against pNP-α-D-galactopyranose, pNP-β-D-fucopyranose or pNP-β-L-fucopyranose. LCC4_2_103, LCC4_2_67, LCC4_2_71, LCC4_2_293 and LCC4_2_269 only hydrolyzed pNP-α-L-fucopyranose, whereas six GH29s (LCC4_1_2701, LCC4_2_230, LCC4_2_169, LCC4_2_171, LCC4_2_22, LCC4_2_218) are characterized by a dual activity towards pNP-α-L-fucopyranose and pNP-β-D-galactopyranose. GH29s LCC4_2_210 showed no activity against any of the five 4-nitrophenyl substrates, whereas LCC4_2_211 exclusively hydrolyzed pNP-β-D-galactopyranose (Fig. S1).

A similar pattern for linkage specificity was also observed using nine human blood and milk oligosaccharides (Tab. 1). LCC4_2_67, LCC4_2_293 and LCC4_2_211 were not able to cleave any of the tested substrates despite their ability of hydrolyzing pNP-α-L-fucopyranose and -β-D-galactopyranose, respectively. LCC4_2_210 was not active ony any of the pNP-substrates, but cleaves α-L-1,2/3-linked fucose from reduced XFG xyloglucan, 2'-fucosyllactose and 3-fucosyllactose. LCC4_1_2701, LCC4_2_71 and LCC4_2_269 exclusively cleave α-L-1,2-linked fucose, whereas LCC4_2_22 and LCC4_2_169 hydrolyze α-L-1,3 linkages. Besides LCC4_2_210, LCC4_2_171 and LCC4_2_230 are capable of cleaving α-L-1,3- and α-L-1,4-linkages. The activity of the GH29s

Lastly, we used pNP-α-L-fucopyranose to determine the biochemical specs of each enzyme (Tab. S1). GH29s from '*Lentimonas'* sp. CC4 are most active in acidic pH of 5.4 or 5.7, which has been reported for other GH29s (Cao et al. 2014). LCC4_2_71, LCC4_2_171 and LCC4_2_218 are characterized by a broader pH tolerance of pH 5.4 to 6.6 or even pH 8. Based on the signal peptide prediction, GH29s are probably excreted and located in the periplasm or extracellular. Therefore, we tested the effect of sodium chloride as well as divalent cations present in the seawater on enzyme activity. However, the activity of only two-thirds of the GH29s was increased in the presence of ≧ 0.2 M NaCl. A screening of co-factors showed that CuCl2, ZnCl2, MnCl2 and Cu(II)SO4 reduced the enzyme activity of most GH29s, whereas MnCl2 increased the activity of LCC4_2_71. The activity of LCC4_2_293, LCC4_2_269, LCC4_2_230, LCC4_2_211 and LCC4_2_210 was unaffected, suggesting that the enzymes do not require any co-factors to be active.

Marine GH29s cleave fucose from macroalgal fucoidans

Enzymes involved in the utilization of fucoidans are characterized by a pronounced substrate specificity (Berteau et al. 2002). Therefore, we tested the activity of GH29s towards structurally different fucoidans from nine macroalgal species. Fucose release was observed by HPAEC-PAD

(Tab. 2). Most GH29s from '*Lenitmonas'* sp. CC4 were active on *F. vesiculosus* fucoidan, which is considered a homofucan mainly consisting of α-1,3/4-linked fucose (Nishino et al. 1994) and 30 – 50 % sulfate (Fitton et al. 2015; Sichert et al. 2020). While most GH29s are characterized by a pronounced substrate specificity, some enzymes act on several macroalgal species. For instance, LCC4_1_2701, LCC4_2_67 and LCC4_2_210 act on *F. vesiculosus* as well as *D. potatorum*, *U. pinnatifida* and *C. okamurans*, respectively. *C. okamurans* and *U. pinnatifida* are heteroglycans, characterized by a greater structural and compositional complexity than *F. vesiculosus*. *C. okamurans* is composed of α-1,3-fucose, which contains branches of α-1,2-fucose, acetylation as well as sulfate groups at O4 (Nagaoka et al. 1999; Lim et al. 2019). In contrast, *U. pinnatifida* is a galactofucan consisting of α-1,3-fucose and β-1,3/4/6-galactose (Lee et al. 2004; Hemmingson et al. 2006; Skriptsova et al. 2010; Synytsya et al. 2010; Zhao et al. 2018). Even though several GH29s can act on *U. pinnatifida,* the yield of fucosidase activity differs. For instance, LCC4_2_269, LCC4_2_103, LCC4_2_67, and LCC4_2_293 hydrolyze 1.8, 1.1, 0.98, and 0.52 µg fucose per mg fucoidan, respectively. As *U. pinnatifida* contains ~ 35.7 % fucose (Sichert et al. 2020), these enzymes are only capable of cleaving ≤ 0.5 % of the total fucose content, suggesting that their activity is limited to the amount of terminal fucose. Presumably, the yield of exo-acting GH29s increases in the presence of enzymatic partners such as sulfatases (Berteau et al. 2002).

Tab. 1. Activity screening against fucose-containing oligosaccharides. GH29s were incubated with oligosaccharides overnight and digestion products were analyzed with HPAEC-PAD. ND: not determined, -: indicates absence of activity.

Enzyme	α-1,2-fucose	α-1,3-fucose	α-1,4-fucose
LCC4_1_2701	BG-H DP2	-	-
LCC4_2_22	-	$Lewis_X$ DP3	-
LCC4_2_67	-	-	-
LCC4_2_71	BG-H DP2	-	-
LCC4_2_103	$Lewis_Y$ DP4		-
LCC4_2_161	ND	ND	ND
LCC4_2_169	-	3-fucosyllactose DP3	-
LCC4_2_171	BG-H DP2	3-fucosyllactose DP3	-
LCC4_2_210	2'-fucosyllactose DP3 Reduced XFG DP7	3-fucosyllactose DP3	-
LCC4_2_211	-	-	-
LCC4_2_218	ND	ND	ND
LCC4_2_230	Reduced XFG DP7	3-fucosyllactose DP3	-
LCC4_2_233	ND	ND	ND
LCC4_2_269	BG-H DP2 Reduced XFG DP7	-	-
LCC4_2_293	-	-	-

Tab. 2. Activity screening against macroalgal fucoidans. GH29s were incubated with fucoidans overnight and digestion products were analyzed with HPAEC-PAD. ND: not determined, +/- : indicates presence/absence of activity.

Enzyme	Fucales				Laminariales				Ectocarpales
	F. vesiculosus	*F. serratus*	*D. potatorum*	*S. fusiforme*	*E. maxima*	*U. pinnatifida*	*M. pyrifera*	*L. nigrescens*	*C. okamurans*
LCC4_1_2701	+	-	+	ND	-	-	ND	ND	-
LCC4_2_22	+	-	-	-	-	-	-	-	-
LCC4_2_67	+	ND	ND	ND	ND	+	ND	ND	-
LCC4_2_71	-	-	-	ND	-	-	ND	ND	-
LCC4_2_103	-	ND	ND	ND	ND	+	ND	ND	-
LCC4_2_161	ND	ND	ND	ND	ND	ND	ND	ND	ND
LCC4_2_169	-	-	-	ND	-	-	ND	ND	-
LCC4_2_171	+	-	-	ND	-	-	ND	ND	-
LCC4_2_210	+	-	-	ND	-	-	ND	ND	+
LCC4_2_211	+	-	-	ND	-	-	ND	ND	-
LCC4_2_218	ND	ND	ND	ND	ND	ND	ND	ND	ND
LCC4_2_230	+	-	-	ND	-	-	ND	ND	-
LCC4_2_233	ND	ND	ND	ND	ND	ND	ND	ND	ND
LCC4_2_269	-	ND	ND	ND	ND	+	ND	ND	-
LCC4_2_293	-	ND	ND	ND	ND	+	ND	ND	-

Crystal structures of GH29s

Of the 15 GH29s, 10 were expressed in suitable amounts to perform protein crystallization. For five out of eight crystallized GH29s diffraction data was collected. The structures for LCC4_2_22, LCC4_2_71, LCC4_2_169 and LCC4_1_2701 were solved by molecular replacement to 1.8, 2.1, 1.6 and 2.7, respectively (Fig. 2). In addition, mutants of LCC4_2_22 were soaked with the human blood oligosaccharides $Lewis_X$ and $Lewis_A$, resulting in an apo-structure with α-L-fucose in the active site (Fig. 3). All four structures possess a characteristic pocket-like active site for exo-acting enzymes (Davies et al. 1997; Hettle et al. 2018). The active site is formed by twelve conserved sterically residues. Based on the dimensions as well as the anionic electrostatic potential of the -1 subsite, these GH29s hydrolyze desulfated fucose from the glycan backbone.

LCC4_2_22, LCC4_2_169 and LCC4_1_2701 adopt a two domain-fold composing of a β-sandwich domain consisting of 7-8 antiparallel β-strands packed in two β-sheets at the C-terminus and the N-terminal $(\beta/\alpha)_{8/6}$ TIM-barrel domain containing the active site. In contrast, LCC4_2_71 is missing the C-terminal domain. Compared to other GH29 structures, 22_GH29 possess an additional loop at the very C-terminus consisting of 48 aa, which is stretching above the β-sandwich and facing towards the active site (Fig. S2A). Overall, the C-terminus is the most variable regions in GH29s (Fig. S2B) and often characterized by a putative CBM. About 51 % of the GH29s from '*Lentimonas'* sp. CC4 possess a C-terminal domain. Putative N- and/or C-terminal binding domains have been reported for GH29s (Summers et al. 2016; Dong et al. 2017), but their mechanistic motif remains unsolved.

Novel catalytic residue in marine GH29s

Previous studies reported the catalytical residues to be a highly conserved Asp and Glu as nucleophile and Bronsted acid/base (Sulzenbacher et al. 2004), respectively. In '*Lentimonas'* sp. CC4, 20 out of 35 GH29s are characterized by a pair of aspartic acids as putative catalytic residues (Fig. 2). For instance, 22_GH29 possess Asp259 which is sterically conserved at the same position as the catalytic Glu from other GH29 structures. This residue is located one flexible loop facing towards the active site. In 22_GH29, the catalytic residues are 6.2 Å apart (Fig. 3). To verify both catalytic residues, we constructed mutants of 22_GH29. Compared to the wild type, both mutants have between 20-30% residual activity (Fig. S3), strongly suggesting that 259Asp is indeed a new catalytic Bronsted acid/base. The DNA base triplet of the 22_GH29 gene encoding the Asp259 is GAC and might have evolved by one-point mutation from GAA or GAG encoding Glu to GAC encoding Asp. This is the first study reporting the Bronsted acid/base to be an Asp instead of a Glu.

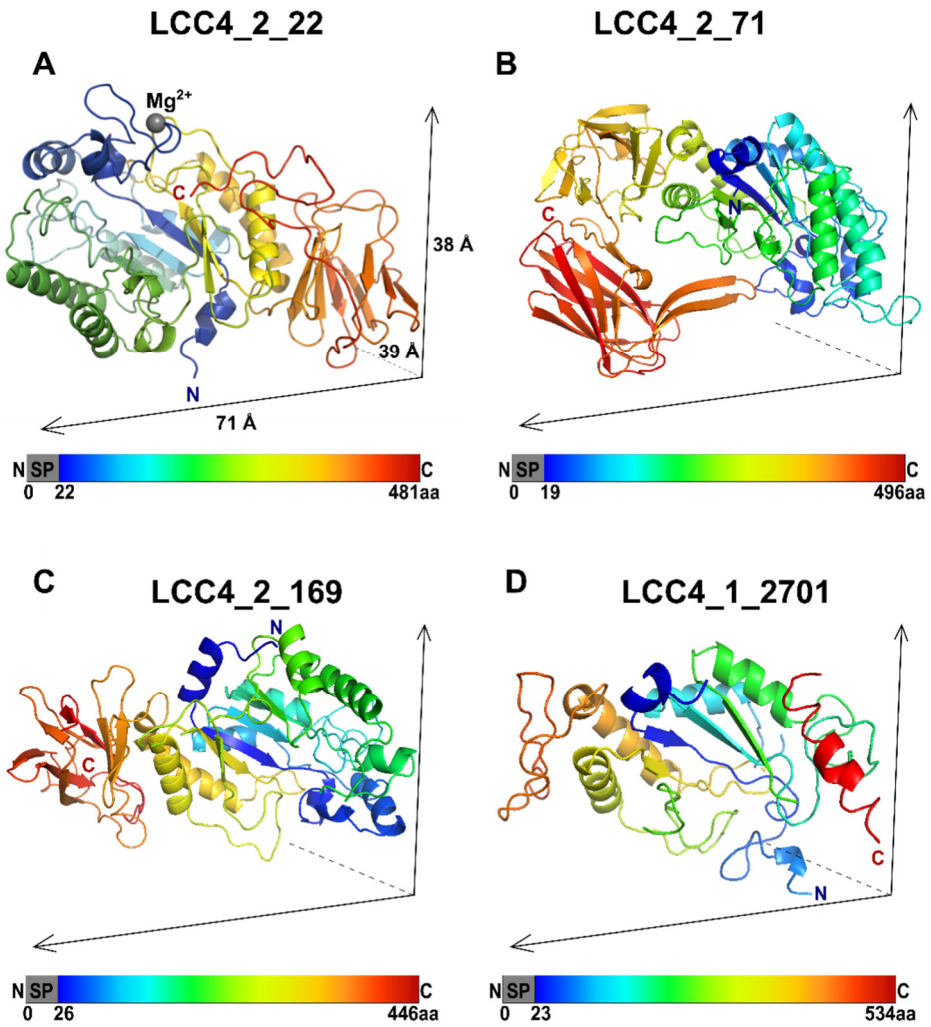


Fig. 2. Crystal structures of GH29s from '*Lentimonas*' sp. CC4. Secondary structures are shown in cartoon representation from N- (blue) to C- (red) terminus. Magnesium ion is shown as grey sphere.

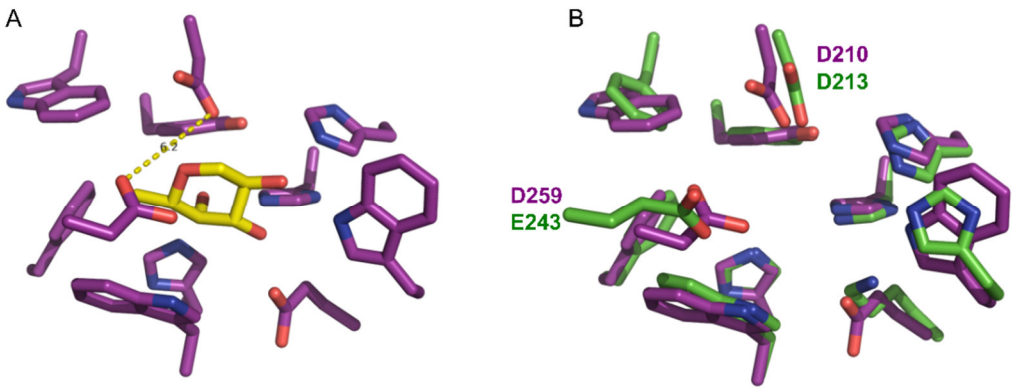


Fig. 3. Insights into active site of marine GH29s. (A) LCC4_2_22_GH29 in complex with α-L-fucose (yellow). (B) Superimposition of active site residues of solved GH29s structures LCC4_2_22 (purple) and LCC4_2_169 (green). Active site residues are shown as sticks. Catalytic residues are highlighted with sequence position.

Conclusion

Family GH29 is extensively studied investigating their role in the degradation of fucosylated glycans by the gut microbiota (Hobbs et al. 2019), whereas marine GH29s remain poorly studied. Nevertheless, together with the exo-acting family GH95, GH29s account to one of the most abundant CAZyme families in marine bacteria associated with algal degradation (Teeling et al. 2012; Kim et al. 2016; Sichert et al. 2020). '*Lenitmonas'* sp. CC4 encodes 35 GH29s, of which 15 were biochemically characterized in the present study. We identified striking differences among GH29s regarding their substrate and linkage specificity, suggesting that these enzymes are involved in several enzymatic pathways specialized on the source of fucoidan. '*Lenitmonas'* sp. CC4 possess GH29s, which exclusively hydrolyze α-1,2- or α-1,3-, whereas others are capable of cleaving both linkages as well as β-D-galactose. A similar mechanistic motif was also overserved for their putative neutral substrate. While some GH29s can only act one specific macroalgae species, others are capable of cleaving fucose from structurally different fucoidans. '*Lentimonas'* sp. CC4 is a valuable resource for the discovery of enzymes with a high biochemical and biotechnological novelty. In particular, the GH29s which are characterized by the pair of aspartic acids as novel catalytic residues represent interesting candidates with potential undescribed substrate and linkage specificity. This underpins that '*Lentimonas'* sp. CC4 can be further explored to get obtain new fucoidan active enzymes.

Experimental procedure

Materials

All chemicals were purchased from Sigma Aldrich (Germany) if not stated otherwise.

Production of recombinant plasmid and protein purification. Fucosidases of interested were cloned (Tab. S2), expressed and purified as previously described (Gerlach et al. 2020). In short, fucosidases were fused to a hexa polyhistidine tag at the N-terminus (Hochuli et al. 1988), neglecting the signal peptide (Nielsen 2017) and ligated into expression vector peT28a (69864, Merck Millipore). Protein expression was performed using *E. coli* BL21(DE3) (C2987I+C2527H, New England Biolabs) using either 1 mM IPTG induction in LB medium or autoinduction in ZYP medium (Studier 2005). *E. coli* cells were harvested by centrifugation and chemically lysed. Protein purification was done via immobilized affinity metal chromatography (IMAC) using 5 ml Talon column (GE Healthcare) operated at Äkta start (GE Healthcare).

Enzyme characterization.

Thermal stability

To ensure that all following enzymatic assays were performed under suitable temperature conditions, the thermal stability was investigated by dynamic light scattering (DLS). Therefore, triplicates of 30 µl purified protein were sealed with oil and measured from 25 – 80 °C with 0.1° min^{-1} increase in a DynaPro plate reader-II (Wyatt Technology).

State of oligomerization

Analytical size exclusion chromatography (SEC) was performed to confirm the theoretical molecular weight and to determine the state of oligomerization. ~ 100 µl IMAC purified protein was loaded onto a high-resolution column ENrich™ SEC 650 10 x 300 (Bio Rad) operated at NGC™ Chromatography System (Bio Rad) at 1 mL min^{-1} flow rate of 20 mM Tris pH 8, 250 mM NaCl. The absorbance at 280 nm of ~100 µL Gel Filtration Standard (1511901, Bio Rad) and protein sample was measured (Fig. S5).

pH, temperature, NaCl and cofactor screening

The biochemical characteristics of each fucosidase was determine photometrically at 405 nm in FLUOstar Omega microplate reader (BMG Labtech GmbH). 0.05 mg ml^{-1} enzyme was incubated for 15 or 30 minutes with 1 mM substrate at 30 °C. First, the effect of pH, followed by NaCl and divalent cations were determined in triplicates. The following chromogenic substrates were used: cNP-α-L-fucopyranoside (Carbosynth), pNP-β-L-fucopyranoside, pNP-β-D-fucopyranoside, pNP-β-D-galactopyranoside and pNP-α-D- galactopyranoside.

Substrate and linkage specificity via HPEAC-PAD

Substrate and linkage specificity were analyzed via high-performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD). The following substrates were tested: blood group H disaccharide (BGH, Carbosynth), blood group B tetraose type 5 (BGB, Elicityl), $Lewis_y$ tetraose (Le_y, Elicityl), $Lewis_b$ tetraose (Le_b, Elicityl), $Lewis_x$ triaose (Le_x, Elicityl), $Lewis_a$ triaose (Le_a, Carbosynth), 3-fucosyllactose (3-FL, Elicityl), reduced XFG xyloglucan (XFGol, Elicityl), 2‘-fucosyllactose (2’-FL, Carbosynth), as well as ion exchange chromatography-purified fucoidans (Sichert et al. 2020) from different macroalgal species: *Fucus serratus* (FS, Carbosynth), *Fucus vesiculosus* (FV, Carbosynth), *Undaria pinnatifida* (UP, Carbosynth), *Macrocystis pyrifera* (MP, Carbosynth), *Ecklonia maxima* (EM, Carbosynth), *Sargassum fusiforme* (SF, Carbosynth), *Cladosiphon okamurans* (CO, Carbosynth) *Lessonia nigrescens* (LN, Carbosynth), and *Durvillaea potatorum* (DP, Carbosynth).

1 mg ml^{-1} GH29 was incubated with 9 mg mL^{-1} substrate, 20 mM Na-cirtrate pH 5.7 1 mg ml^{-1} BSA, 0.2 M NaCl at 30 °C and 250 rpm and the reaction was terminated at 99 °C for 10 min. HPAEC-PAD was carried out on a Dionex ICS5000+ system (Dionex) using a CarboPac™ PA10 (2 mm x 250 mm, Dinoex) analytical column and a CarboPac™ PA10 (2 mm x 50 mm, Dinoex) guard column. Separation of neutral and acidic monomers was carried out using a previous described protocol (Engel and Händel 2011). In short, neutral monosaccharides were separated by a linear flow of 18 mM NaOH for 20 minutes, followed by a gradient of to 100 mM NaOH and 500 mM NaAc for 15 minutes to elute acidic sugars. To quantify the released monosaccharides, we used an external standard mixture composed of fucose, rhamnose, galctosamine, arabinoase, glucosamine, galactose, glucose, mannose, xylose, muramic acid, galacturonic acid, glucuronic acid and mannuronic acid.

Identification and mutagenesis of catalytic residues from 22_GH29

The catalytic residues of family GH29 are highly conserved as Asp as nucleophile and Glu/Asp as Bronsted acid/base (Sulzenbacher et al. 2004) arranged in the TIM barrel domain of the enzyme. The nucleophile was identified via multiple sequence alignment using ESPript V3.0 (Robert and Gouet 2014), whereas the Bronsted acid/base can only be found via structural alignment, which was performed in PyMOL Version 2.0 (Schrödinger, LLC).

The putative catalytic residues were identified via structural alignment as Asp210 and Asp259 with 2wvs. Single-side directed mutagenesis of each catalytic residue into a Ala residue was performed according to Liu & Naismith (2008) using purified recombinant plasmid of 22_GH29 (Tab. S1). Mutated recombinant plasmids were verified by gel electrophoresis and Sanger-sequencing prior transformation into *E. coli* BL21(DE3).

Protein crystallization and X-ray data collection

Fucosidases were screen for suitable crystallization conditions using a crystallization robot (Phoenix, ARI) and the sitting-drop vapor diffusion approach. Conditions that showed crystallized protein were repeated and, optimized with slight modifications to the original recipe using hanging-drop vapor diffusion at 16 °C. 24-well hanging drop crystal trials were set up containing 500 µL mother liquor and three different dilutions (1:2, 2:1 and 1:1) of protein and crystallization solution with a total volume of 3 µL.

Protein crystals were soaked in 25-30 % glycerol in the mother liquor and flash-frozen to cryo-temperatures using liquid nitrogen. 22_GH29 was soaked with substrate prior freezing. Data collection was done at the PETRAIII (Burkhardt et al. 2016) at the beam line P13 (Cianci et al. 2017) from European Molecular Biology Laboratory (EMBL, Hamburg), beam line P11 (Meents

et al. 2013) from Deutsches Elektronen Synchrotron (DESY, Hamburg) and beam line MX from the Diamond Light Source.

Structure determination

Structures were solved by molecular replacement using the following software packages in CCP4i2 (Winn et al. 2011) and CCP4 Cloud (Krissinel et al. 2018): iMOSFLM (Battye et al. 2011), DIALS (Winter et al. 2018), AIMLESS (Evans 2011), MrBUMP (Keegan and Winn 2007), and PHASER (McCoy et al. 2007). The structural homologues were identified by protein blast in NCBI using PDB data base (Berman 2000). The models were built using AUTOBUILD and REFINE in the Phenix crystallography suite (Adams et al. 2010) as well as BUCCANER (Cowtan 2006), PARROT (Cowtan 2010), and CCP4build (Winn et al. 2011). Refinement was done based on a Maximum Likelihood method using REFMAC V5.5 (Murshudov et al. 1997) and COOT V0.8.7.1 (Emsley et al. 2010) for manual finalization. Data processing and statistics are shown in Table S3. The degree of conserved residues was determined using ConSurf (Glaser et al. 2003) and DALI (Holm and Elofsson 2019).

Data availability

Refined crystal structure coordinates and experimental structure factors are available at the Protein Data Bank.

Author's contribution

NG, AS and JHH designed the study. NG, AS and TvR cloned and expressed the enzymes. NG and AP performed biochemical analyses. NG and TvR crystallized the enzymes. NG and CSR collected X-ray diffraction data. NG, RK and CSR determined the protein structures. AS conducted phylogenetic analysis. NG designed mutants. NG and AS visualized the data. NG, AS and JHH wrote the manuscript with contribution of all authors.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements and funding

We acknowledge EMBL, a member of the Helmholtz Association HGF, and DESY (Hamburg, Germany) for the provision of experimental facilities. Parts of this research were carried out at PETRAIII and we would like to thank the beamline scientist in operating the beamline P13 and P11. Part of the structural studies were also performed on beamline I03 at the Diamond Light Source, Harwell, UK. We are grateful to Dr. Neil Paterson and Dr. David Aragao assistance while using this beamline. We would like to thank the 2nd CCP4/BGU Workshop on Advanced Methods for Macromolecular Structure Determination which took place at Ben-Gurion University during February - March 2020 help and contribution in the structure determination. We also thank Alek Bolte for assistance with HPEAC-PAD as well as Tina Trautmann and Mandy Kunze for lab assistance. J.H.H. acknowledges funding from the Max-Planck-Gesellschaft and from the Emmy Noether Program of the DFG (grant number HE 7217/1-1).

Abbreviations

DP: degree of polymerization, cNP: *2*-Chloro-4-nitrophenyl, pNP: 4-nitrophenol, DLS: dynamic light scattering, IMAC: immobilized meal affinity chromatography, SEC: size exclusion chromatography, HPAEC-PAD: high performance anion exchange chromatography pulsed amperometric detection

Supplementary information

In appendix.

Discussion and future perspectives

The overall goal of this dissertation was to gain molecular insights into the utilization of anionic polysaccharides by marine heterotrophic bacteria. This work combines biochemical characterization with structural biology to shed light on the mechanistic motif of ulvan- and fucoidan-degrading enzymes.

X-ray crystallography as a powerful tool for glycobiochemistry

Structural biology explores the structure of macromolecules such as proteins or nucleic acids. This research field requires the combination of molecular biology, biochemistry and biophysics to enhance our understanding of the function and interaction of molecules. In the last three decades, technical innovations have led to an explosion of structure determinations in the PDB (PDB Statistics). X-ray crystallography is the most commonly used approach to study the 3D structure of proteins and to gain insights into their function on a molecular level (PDB Statistics). The fundamental step of this technique is at the same time often the bottleneck – generating highly diffracting protein crystals. Protein crystallization is a complex, multiparametric process, which can be very challenging and time-consuming (Smyth 2000; Chayen and Saridakis 2008). Progress in experimental procedures range from protein purification (Magnusdottir et al. 2009) to the usage of novel phasing and nucleation agents (Engilberge et al. 2017; Engilberge et al. 2018; Engilberge et al. 2019), streaking and seeding techniques (Stura and Wilson 1991; Bergfors 2003; Stewart et al. 2011) and *in vivo* crystallization (Koopmann et al. 2012; Gati et al. 2014).

In chapter 1, the enzymatic pathway to degrade the cell wall polysaccharide ulvan of green seaweed has been elucidated. The collaboration of biochemistry, proteomic and transcriptomic approaches in combination with structural biology revealed detailed and novel insights into ulvan utilization. The enzymatic machinery from *Formosa agariphila* KMM3901^T consists of seven glycoside hydrolases, three sulfatases, and two polysaccharide lyases that sequentially break down ulvan into fermentable monosaccharides (Reisky 2019). The structures of three of the 12 enzymes was solved including an endo-acting GH39 with a novel activity towards α-1,4-L-rhamnose (Fig. 8), an exo-acting β-glucuronidase of family GH2, and an arylsulfatase of subfamily S1_8, which removes the sulfate group of an ulvan disaccharide. However, the structure of key enzymes from the pathway still remains unsolved.

I focused on the two ulvan lyases, that initiate the hydrolytic process (Reisky et al. 2019). Even though both lyases crystallized and diffraction data was collected, it was not possible to solve the structures. A combination of several factors impeded structure determination. (i) Crystallization

conditions: In order to find suitable crystallization conditions, multiple protein concentrations are screened against a large variety of crystallization solutions, resulting in hundreds of possible combinations (Chayen and Saridakis 2008). Both enzymes did crystallize in a variety of conditions, but only two to three yielded moderate diffraction after several, time-consuming optimization steps (Fig. S1), while the quality of the diffraction data remained moderate. Poorly crystallizing or flexible regions can be removed by protein engineering, i.e. truncation of the enzyme, to improve protein expression and/or diffraction (Collins et al. 2018). For ulvan lyase P30_PL28, the full construct without the N-terminal signal peptide resulted into weak expression, whereas the truncated version also missing the C-terminal signal peptide for the type IX secretion system (T9SS) showed high expression (Reisky et al. 2018). To obtain better diffracting crystals, PL28 was further truncated to its catalytic domain only. Removing its C-terminal binding domain resulted in weak expression that was not suitable for protein crystallization. In addition, I attempted different techniques (i.e. hanging drop and sitting drop vapor diffusion vs. microbatch method), (streak) seeding, co-crystallization with ulvan as well as additives and nucleation agents to improve the crystallization process and consequently, the resolution. Finally, I was able to collect medium resolution data sets at 2.2 – 3 Å for both enzymes. (ii) Structure determination: The 3D structure of enzymes can be solved using different approaches. If the structure of a distant relative is known, molecular replacement is the most suitable technique. As a rule of thumb, a successful structural homologue covers ~ 50 % of the target enzymes, which should diffract to ≤ 2 Å. The amino acid sequence identity between both enzymes determines the lower limit for molecular replacement and should be > 25-30 % (Chothial and Lesk 1986; Abergel 2013; Brito and Archer 2020). However, sequence identity is only an indication. Structures have been solved via molecular replacement with sequence identity of less than 20 % (Simpkin et al. 2019), whereas in other cases much higher values remained much more challenging. P30_PL28 shares 66 % sequence identity with its closest structural homologue, whereas there is no such homologue for P10_PLnc. Therefore, the second approach of experimental phasing was pursued. I performed selenomethionine incorporation, co-crystallization and soaking with heavy metals and lanthanides, but regardless of the technique used to determine the structure, the (iii) data quality is the decisive factor. The quality of a crystal is impacted during the freezing process, throughout data collection in form of radiation damage (Pflugrath 2015) and by its crystallization conditions. All of this results in a variety of challenging protein specimen regardless of the resolution. Such factors include intrinsic disorder, data pathologies (e.g. anisotropy, twinning, translational non-crystallographic symmetry (tNCS)), high non-crystallographic symmetry (NCS), alerted confirmations of target and model protein, etc. In conclusion, the challenges associated with both lyases impaired determination of their structures and reflect challenging cases for macromolecular crystallography.

Another riveting case is the endo-acting α-1,4-L-rhamnosidase. P31_GH39 crystallized in a variety of conditions, but the obtained crystals were clustered and too small for macromolecular crystallography, regardless of the protein concentration or the ratio of protein and crystallization condition. Only one condition resulted into big, hexagonal crystals (Fig. S1) diffracting to ~ 2 Å. However, the growth of these crystals took roughly one year, demonstrating the uncertainty of success in protein crystallization. Eventually, the structure of the P31_GH39 was solved. This is the first structure of a GH39 with an endo-acting α-1,4-L-rhamnosidase activity. So far, family GH39 has only been described as β-xylosidases (EC 3.2.1.37), β-galactosidases (EC 3.2.1.23), β-glucosidases (EC 3.2.1.21), α-L-arabinofuranosidase (EC 3.2.1.55) and α-L-iduronidase (EC 3.2.1.76) (Lombard et al. 2014a).

P31_GH39 consists of a β-sandwich domain at the C-terminus and a N-terminal TIM-barrel domain harboring the active site (Fig. 8A). Comparable to the other ulvan-utilizing enzymes from *F. agariphila* (Reisky et al. 2019), P31_GH39 is characterized by a cationic active site (Fig. 8B+D), reflecting the interaction of the enzymes with anionic substrates. The overall fold of GH39s can be also found in other CAZymes such as fucosidases of family GH29 (Lombard et al. 2014a). In both enzyme families, the acid/base of the catalytic pair can be found by multisequence alignment, whereas the nucleophile can only be found by structural alignment, again underlining the importance of structural biology. Families GH29 and GH39 are retaining enzymes, whose catalytic residues are typically ~ 5.5 Å apart (McCarter and Stephen Withers 1994). In P31_GH39, this distance is only 3.8 Å (Fig. 8C). For a detailed understanding of the endo-acting α-L-rhamnosidase activity of P31_GH39, enzyme-substrate complexes are needed. Holo-structures do not only clarify the catalytic mechanism, but also resolve the mechanistic motifs of the C-terminally β-sandwich domain. Bioinformatic analyses identified this domain as putative carbohydrate-binding module (CBM) of families CBM32 and CBM35 as well as CBM9, CBM13 and CBM35 for families GH29 (Abbott et al. 2007; Dong et al. 2017) and GH39 (Yu et al. 2015), respectively. Many GHs are multimodular enzymes, which often possess surface binding sites (SBS) located close to the catalytic domain with a certain distance to the active site or CBMs, which are auxiliary binding sites in addition to the catalytic domain. For many CAZyme families, such putative carbohydrate-binding domains are reported to be involved in substrate recognition or binding, but they are not well studied yet (Cuyvers et al. 2012; Ficko-Blean and Boraston 2012; Cockburn et al. 2014). To date, no biochemical or structural study elucidated the mechanistic motif of putative CBMs for neither family GH29 nor GH39. Presumably, this domain is involved in substrate recognition or binding (Yu et al. 2015).

Even though the tertiary structure of families GH29 and GH39 has already been described (Yang et al. 2004; Ramos Santos et al. 2012; Maita et al. 2013; Yu et al. 2015; Ali-Ahmad et al. 2017; Jones et al. 2017), each additional structure points out structural differences and enhances our

knowledge and thus improves the accuracy and reliability of other tools predicting substrate binding and protein-protein interaction (Laskowski et al. 2018; Holm and Elofsson 2019) or building models (Guex et al. 2009; Kelley et al. 2015).

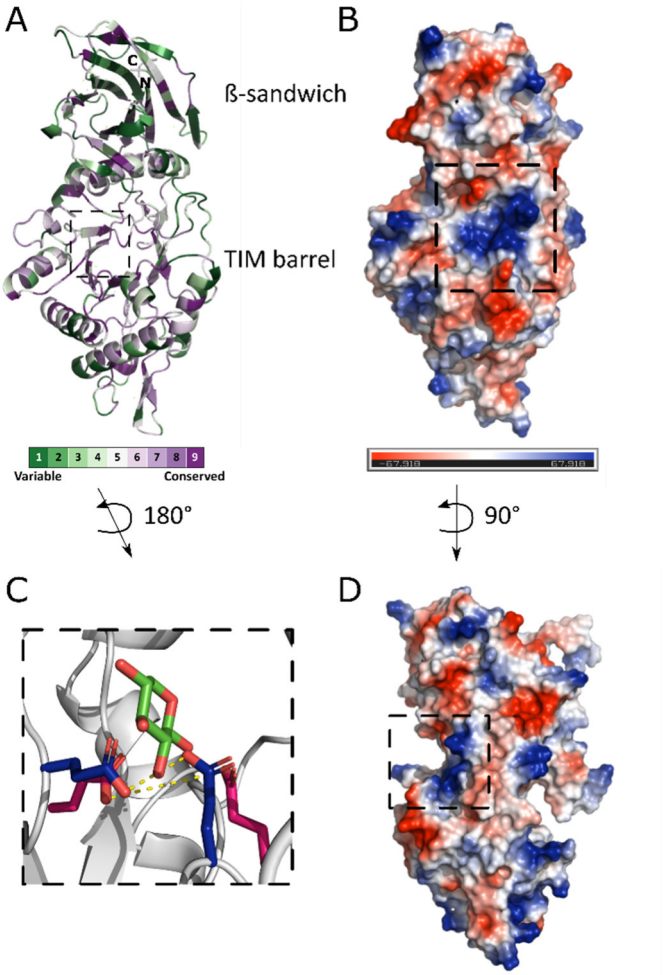


Figure 8. Structure of 31_GH39 with novel endo α-1,4-L-rhamnosidase activity. (A) Cartoon representation of 31_GH39. Degree of conservation is shown as colour score. (B, D) Surface view identifies highly cationic (blue) active site. (C) Catalytic residues Glu199 and Glu307 (blue) identified by structural alignment with GH39 β-D-xylosidase (PDB:1PX8, magenta) from *T. saccharolyticum* B6A-RI in complex with β-D-xylopyranose (green).

Unpuzzling fucoidan degradation by '*Lentimonas*' sp. CC4

'Lentimonas' sp. CC4 is an extraordinary glycan degrader, specialized on the sulfated algal polysaccharides carrageenan and fucoidan. Fucoidan utilization by this isolate is almost exclusively coordinated on a 0.89 mbp plasmid harboring 113 sulfatases, 100 glycoside hydrolases and 17 carbohydrate esterases. These putative fucoidan-degrading enzymes are individually upregulated

depending on the type and source of fucoidan. The regulation of these enzymes in several operons as well as its ability to utilize several fucoidans at different rates (Sichert et al. 2020) are strong indicators for the existence of several fucoidan-utilizing pathways. The complete enzymatic pathway to catabolize fucoidan has not been elucidated for any organisms yet. In order to unpuzzle fucoidan utilization by '*Lentimonas'* sp. CC4, I propose three possible strategies for future studies.

(i) In-depth characterization of specific CAZyme families: The genomes of the fucoidan-degrading bacteria '*Lentimonas'* sp. CC4 and *Kritimatiellales* sp. F21 are highly enriched in the families GH29, GH95, GH141, S1_15, S1_16, S1_17, S1_25 and CE7, suggesting that these families are crucial for the degradation process (van Vliet et al. 2019; Sichert et al. 2020). Thus, focusing on these families is a valuable starting point to elucidate fucoidan degradation.

The present work focused on exo-acting enzymes, which have been shown to initiate the degradation of fucosylated glycans in the human gut microbiome (Hobbs et al. 2019). '*Lentimonas'* sp. CC4 possess 50 homologues of fucosidases of the families GH29 and GH95 alone (Sichert et al. 2020). In chapter 2, the substrate specificity of 15 GH29s and one GH95 towards nine macroalgal fucoidans was determined. Based on the biochemical characterization, I hypothesize that GH95s mainly cleave α-1,2-linked fucose, whereas GH29s target α-1,3 and α-1,4 linkages in fucoidans. However, their ability to act on fucoidan is limited. Single exo-acting enzymes were only capable of removing less than one percent of the total ~ 20 – 50 % fucose content available in macroalgal fucoidans (Fitton et al. 2015; Sichert et al. 2020). Exo-acting GH29s and GH95s can only be active on accessible fucose either located at the terminus of the glycan backbone or attached to it as branches. Decorated fucose with regards to sulfation or acetylation is not accessible for these enzymes and requires enzymatic partners such as esterases or sulfatases (Berteau et al. 2002) to act first. Fucoidan-degraders such as *Rhodopirellula baltica* SWK7, *Lentisphaera araneosa* HTCC2155T and *Kiritimatiellales* F1 encode between 200 and more than 500 sulfatases (Thrash et al. 2010; Wegner et al. 2013; van Vliet et al. 2019). Previous studies showed, that S1_17 cleaves Fuc3S, whereas family S1_25 acts on Fuc2S (Silchenko et al. 2018). Subfamilies S1_16 and S1_15 might be specialized on other linkages and thus, increase the overall yield of desulfation similar to fucosidases of families GH29 and GH95. The substrate and/or linkage specificity of GHs and sulfatases is a result of the active site architecture (Davies et al. 1997; Hettle et al. 2018). For example, the human blood oligosaccharides $Lewis_A$ and $Lewis_X$ both consist of fucose, glucose and galactose. Therefore, they are widely used as substrates to test activity of fucosidases and galactosidases. In $Lewis_X$, fucose is linked via an α-1,3-linkage to glucose, whereas $Lewis_A$ exhibits an α-1,4-linkage at this position. Several studies reported linkage specificities not only between the families GH29 and GH95, but also within these families (Katayama et al. 2004; Ashida et al. 2009; Hobbs et al. 2019).

Bacteria encoding several families with the same activity (e.g. fucosidases) plus having many homologues of the same family with complementary linkage specificities enables organisms to exploit a wide range of glycans (Hobbs et al. 2019). As fucoidans form a versatile glass of glycans, a similar scenario is possible for fucoidan-degraders targeting structurally different FCSPs. Not many bacteria possess the enzymatic repertoire to utilize fucoidans. It appears that these bacteria adapted to a complex ecological niche presumably to limit or even avoid interspecies competition for carbon sources. However, enzymatic pathways to utilize complex glycans is cost intensive. For instance, Bacteriodetes and Proteobacteria specialized on sulfated mannan-, xylose- or fucose-containing polysaccharides possess PULs which harbor more enzyme families and homologue of the same families compared to simpler glycans such α- and β-glucans or alginate. Consequently, this results into larger genome sizes and higher energy demands (Krüger et al. 2019; Kappelmann et al. 2019). Such substrate-specialized niche partitioning is not only foung among species but also within species such as *Alteromonas* (Ivars-Martínez et al. 2008; Ivars-Martinez et al. 2008), *Polaribacter* (Avcı et al. 2020) and *Vibrio* (Hehemann et al. 2016; Corzett et al. 2018).

Eventually, the combination of several exo-acting enzymes such as esterases, fucosidases and sulfatases leads to the de-branching and un-decoration of fucoidans. Thereby, the unbranched polymer is accessible for other enzymes such as endo-acting enzymes of family GH107 producing fucoidan-oligosaccharides.

(ii) Identification of endo-fucosidases: Focusing on endo-acting enzymes that initiate the glycan degradation might serve an alternative to the exo-strategy. For example, in *Formosa agariphila* KMM3901^T, the catabolism of the anionic polysaccharide ulvan is encoded in a PUL of 39 enzymes (Mann et al. 2013), of which 12 CAZymes and sulfatases sequentially degrade the glycan into fermentable monosaccharides (Reisky 2019). In spite of branches and sulfation, the utilization process is initiated by three endo-acting CAZymes, including two polysaccharide lyases and a glycoside hydrolase of family GH39. These enzymes generate similar, yet structurally different oligosaccharides, which are the starting point of an enzymatic cascade of substrate specific endo- and exo-acting proteins (Reisky et al. 2019). A similar enzyme cascade may be possible for fucoidan catabolism.

'*Lentimonas'* sp. CC4 possess four GH107s, which are co-located (Fig. 8A) together with other fucoidan-degrading CAZyme families and sulfatases (Sichert et al. 2020). Via heterologous gene expression, I produced these four GH107s together with four hypothetical enzymes sharing certain sequence similarities. I investigated various expression system (IPTG vs. auto-induction and codon optimized vs. non-optimized enzymes in the expression strains *E.coli* BL21, Rosetta, Shuffle and Lemo21) and found insoluble expression in form of inclusion bodies. For none of the eight putative endo-fucoidanases any of the tested expression conditions yielded significant amounts of active

enzyme. Similar observations were also reported for GH107 FFA2 from *F. algae* (Silchenko et al. 2016). These results might be explained by the multimodular domain architecture of family GH107. Family GH107 consists of an N-terminal secretion signal, the catalytic domain and different immunoglobulin (Ig)-like domains. Some members have additional C-terminal PorSS signals and polycystic kidney disease (PKD) domains (Vickers et al. 2018; Schultz-Johansen et al. 2018). So far, this complex domain architecture has been barely investigated. Some studies report that heterologous gene expression was successful or enhanced by truncated versions without other domains than the catalytic domain (Vickers et al. 2018; Schultz-Johansen et al. 2018). Hence, future studies focusing on endo-fucoidanases from '*Lentimonas'* sp. CC4 should start by investigating the impact of different enzyme boundaries on protein expression first.

Alternatively, native purification or gene knockout might be successful ways to investigate *in vivo*, which enzymes are crucial to initiate fucoidan degradation. Native purification from marine bacteria and invertebrates identified several endo-fucoidanases (Silchenko et al. 2013; Silchenko et al. 2014; Kim et al. 2015) as well as a fucoidan-esterase (Nagao et al. 2017), whereas single and double mutants of *Zobellia galactanivorans* helped to reveal the complete utilization system of carrageenan (Ficko-Blean et al. 2017), glycosaminoglycans (Ndeh et al. 2020) and RG-II/pectin (Ndeh et al. 2017).

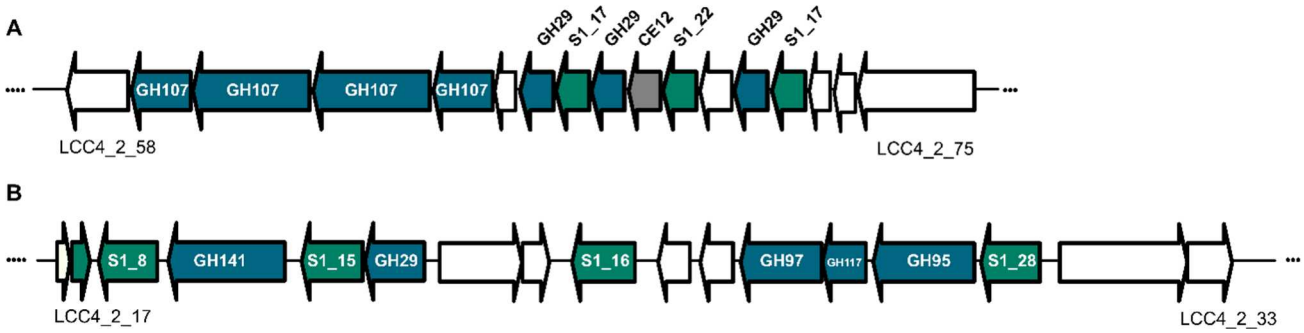


Figure 9. Fucoidan-utilizing gene clusters from '*Lentimonas'* sp. CC4. (A) Putative endo-acting fucoidanases of family GH107. (B) Two fucoidan-operons of chapter 2.1 (left) and 2.2 (right). Glycoside hydrolases are shown in blue, sulfatases in green and carbohydrate esterases in grey. Operons characterized by (A) endo-acting and (B) exo-acting fucoidanases.

(iii) Focus on fucoidan operons: The structural complexity of fucoidans and the theoretically high number of putative fucoidan-degrading enzymes is at the same time fascinating and exciting but also challenging and overwhelming. If I could start the fucoidan project all over again, but with the knowledge about '*Lentimonas'* sp. CC4 that we have now, I suggest focusing on single operons. In my opinion, the most promising approach to unravel fucoidan degradation by this isolate is by choosing one source of fucoidan, which upregulates CAZymes and sulfatases in

an operon-like cluster. Even if these synergistic enzymes would not lead to the complete degradation, they might produce defined oligosaccharides that could be further used in biochemical and structural experiments to investigate the mechanistic motif of other enzymes. The two operons harboring 21_S1_15 and 22_GH29 as well as 30_GH95 from chapters 2.1. and 2.2 appear as promising candidates (Fig. 9B). These regions stand out by encoding putative exo-acting fucoidan-degrading enzymes of various CAZyme and sulfatase families. In chapter 1, fucosidase activity was observed for 22_GH29 and 30_GH95 with a pronounced substrate specificity for macroalgal species. Promising candidates in these operons include families S1_8, GH97, GH117 and GH141. To date, all four enzyme families contain a high number of identified members of which only few have been biochemically and structurally characterized. For instance, families GH97 and GH141 have 2033 and 424 entries in the CAZy database of which only eight and two enzymes have been characterized, respectively. Families GH97 and GH117 are α-galactosidases and α-glucosidases (Lombard et al. 2014a), and galactose is an abundant monosaccharide in certain fucoidans (Tab. 1). For instance, fucoidans from *S. fusiforme* contains α-1,3/4/6-linked galactose (Hu et al. 2014; Hu et al. 2016), whereas *U. pinnatifida* possess β-1,3/4/6- linked galactose (Lee et al. 2004). Fucoidan-degrading bacteria like '*Lentimonas'* sp. CC4 possess a high number of enzymes from family GH141 (Sichert et al. 2020), which are endo-acting xylanases (Heinze et al. 2017) and α-L-fucosidases with an unknown mechanism (Ndeh et al. 2017). Endo-acting galactosidases or xylanases such as family GH141 are alternative candidates to endo-acting family GH107, which might act on galacto- and xylofucans found in *U. pinnatifida* (Lee et al. 2004; Hemmingson et al. 2006; Skriptsova et al. 2010; Synytsya et al. 2010; Zhao et al. 2018) and *P. plantagenia* (Bilan et al. 2014). P32_S1_8 from *Formosa agariphila* KMM3901^T is an exo-acting xylose 2-sulfate sulfatase (Reisky et al. 2019). Even though sulfated xylose has not been reported for fucoidans yet, subfamily S1_8 remains an interesting candidate.

To conclude, all the presented strategies have potential to resolve fucoidan catabolism in '*Lentimonas'* sp. CC4. The most promising approach might be a research team working on multiple strategies to generate an enzyme cocktail that can degrade fucoidan into fermentable monosaccharides.

Classification of glycoside hydrolase subfamilies

Fucoidan-degrading bacteria enriched in specialized CAZyme and sulfatase families are not only a clear indication that these enzymes play an important role in fucoidan degradation, but also raise the question why so many homologues of the same families are needed. The structural diversity of fucoidans and the substrate specificities of known fucoidan-degrading enzymes suggest that bacteria perform an individual multistep pathway adapted to the type and source of fucoidan (Sichert et al. 2020).

In chapter 2, exo-acting fucosidases of families GH29 and GH95 were biochemically and structurally characterized to gain insights into fucoidan degradation by *Lentimonas* sp. CC4. Even though both families have been proposed as important CAZymes involved in the utilization of macroalgal polysaccharides by heterotrophic bacteria (Teeling et al. 2012; Kappelmann et al. 2019), I demonstrated the fucoidan-degrading capacity of family GH95 for the first time. Moreover, the comparative approach of 15 GH29s provides advanced knowledge on substrate and linkage specificity for marine members. The data adds to the knowledge that the diversity of CAZyme and sulfatase families cannot only be explained by their sequence identity (Lombard et al. 2014a; Barbeyron et al. 2016a), but also reflects other aspects of their function and role in their respective environment. For instance, alginate lyases of family PL7 are divided into six subfamilies, which partially reflect their substrate specificity as well as the habitat they originate from (Thomas et al. 2013). In contrast, sulfatases have been only classified into subfamilies of S1 based on substrate specificity (Barbeyron et al. 2016a). GHs have been extensively studied and further grouped into clans, reflecting the tertiary structure (Lombard et al. 2014a), or into subfamilies based on sequence ident and similar substrate specificities. However, the number of families which have been studied in detail is low when compared to the total number of families and members known. Out of the 163 GH families known to date, only the families GH5, GH13, GH16, GH30, GH43 and GH128 are further classified into subfamilies (Andrews et al. 2000; Stam et al. 2006; St John et al. 2010; Aspeborg et al. 2012; Mewis et al. 2016; Viborg et al. 2019).

To date, fucoidan-utilizing CAZymes of families GH29 and GH95 have not been classified into subfamilies (Lombard et al. 2014a). Previous studies proposed two subfamilies for family GH29 based on substrate specificity and phylogenetic clustering. The proposed subfamily A is active on 4-nitrophenyl-α-L-fucopyranoside (pNP-α-L-Fuc), whereas the proposed subfamily B hydrolyzes the α-L-1,3/4-fucosidic bond but not pNP (Ashida et al. 2009; Sakurama et al. 2012a; Sakurama et al. 2012b). This classification however has not been implemented in the CAZy database yet (Lombard et al. 2014a). The biochemical and structural characterization of marine GH29s in chapter 2.3. may provide a first step for the classification of GH29s.

In contrast to family GH29, GH95s even are less studied. To date, there are only ten entries in the CAZy data base of which two also contain structural details. Family GH95 is characterized by a similar substrate specificity to families GH29 or GH27 but adopts an overall fold as a $(\alpha/\alpha)_6$-helical barrel fold shared with those others of the GH-L clan (GH15, GH65 and GH125), and the β-linkage-cleaving mode of GH94. The active site is highly conserved, consisting of glutamic acid (E) and asparagine (N) as acid / base pair (Nagae et al. 2007). So far, GH95s were either described as α-L-fucosidases or -galactosidases (Rogowski et al. 2015; Hobbs et al. 2019). In chapter 2.2., I showed novel biochemical and structural features of 30_GH95 which have not been reported for this family yet. 30_GH95 possess a dual activity towards 1,2/3-linked fucose and β-D-galactose. It has been proposed that the so-called *H/T* polymorphism of the active site determines the substrate specificity of GH95s (Rogowski et al. 2015). 30_GH95 is the first described member of the family who deviates from the highly conserved active site architecture: L-W-*H/T*-**N**-N-E-R-H-W-H-**D**. It is characterized by a *histidine* besides its dual activity and the first **aspartic acid** residue, which orientates the side chain of the catalytic acid (Nagae et al. 2007), is replaced by an **asparagine**. Based on these two novel structural features, I extended the sequence analysis of GH95s including all homologues from '*Lentimonas'* sp. CC4. The preliminary analysis of 85 sequences resulted in a revised active site residue architecture: X-X-*X*-**N/D**-N-D/E-R-H-Y/W-X-**D/E**, of which X represents amino acids with more than two alternative residues. I propose that the orientation of the side chain of the acid and the activation of the base are not exclusively carried out by a **D** and **N**, respectively. In 6 and 54 % of the sequences, these residues are exchanged by **E** and **D**, respectively. The catalytic pair of N and E are highly conserved among GH95s. Only one GH95 from the agar-degrading, marine Gammaproteobacterium *Gayadomonas joobiniege* possess an aspartic acid at the position of the catalytic acid. The amino acid which determines the previously proposed *H/T* polymorphism is, together with the second histidine, the most versatile residue in the active site. In about 80 % of both positions a histidine is found, followed by ~ 10 % threonine and glycine, respectively, as well as other polar and hydrophobic amino acids (W, M, G, L and R, S, Y, T).

The preliminary phylogenetic analyses revealed that GH95s fall in several clades and can be further distinguished into two subgroups (Fig. 10). Subgroup A contains only marine bacteria from the superphylum PVC and the phylum Bacteroidetes which possess an extended loop linking the $(\alpha/\alpha)_6$-helical barrel harboring the active site with the C-terminal β-sandwich domain. In addition, they are characterized by a histidine for the *H/T* polymorphism and the residue which orientates the side chain of the catalytic acid is an **D**. In contrast, subgroup B possess a H, T or other another residue for the *H/T* polymorphism and is characterized by an **N** instead of a D. GH95s belonging to subgroup B originate from marine and terrestrial bacteria, fungi and eukaryote and are missing the linker. GH95s from '*Lentimonas'* sp. CC4 belonging to clades 11, 12 and 13 are interesting cases as they are the only members which are not characterized by a H or T but M, W, L and G instead.

This structural feature together with our biochemical analyses of 30_GH95 in chapter 2.2. suggest that the *H/T* polymorphism does not determine the substrate specificity but is presumably involved in substrate binding and/or recognition. Moreover, clade 13 is the only clade which possess an **E** instead of a D, which activates the catalytic base, and both tryptophans are replaced by Y, A or F.

In conclusion, the in-depth characterization of families GH29 and GH95 is the basis to divide these families into subfamilies. Investigating special cases such as clade 13 from family GH95 might reveal novel functions.

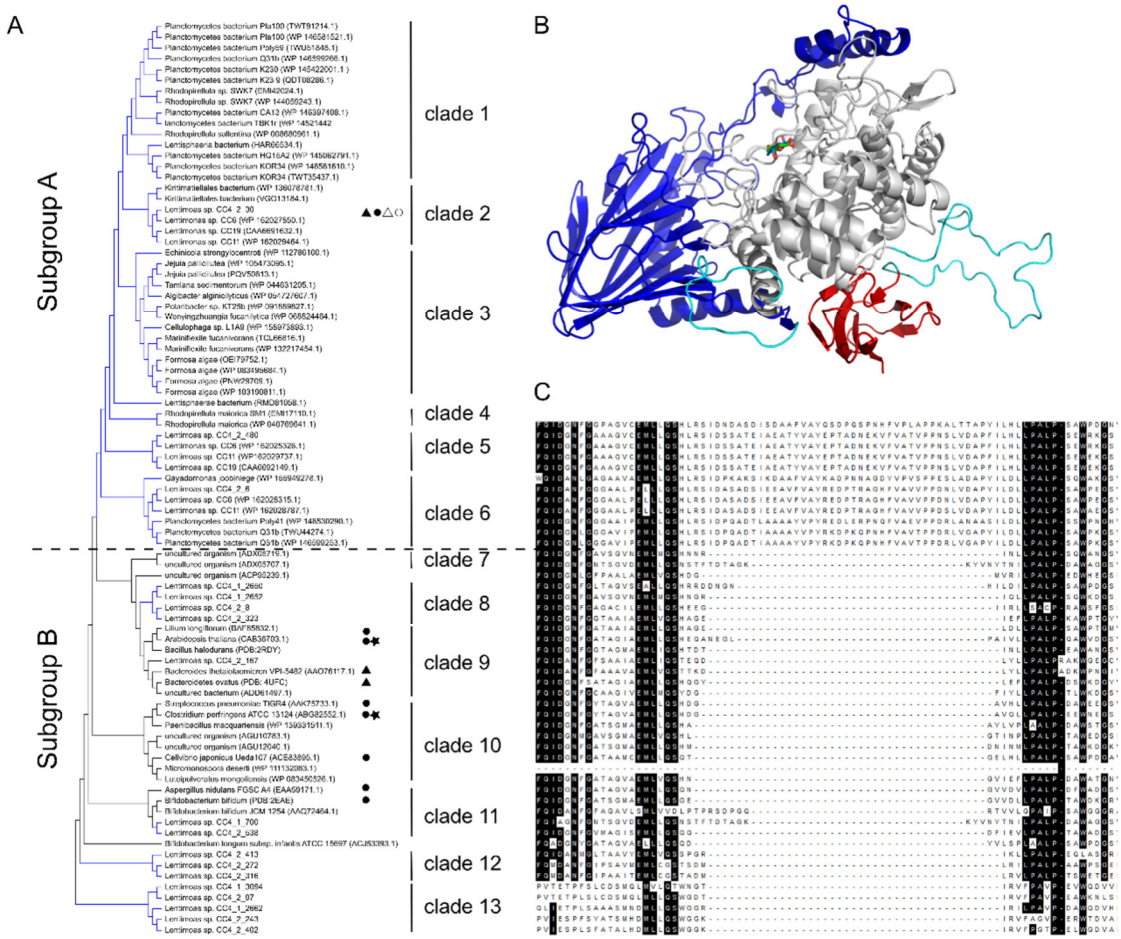


Figure 10. GH95s cluster into clades. (A) The unrooted phylogenetic tree is based on a CLUSTALW alignment and the maximum likelihood approach using the WAG model (Whelan and Goldman 2001) constructed in MEGA V7 (Kumar et al. 2016). A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 1,1994)). The rate variation model allowed for some sites to be evolutionarily invariable ([+*I*], 0,21% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Marine isolates are indicated by a blue branch. The activity of the enzymes: fucosidase (●), galactosidase (▲), no activity against pNP (star), pNP activity against fucose (○) and galactose (Δ). (B) Cartoon representation of model 30_GH95 from '*Lentimonas'* sp. CC4: super sandwich at N-terminus (blue) and loop (turquois, amino acid ~ 707 – 746) connecting the (α/α)6-helical

barrel harboring the active site (grey) with the C-terminal β-sandwich domain (red). The model was generated by Phyre2 (Kelley et al. 2015). Confidence in the model: 769 residues (94 % coverage) modelled at > 90 % accuracy. (C) Multiple sequence alignment: Conserved resides at 90 % level or higher are indicated with a black background.

Quantification of anionic polysaccharides in the oceans

Anionic polysaccharides have been detected in DOC and POC from various water depths all over the world as well as in *in situ* cultures of e.g. single algal species or mesocosms experiments (Verdugo et al. 2004; Hirokawa et al. 2005; Thornton 2009; Underwood et al. 2010). The most commonly used detection method is staining using alcian blue (Alldredge et al. 1993) or ruthenium red (Mitulla et al. 2016), which bind to negatively charged polysaccharides. Alternatively, the content of sulfate and uronic acids (Li et al. 2007a; Lee et al. 2016) as well as the ability to bind cations (De Jong et al. 1976) can be considered as proxies for anionic polysaccharides. These methods do not distinguish between individual polysaccharides and thus, cannot be used as analytical techniques to determine the amount of specific anionic polysaccharides in marine samples. In contrast, enzyme-based approaches are structure-specific and allow quantification of individual polysaccharides within marine organic matter samples (Becker et al. 2017; Scheschonk et al. 2019; Becker et al. 2020). To be able to use the enzymatic approach, the complete degradation pathway of the particular glycan needs to be solved. Anionic polysaccharides like fucoidans and ulvan are complex glycans characterized by a high structural and compositional variety. Their complete degradation by heterotrophic bacteria is either not completely elucidated yet, as it is the case for fucoidans, or requires a multistep enzymatic hydrolysis as it is the case for ulvan utilization (Reisky et al. 2019). An ideal anionic model glycan must be simpler in its structure and degradation pathway. Alginate originated from brown macroalgae is a promising candidate to start developing an enzymatic assay for the quantification of anionic polysaccharides in the ocean.

Alteromonas macleodii is a globally distributed Gammaproteobacterium (Ivars-Martinez et al. 2008), which reaches high abundance in bacterial blooms when organic nutrients are present (Allers et al. 2008; McCarren et al. 2010). Its hydrolytic capacity to degrade easily accessible substrates such as glucose (Allers et al. 2007) and laminarin, complex glycans like alginate and pectin (Neumann et al. 2015; Wietz et al. 2015; Koch et al. 2019a) as well as gel particles (Mitulla et al. 2016), makes it an attractive model organism for alginate degradation. The alginolytic utilization system (AUS) in *A. macleodii* is highly conserved among several strains. AUSs are similar organized to PULs in Bacteroidetes species but missing the characteristic SusCD pair. The model strain *A. macleodii* 83-1 possesses an AUS located on a 24 kp genomic island, which harbors among other enzymes five alginate lyases (Aly) of families PL6, PL7, and PL17 (Neumann et al. 2015).

Each Aly family was produced as recombinant plasmid and biochemical characterized (Fig. S2). The biochemical characterization is crucial to develop the enzymatic kit with maximum robustness, efficiency and accuracy. Since alginate is a heteroglycan consisting of β-D-mannuronate (M) and α-L-guluronate (G), the substrate specificity of each Aly was determined using self-made oligosaccharides enriched in M-, G- or mixed linked MG-resides (Haug et al. 1966). The Alys of families PL6, PL7, and PL17 were active on all tested substrates, but exhibited specificity towards G-enriched alginate (Fig. S3).

Next, different types and sources of alginate were used including bacterial alginate derived from *Azotobacter vinelandii*, alginate particles, DOM samples from a mesocosm microalgae bloom in Bergen, Norway and POM samples extracted from anoxic sediment core from the Shaban Deep, Red Sea. The Alys were incubated with bacterial alginate, alginate particles and OM-samples and analyzed via HPEAC-PAD (Ballance et al. 2005), light microscopy and carbohydrate microarrays (Vidal-Melgosa et al. 2015), respectively. Our preliminary experiments showed that these enzymes are not only active on dissolved alginate but also capable of degrading alginate gel particles (Fig. S4), whereas no activity was detected using bacterial alginate or for DOM and POM samples (data not shown). The absence of activity towards bacterial alginate indicates the need of enzymatic or chemical deacetylation prior to hydrolysis.

The carbohydrate microarray analysis of the POM and DOM samples showed a positive signal for the monoclonal antibody BAM7, which recognizes MG-regions of alginate (Torode et al. 2016), revealing the presence of this polymer in the sample. The enzyme treatment of the carbohydrate arrays should abolish or diminish the BAM7 binding to the sample. However, BAM7 binding remained for the enzyme treatment. These results were surprising and raised the question which exact regions of the carbohydrate are recognized by BAM7. It is likely, that the regions of an acetylated alginate backbone are recognized by the monoclonal antibody BAM7 resulting in a positive signal, while Alys would not be able to degrade the same structure. Alginate detected in sediment samples might not have an algal-derived origin but could be the result of bacterial biofilm production similar to alginate-producing pathogens associated with lung diseases (Sabra et al. 2001). Nevertheless, this hypothesis requires additional experimental data.

To further validate our alginate enzyme kit, the yield of the enzymatic kit was compared with the classic approach of acid hydrolysis and subsequent reducing sugar assay. Samples with high content of uronic acids such as alginate require a harsh hydrolysis (Manns et al. 2014). However, the exact hydrolysis conditions must be controlled carefully, because too harsh condition easily degrade monosaccharides, whereas too mild lead to low hydrolysis yields (Panagiotopoulos and Sempéré 2005). Previous studies, focusing exclusively on alginates, observed unequal rates of release of M and G (Lu et al. 2015). I compared mild (1 h at 100 °C) and harsh (24 h at 100 °C)

acid hydrolysis using two different acids (HCl vs. TFA). Quantification using the reducing sugar assay PAHBAH showed that the harsh acid hydrolysis resulted in higher concentration of alginate than the initial concentration, suggesting the degradation of the monosaccharides. These results are in line with previous experiments showing that 1 h at 100 °C is sufficient for pure alginate (Lu et al. 2015). The mild acid hydrolysis with both acids as well as the enzymatic digestion using all three Alys resulted in near complete hydrolysis (Fig. 11).

In conclusion, our preliminary experiments indicate that the enzymatic kit will result in an accurate quantification of alginate as a near complete hydrolysis is achieved even with three guluronate-specific Alys. The enzyme-based method has the potential to distinguish between different sources of alginates (i.e. macroalgal and bacterial), which is not possible with other methods so far. Moreover, this kit would allow specific quantification of macroalgal alginate in environment samples containing complex organic matter mixtures, which remains challenging with currently used approaches which cannot trace individual glycans without prior purification. For future experiments I suggest investigating the time of incubation needed, as for this study, we only analyzed overnight digestion of alginates. Additionally, it is important to test whether the usage of a mannuronante-specific Aly increases the hydrolysis time. Finally, the enzyme-based method can be applied to environmental samples such as crude extracts of brown algae.

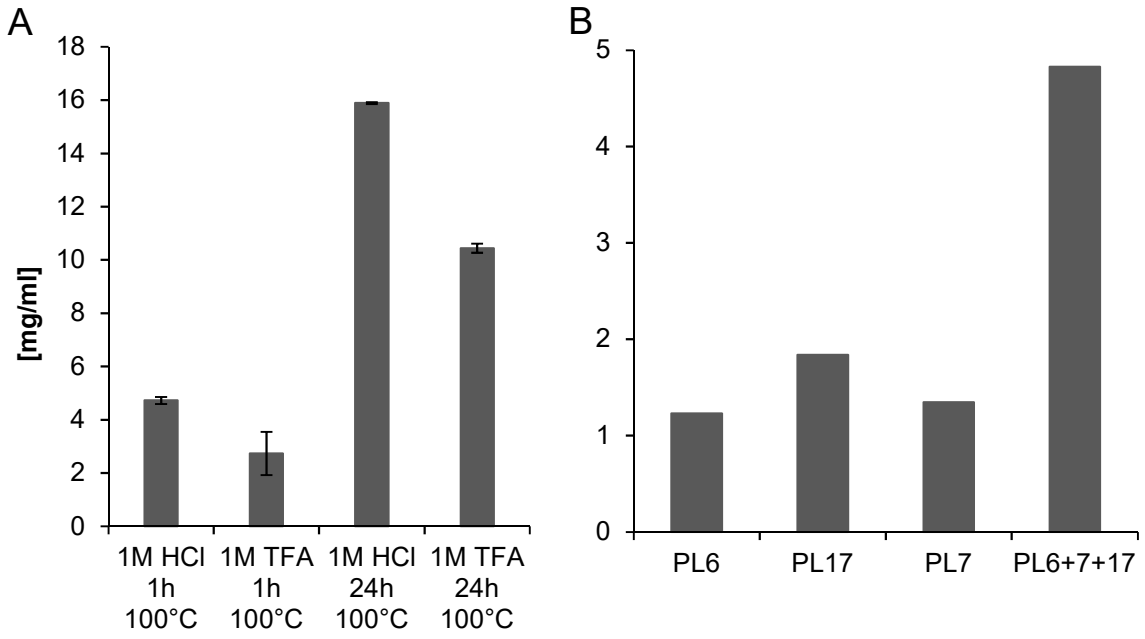


Figure 11. Comparison of (A)acid hydrolysis vs. (B) enzymatic hydrolysis. Low viscosity alginate from brown algae was dissolved in MQ and used in a final concentration of 5 mg ml^{-1}. The enzymatic digest was carried out overnight at 30 °C. Alginate hydrolysis was measured using the reducing sugar assay PAHBAH. Error bars indicate technical triplicates. Enzyme digestion was not setup in triplicates.

Conclusion

The remineralization of algal biomass by heterotrophic bacteria is a crucial step in the marine carbon cycle, but so far little is known about the enzymatic pathways of key degraders. The present study provides detailed insights into the biochemistry, function and structure of bacterial enzymes involved in the degradation of anionic polysaccharides from macroalgae. Anionic polysaccharides are an abundant but at the same time challenging food source for many marine bacteria. Key degraders have adapted and specialized enzymatic machinery to tackle these types of glycans. Consequently, scientists must combine multiple research disciplines ranging from microbial ecology to analytical chemistry, molecular and structural biology to shed light on bacterial glycan utilization. Ultimately, this work provides a better understanding of species interactions and food web dynamics, but also how enzymes can be used as analytical tools to investigate biogeochemical cycles and applied in other field of research. CAZymes and sulfatases represent a fascinating research topic with lots of potential for many more generations of curious scientists to come.

Bibliography

Abbott DW, Eirin-Lopez JM, Boraston AB (2007) Insight into Ligand Diversity and Novel Biological Roles for Family 32 Carbohydrate-Binding Modules. Mol Biol Evol 25:155–167 . https://doi.org/10.1093/molbev/msm243

Abergel C (2013) Molecular replacement: tricks and treats. Acta Crystallogr Sect D Biol Crystallogr 69:2167–2173 . https://doi.org/10.1107/S0907444913015291

Adams PD, Afonine P V., Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr Sect D Biol Crystallogr 66:213–221 . https://doi.org/10.1107/S0907444909052925

Ale MT, Meyer AS (2013) Fucoidans from brown seaweeds: an update on structures, extraction techniques and use of enzymes as tools for structural elucidation. RSC Adv 3:8131–8141 . https://doi.org/10.1039/C3RA23373A

Ale MT, Mikkelsen JD, Meyer AS (2011) Important Determinants for Fucoidan Bioactivity: A Critical Review of Structure-Function Relations and Extraction Methods for Fucose-Containing Sulfated Polysaccharides from Brown Seaweeds. Mar Drugs 9:2106–2130 . https://doi.org/10.3390/md9102106

Ali-Ahmad A, Garron M-L, Zamboni V, Lenfant N, Nurizzo D, Henrissat B, Berrin J-G, Bourne Y, Vincent F (2017) Structural insights into a family 39 glycoside hydrolase from the gut symbiont *Bacteroides cellulosilyticus* WH2. J Struct Biol 197:227–235 . https://doi.org/10.1016/j.jsb.2016.11.004

Alldredge AL, Passow U, Logan BE (1993) The abundance and significance of a class of large, transparent organic particles in the ocean. Deep Sea Res Part I Oceanogr Res Pap 40:1131–1140 . https://doi.org/10.1016/0967-0637(93)90129-Q

Allers E, Gómez-Consarnau L, Pinhassi J, Gasol JM, Šimek K, Pernthaler J (2007) Response of Alteromonadaceae and Rhodobacteriaceae to glucose and phosphorus manipulation in marine mesocosms. Environ Microbiol 9:2417–2429 . https://doi.org/10.1111/j.1462-2920.2007.01360.x

Allers E, Niesner C, Wild C, Pernthaler J (2008) Microbes Enriched in Seawater after Addition of Coral Mucus. Appl Environ Microbiol 74:3274–3278 . https://doi.org/10.1128/AEM.01870-07

Anastyuk SD, Shevchenko NM, Nazarenko EL, Imbs TI, Gorbach VI, Dmitrenok PS, Zvyagintseva TN (2010) Structural analysis of a highly sulfated fucan from the brown alga *Laminaria cichorioides* by tandem MALDI and ESI mass spectrometry. Carbohydr Res 345:2206–2212 . https://doi.org/10.1016/j.carres.2010.07.043

Andrade LR, Leal RN, Noseda M, Duarte MER, Pereira MS, Mourão PAS, Farina M, Amado Filho GM (2010) Brown algae overproduce cell wall polysaccharides as a protection mechanism against the heavy metal toxicity. Mar Pollut Bull 60:1482–1488 . https://doi.org/10.1016/j.marpolbul.2010.05.004

Andrews SR, Charnock SJ, Lakey JH, Davies GJ, Claeyssens M, Nerinckx W, Underwood M, Sinnott ML, Warren RAJ, Gilbert HJ (2000) Substrate specificity in glycoside hydrolase family 10: Tyrosine 87 and leucine 314 play a pivotal role in discriminating between glucose and xylose binding in the proximal active site of pseudomonas cellulosa xylanase 10A. J Biol Chem 275:23027–23033 . https://doi.org/10.1074/jbc.M000128200

Antelmann H, Tjalsma H, Voigt B, Ohlmeier S, Bron S, Dijl JM Van, Hecker M (2001) A Proteomic View on Genome-Based Signal Peptide Predictions The Extracellular Proteome of. Genome Res 11:1484–1502 . https://doi.org/10.1101/gr.182801.way

Ashida H, Miyake A, Kiyohara M, Wada J, Yoshida E, Kumagai H, Katayama T, Yamamoto K (2009) Two distinct α-L-fucosidases from *Bifidobacterium bifidum* are essential for the utilization of fucosylated milk oligosaccharides and glycoconjugates. Glycobiology 19:1010–1017 . https://doi.org/10.1093/glycob/cwp082

Aspeborg H, Coutinho PM, Wang Y, Brumer H, Henrissat B (2012) Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). BMC Evol Biol 12: . https://doi.org/10.1186/1471-2148-12-186

Aspinall GO (1983) The polysaccharides. ACADEMIC PRESS, INC.

Avcı B, Hahnke RL, Chafee M, Fischer T, Gruber-Vodicka H, Tegetmeyer HE, Harder J, Fuchs BM, Amann RI, Teeling H (2017) Genomic and physiological analyses of '*Reinekea forsetii*' reveal a versatile opportunistic lifestyle during spring algae blooms. Environ Microbiol 19:1209–1221 . https://doi.org/10.1111/1462-2920.13646

Avcı B, Krüger K, Fuchs BM, Teeling H, Amann RI (2020) Polysaccharide niche partitioning of distinct Polaribacter clades during North Sea spring algal blooms. ISME J. https://doi.org/10.1038/s41396-020-0601-y

Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil LA, Thingstand F (1983) The Ecological Role of Water-Column Microbes in the Sea. Mar Ecol - Prog Ser 10:257–263 . https://doi.org/10.1021/acs.joc.6b00938

Azam F, Malfatti F (2007) Microbial structuring of marine ecosystems. Nat Rev Microbiol 5:782–791 . https://doi.org/10.1038/nrmicro1747

Bajpai VK, Rather IA, Lim J, Park YH (2014) Diversity of bioactive polysaccharide originated from marine sources: A review. Indian J Geo-Marine Sci 43:1857–1869

Bakunina IY, Nedashkovskaya OI, Alekseeva SA, Ivanova EP, Romanenko LA, Gorshkova NM, Isakov V V., Zvyagintseva TN, Mikhailov V V. (2002) Degradation of fucoidan by the marine proteobacterium *Pseudoalteromonas citrea*. Microbiology 71:49–55 . https://doi.org/10.1023/A:1017994131769

Ballance S, Holtan S, Aarstad OA, Sikorski P, Skjåk-Bræk G, Christensen BE (2005) Application of high-performance anion-exchange chromatography with pulsed amperometric detection and statistical analysis to study oligosaccharide distributions – a complementary method to investigate the structure and some properties of alginates. J Chromatogr A 1093:59–68 . https://doi.org/10.1016/j.chroma.2005.07.051

Barbeyron T, Brillet-Guéguen L, Carré W, Carrière C, Caron C, Czjzek M, Hoebeke M, Michel G (2016a) Matching the diversity of sulfated biomolecules: Creation of a classification database for sulfatases reflecting their substrate specificity. PLoS One 11:1–33 . https://doi.org/10.1371/journal.pone.0164846

Barbeyron T, L'Haridon S, Michel G, Czjzek M (2008) *Mariniflexile fucanivorans* sp. nov., a marine member of the *Flavobacteriaceae* that degrades sulphated fucans from brown algae. Int J Syst Evol Microbiol 58:2107–2113 . https://doi.org/10.1099/ijs.0.65674-0

Barbeyron T, Thomas F, Barbe V, Teeling H, Schenowitz C, Dossat C, Goesmann A, Leblanc C, Oliver Glöckner F, Czjzek M, Amann R, Michel G (2016b) Habitat and taxon as driving forces of carbohydrate catabolism in marine heterotrophic bacteria: example of the model algae-associated bacterium *Zobellia galactanivorans* Dsij T. Environ Microbiol 18:4610–4627 . https://doi.org/10.1111/1462-2920.13584

Bateman A (2002) The Pfam Protein Families Database. Nucleic Acids Res 30:276–280 . https://doi.org/10.1093/nar/30.1.276

Bateman A (2019) UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res 47:D506–D515 . https://doi.org/10.1093/nar/gky1049

Battye TGG, Kontogiannis L, Johnson O, Powell HR, Leslie AGW (2011) iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr Sect D Biol Crystallogr 67:271–281 . https://doi.org/10.1107/S0907444910048675

Becker S, Scheffel A, Polz MF, Hehemann J-H (2017) Accurate Quantification of Laminarin in Marine Organic Matter with Enzymes from Marine Microbes. Appl Environ Microbiol 83:1–14 . https://doi.org/10.1128/AEM.03389-16

Becker S, Tebben J, Coffinet S, Wiltshire K, Iversen MH, Harder T, Hinrichs KU, Hehemann JH (2020) Laminarin is a major molecule in the marine carbon cycle. Proc Natl Acad Sci U S A 117:6599–6607 . https://doi.org/10.1073/pnas.1917001117

Bedê PM, Da Silva MHP, Da Silva Figueiredo ABH, Finotelli PV (2017) Nanostructured magnetic alginate composites for biomedical applications. Polimeros 27:267–272 . https://doi.org/10.1590/0104-1428.2267

Behrenfeld MJ (2014) Climate-mediated dance of the plankton. Nat Clim Chang 4:880–887 . https://doi.org/10.1038/nclimate2349

Berges JA, Falkowski PG (1998) Physiological stress and cell death in marine phytoplankton: Induction of proteases in response to nitrogen or light limitation. Limnol Oceanogr 43:129–135 . https://doi.org/10.4319/lo.1998.43.1.0129

Bergfors T (2003) Seeds to crystals. J Struct Biol 142:66–76 . https://doi.org/10.1016/S1047-8477(03)00039-X

Berman HM (2000) The Protein Data Bank. Nucleic Acids Res 28:235–242 . https://doi.org/10.1093/nar/28.1.235

Bernardes JS, Vieira FRJ, Zaverucha G, Carbone A (2016) A multi-objective optimization approach accurately resolves protein domain architectures. Bioinformatics 32:345–353 . https://doi.org/10.1093/bioinformatics/btv582

Berteau O, Bielicki J, Kilonda A, Machy D, Anson DS, Kenne L (2004) α-L-Fucosidases: Exoglycosidases with Unusual Transglycosylation Properties. Biochemistry 43:7881–7891 . https://doi.org/10.1021/bi036066z

Berteau O, McCort I, Goasdoue N, Tissot B, Daniel R (2002) Characterization of a new α-L-fucosidase isolated from the marine mollusk Pecten maximus that catalyzes the hydrolysis of -L-fucose from algal fucoidan (*Ascophyllum nodosum*). Glycobiology 12:273–282 . https://doi.org/10.1093/glycob/12.4.273

Bilan MI, Grachev AA, Shashkov AS, Nifantiev NE, Usov AI (2006) Structure of a fucoidan from the brown seaweed Fucus serratus L. Carbohydr Res 341:238–245 . https://doi.org/10.1016/j.carres.2005.11.009

Bilan MI, Grachev AA, Ustuzhanina NE, Shashkov AS, Nifantiev NE, Usov AI (2004) A highly regular fraction of a fucoidan from the brown seaweed *Fucus distichus* L. Carbohydr Res 339:511–517 . https://doi.org/10.1016/j.carres.2003.10.028

Bilan MI, Shashkov AS, Usov AI (2014) Structure of a sulfated xylofucan from the brown alga *Punctaria plantaginea*. Carbohydr Res 393:1–8 . https://doi.org/10.1016/j.carres.2014.04.022

Bjursell MK, Martens EC, Gordon JI (2006) Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, *Bacteroides thetaiotaomicron*, to the

suckling period. J Biol Chem 281:36269–36279 . https://doi.org/10.1074/jbc.M606509200

Blair NE, Aller RC (2012) The Fate of Terrestrial Organic Carbon in the Marine Environment. Ann Rev Mar Sci 4:401–423 . https://doi.org/10.1146/annurev-marine-120709-142717

Boraston AB, Bolam DN, Gilbert HJ, Davies GJ (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem J 382:769–781 . https://doi.org/10.1042/BJ20040892

Brito JA, Archer M (2020) Structural biology techniques: X-ray crystallography, cryo-electron microscopy, and small-angle X-ray scattering. In: Practical Approaches to Biological Inorganic Chemistry, 2. Elsevier, pp 375–416

Buchan A, LeCleir GR, Gulvik CA, Gonzalez JM, González JM (2014) Master recyclers: features and functions of bacteria associated with phytoplankton blooms. Nat Rev Microbiol 12:686–698 . https://doi.org/10.1038/nrmicro3326

Burkhardt A, Pakendorf T, Reime B, Meyer J, Fischer P, Stübe N, Panneerselvam S, Lorbeer O, Stachnik K, Warmer M, Rödig P, Göries D, Meents A (2016) Status of the crystallography beamlines at PETRA III. Eur Phys J Plus 131:0–8 . https://doi.org/10.1140/epjp/i2016-16056-0

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:421 . https://doi.org/10.1186/1471-2105-10-421

Cao H, Mikkelsen M, Lezyk M, Bui L, Tran V, Silchenko A, Kusaykin M, Pham T, Truong B, Holck J, Meyer A (2018) Novel Enzyme Actions for Sulphated Galactofucan Depolymerisation and a New Engineering Strategy for Molecular Stabilisation of Fucoidan Degrading Enzymes. Mar Drugs 16:422 . https://doi.org/10.3390/md16110422

Cao H, Walton JD, Brumm P, Phillips GN (2014) Structure and Substrate Specificity of a Eukaryotic Fucosidase from *Fusarium graminearum*. J Biol Chem 289:25624–25638 . https://doi.org/10.1074/jbc.M114.583286

Cartmell A, Lowe EC, Baslé A, Firbank SJ, Ndeh DA, Murray H, Terrapon N, Lombard V, Henrissat B, Turnbull JE, Czjzek M, Gilbert HJ, Bolam DN (2017) How members of the human gut microbiota overcome the sulfation problem posed by glycosaminoglycans. Proc Natl Acad Sci U S A 114:7037–7042 . https://doi.org/10.1073/pnas.1704367114

Caspi R, Billington R, Fulcher CA, Keseler IM, Kothari A, Krummenacker M, Latendresse M, Midford PE, Ong Q, Ong WK, Paley S, Subhraveti P, Karp PD (2018) The MetaCyc database of metabolic pathways and enzymes. Nucleic Acids Res 46:D633–D639 . https://doi.org/10.1093/nar/gkx935

Cesário MT, da Fonseca MMR, Marques MM, de Almeida MCMD (2018) Marine algal carbohydrates as carbon sources for the production of biochemicals and biomaterials. Biotechnol Adv 36:798–817 . https://doi.org/10.1016/j.biotechadv.2018.02.006

Chayen NE, Saridakis E (2008) Protein crystallization: from purified protein to diffraction-quality crystal. Nat Methods 5:147–153 . https://doi.org/10.1038/nmeth.f.203

Chen F, Chang Y, Dong S, Xue C (2016) *Wenyingzhuangia fucanilytica* sp. nov., a sulfated fucan utilizing bacterium isolated from shallow coastal seawater. Int J Syst Evol Microbiol 66:3270–3275 . https://doi.org/10.1099/ijsem.0.001184

Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: All-atom structure validation for macromolecular crystallography. Acta Crystallogr Sect D Biol Crystallogr 66:12–21 . https://doi.org/10.1107/S0907444909042073

Chevolot L, Mulloy B, Ratiskol J, Foucault A, Colliec-Jouault S (2001) A disaccharide repeat unit is the major structure in fucoidans from two species of brown algae. Carbohydr Res 330:529–535 . https://doi.org/10.1016/S0008-6215(00)00314-1

Chin W-C, Orellana M nica V., Verdugo P (2003) Spontaneous assembly of marine dissolved organic matter into polymer gels. Lett to Nat 421:177–182 . https://doi.org/10.1038/nature01282.1.

Chin W-C, Orellana M V., Verdugo P (1998) Spontaneous assembly of marine dissolved organic matter into polymer gels. Nature 391:568–572 . https://doi.org/10.1038/35345

Chothial C, Lesk AM (1986) The relation between the divergence of sequence and structure in proteins. EMBO J 5:823–826

Ciais P, Sabine C, Bala G, Bopp L, Brovkin V, Canadell J, Chhabra A, DeFries R, Galloway J, Heimann M, Jones C, Quéré C Le, Myneni RB, Piao S, P. Thornton (2013) Carbon and other biogeochemical cycles. Clim Chang 2013 Phys Sci Basis Work Gr I Contrib to Fifth Assess Rep Intergov Panel Clim Chang 9781107057:465–570 . https://doi.org/10.1017/CBO9781107415324.015

Cianci M, Bourenkov G, Pompidor G, Karpics I, Kallio J, Bento I, Roessle M, Cipriani F, Fiedler S, Schneider TR (2017) P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable beam focusing. J Synchrotron Radiat 24:323–332 . https://doi.org/10.1107/S1600577516016465

Cobucci-Ponzano B, Trincone A, Giordano A, Rossi M, Moracci M (2003a) Identification of the Catalytic Nucleophile of the Family 29 α-L-Fucosidase from *Sulfolobus solfataricus* via Chemical Rescue of an Inactive Mutant †. Biochemistry 42:9525–9531 . https://doi.org/10.1021/bi035036t

Cobucci-Ponzano B, Trincone A, Giordano A, Rossi M, Moracci M (2003b) Identification of an Archaeal α-L-Fucosidase Encoded by an Interrupted Gene. J Biol Chem 278:14622–14631 . https://doi.org/10.1074/jbc.M211834200

Cockburn D, Wilkens C, Ruzanski C, Andersen S, Willum Nielsen J, Smith AM, Field RA, Willemoës M, Abou Hachem M, Svensson B (2014) Analysis of surface binding sites (SBSs) in carbohydrate active enzymes with focus on glycoside hydrolase families 13 and 77 - a mini-review. Biol 69:705–712 . https://doi.org/10.2478/s11756-014-0373-9

Colin S, Deniaud E, Jam M, Descamps V, Chevolot Y, Kervarec N, Yvin JC, Barbeyron T, Michel G, Kloareg B (2006) Cloning and biochemical characterization of the fucanase FcnA: Definition of a novel glycoside hydrolase family specific for sulfated fucans. Glycobiology 16:1021–1032 . https://doi.org/10.1093/glycob/cwl029

Collén PN, Jeudy A, Sassi J-F, Groisillier A, Czjzek M, Coutinho PM, Helbert W (2014) A Novel Unsaturated β-Glucuronyl Hydrolase Involved in Ulvan Degradation Unveils the Versatility of Stereochemistry Requirements in Family GH105. J Biol Chem 289:6199–6211 . https://doi.org/10.1074/jbc.M113.537480

Collen PN, Sassi JF, Rogniaux H, Marfaing H, Helbert W (2011) Ulvan lyases isolated from the flavobacteria *Persicivirga ulvanivorans* Are the first members of a new polysaccharide lyase family. J Biol Chem 286:42063–42071 . https://doi.org/10.1074/jbc.M111.271825

Collins PM, Douangamath A, Talon R, Dias A, Brandao-Neto J, Krojer T, von Delft F (2018) Achieving a Good Crystal System for Crystallographic X-Ray Fragment Screening. Methods Enzymol 610:251–264 . https://doi.org/10.20944/preprints201809.0383.v1

Connors J, Krzywinski M, Schein J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos : An information aesthetic for comparative genomics. Genome Res 19:1639–1645 .

https://doi.org/10.1101/gr.092759.109.19

Cornuault V, Manfield IW, Ralet MC, Paul Knox J (2014) Epitope detection chromatography: A method to dissect the structural heterogeneity and inter-connections of plant cell-wall matrix glycans. Plant J 78:715–722 . https://doi.org/10.1111/tpj.12504

Cory RM, Harrold KH, Neilson BT, Kling GW (2015) Controls on dissolved organic matter (DOM) degradation in a headwater stream: The influence of photochemical and hydrological conditions in determining light-limitation or substrate-limitation of photo-degradation. Biogeosciences 12:6669–6685 . https://doi.org/10.5194/bg-12-6669-2015

Corzett CH, Elsherbini J, Chien DM, Hehemann J-H, Henschel A, Preheim SP, Yu X, Alm EJ, Polz MF (2018) Evolution of a Vegetarian Vibrio: Metabolic Specialization of *Vibrio breoganii* to Macroalgal Substrates. J Bacteriol 200:1–10 . https://doi.org/10.1128/JB.00020-18

Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr Sect D Biol Crystallogr 62:1002–1011 . https://doi.org/10.1107/S0907444906022116

Cowtan K (2010) Recent developments in classical density modification. Acta Crystallogr Sect D Biol Crystallogr 66:470–478 . https://doi.org/10.1107/S090744490903947X

Croci DO et al. (2011) Fucans, but Not Fucomannoglucuronans, Determine the Biological Activities of Sulfated Polysaccharides from *Laminaria saccharina* Brown Seaweed. PLoS One 6:e17283 . https://doi.org/10.1371/journal.pone.0017283

Crooks G, Hon G, Chandonia J, Brenner S (2004) WebLogo: a sequence logo generator. Genome Res 14:1188–1190 . https://doi.org/10.1101/gr.849004.1

Cuyvers S, Dornez E, Delcour JA, Courtin CM (2012) Occurrence and functional significance of secondary carbohydrate binding sites in glycoside hydrolases. Crit Rev Biotechnol 32:93–107 . https://doi.org/10.3109/07388551.2011.561537

Daniel R, Berteau O, Chevolot L, Varenne A, Gareil P, Goasdoue N (2001) Regioselective desulfation of sulfated L-fucopyranoside by a new sulfoesterase from the marine mollusk *Pecten maximus*. Eur J Biochem 268:5617–5626 . https://doi.org/10.1046/j.1432-1033.2001.02497.x

Davidson IW, Sutherland IW, Lawson CJ (2009) Localization of O-Acetyl Groups of Bacterial Alginate. J Gen Microbiol 98:603–606 . https://doi.org/10.1099/00221287-98-2-603

Davies G, Gilbert H, Henrissat B, Svensson B, Vocadlo D, Williams S (2018) Ten years of CAZypedia: A living encyclopedia of carbohydrate-active enzymes. Glycobiology 28:3–8 . https://doi.org/10.1093/glycob/cwx089

Davies GJ, Wilson KS, Henrissa B (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. Biochem J 321:557–559 . https://doi.org/10.1042/bj3210557

De Jesus Raposo MF, De Morais AMB, De Morais RMSC (2015) Marine polysaccharides from algae with potential biomedical applications. Mar Drugs 13:2967–3028 . https://doi.org/10.3390/md13052967

De Jong EW, Bosch L, Westbrowk P (1976) Isolation and Characterization of a Ca^{2+}-Binding Polysaccharide Associated with Coccoliths of *Emiliania huxleyi* (Lohmann) Kamptner. Eur J Biochem 70:611–621 . https://doi.org/10.1111/j.1432-1033.1976.tb11052.x

de Kruif CG, Weinbreck F, de Vries R (2004) Complex coacervation of proteins and anionic polysaccharides. Curr Opin Colloid Interface Sci 9:340–349 . https://doi.org/10.1016/j.cocis.2004.09.006

Demydchuk M, Hill CH, Zhou A, Bunkóczi G, Stein PE, Marchesan D, Deane JE, Read RJ (2017)

Insights into Hunter syndrome from the structure of iduronate-2-sulfatase. Nat Commun 8: . https://doi.org/10.1038/ncomms15786

Deniaud-Bouët E, Hardouin K, Potin P, Kloareg B, Hervé C (2017) A review about brown algal cell walls and fucose-containing sulfated polysaccharides: Cell wall context, biomedical properties and key research challenges. Carbohydr Polym 175:395–408 . https://doi.org/10.1016/j.carbpol.2017.07.082

Deniaud-Bouët E, Kervarec N, Michel G, Tonon T, Kloareg B, Hervé C (2014) Chemical and enzymatic fractionation of cell walls from Fucales: Insights into the structure of the extracellular matrix of brown algae. Ann Bot 114:1203–1216 . https://doi.org/10.1093/aob/mcu096

Descamps V, Colin S, Lahaye M, Jam M, Richard C, Potin P, Barbeyron T, Yvin J-C, Kloareg B (2006) Isolation and Culture of a Marine Bacterium Degrading the Sulfated Fucans from Marine Brown Algae. Mar Biotechnol 8:27–39 . https://doi.org/10.1007/s10126-005-5107-0

Dierks T, Miech C, Hummerjohann J, Schmidt B, Kertesz MA, von Figura K (1998) Posttranslational Formation of Formylglycine in Prokaryotic Sulfatases by Modification of Either Cysteine or Serine. J Biol Chem 273:25560–25564 . https://doi.org/10.1074/jbc.273.40.25560

Dong S, Chang Y, Shen J, Xue C, Chen F (2017) Purification, expression and characterization of a novel α-L-fucosidase from a marine bacteria *Wenyingzhuangia fucanilytica*. Protein Expr Purif 129:9–17 . https://doi.org/10.1016/j.pep.2016.08.016

Doublier J-L, Garnier C, Renard D, Sanchez C (2000) Protein–polysaccharide interactions. Curr Opin Colloid Interface Sci 5:202–214 . https://doi.org/10.1016/S1359-0294(00)00054-6

Draget K, Smidsrød O, Skjåk-Bræk G (2005) Alginates from algae. In: Biopolymers. pp 1–30

Duarte CM, Cebrián J (1996) The fate of marine autotrophic production. Limnol Oceanogr 41:1758–1766 . https://doi.org/10.4319/lo.1996.41.8.1758

Duarte CM, Middelburg JJ, Caraco N (2004) Major role of marine vegetation on the oceanic carbon cycle. Biogeosciences Discuss 1:659–679 . https://doi.org/10.5194/bgd-1-659-2004

Dubin PL (1988) Electrostatic Effects. J Chromatogr Libr 40:55–75 . https://doi.org/10.1016/S0301-4770(08)60707-1

Eddy SR (2011) Accelerated Profile HMM Searches. PLoS Comput Biol 7:e1002195 . https://doi.org/10.1371/journal.pcbi.1002195

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797 . https://doi.org/10.1093/nar/gkh340

Egan S, Harder T, Burke C, Steinberg P, Kjelleberg S, Thomas T (2013) The seaweed holobiont: Understanding seaweed-bacteria interactions. FEMS Microbiol Rev 37:462–476 . https://doi.org/10.1111/1574-6976.12011

El Gamal AA (2010) Biological importance of marine algae. Saudi Pharm J 18:1–25 . https://doi.org/10.1016/j.jsps.2009.12.001

Emerson S, Hedges J (2008) Chemical oceanography and the marine carbon cycle. Chem Oceanogr Mar Carbon Cycle 9780521833:1–461 . https://doi.org/10.1017/CBO9780511793202

Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta Crystallogr Sect D Biol Crystallogr 66:486–501 . https://doi.org/10.1107/S0907444910007493

Eneyskaya E V., Kulminskaya AA, Kalkkinen N, Nifantiev NE, Arbatskii NP, Saenko AI,

Chepurnaya O V., Arutyunyan A V., Shabalin KA, Neustroev KN (2001a) An alpha-L-fucosidase from *Thermus* sp. with unusually broad specificity. Glycoconj J 18:827–34 . https://doi.org/10.1023/a:1021163720282

Eneyskaya E V., Kulminskaya AA, Kalkkinen N, Nifantiev NE, Arbatskii NP, Saenko AI, Chepurnaya O V., Arutyunyan A V., Shabalin KA, Neustroev KN (2001b) An α-L-fucosidase from *Thermus* sp. with unusually broad specificity. Glycoconj J 18:827–834 . https://doi.org/10.1023/A:1021163720282

Engel A (2000) The role of transparent exopolymer particles (TEP) in the increase in apparent particle stickiness (α) during the decline of a diatom bloom. J Plankton Res 22:485–497 . https://doi.org/10.1093/plankt/22.3.485

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. Mar Chem 127:180–191 . https://doi.org/10.1016/j.marchem.2011.09.004

Engel A, Thoms S, Riebesell U, Rochelle-Newall E (2004) Polysaccharide aggregation as a potential sink of marine dissolved organic carbon. 428:27–30 . https://doi.org/10.1038/nature02506.1.

Engilberge S, Riobé F, Di Pietro S, Lassalle L, Coquelle N, Arnaud C-A, Pitrat D, Mulatier J-C, Madern D, Breyton C, Maury O, Girard E (2017) Crystallophore: a versatile lanthanide complex for protein crystallography combining nucleating effects, phasing properties, and luminescence. Chem Sci 8:5909–5917 . https://doi.org/10.1039/C7SC00758B

Engilberge S, Riobé F, Wagner T, Di Pietro S, Breyton C, Franzetti B, Shima S, Girard E, Dumont E, Maury O (2018) Unveiling the Binding Modes of the Crystallophore, a Terbium-based Nucleating and Phasing Molecular Agent for Protein Crystallography. Chem - A Eur J 24:9739–9746 . https://doi.org/10.1002/chem.201802172

Engilberge S, Wagner T, Santoni G, Breyton C, Shima S, Franzetti B, Riobé F, Maury O, Girard E (2019) Protein crystal structure determination with the crystallophore, a nucleating and phasing agent. J Appl Crystallogr 52:722–731 . https://doi.org/10.1107/S1600576719006381

Espinel-Ingroff A, Fothergill A, Ghannoum M, Manavathu E, Ostrosky-Zeichner L, Pfaller MA, Rinaldi MG, Schell W, Walsh TJ (2007) A Revised Structure for Fucoidan May Explain Some of Its Biological Activities. J Clin Microbiol 45:2180–2182 . https://doi.org/10.1128/JCM.00399-07

Evans PR (2011) An introduction to data reduction: Space-group determination, scaling and intensity statistics. Acta Crystallogr Sect D Biol Crystallogr 67:282–292 . https://doi.org/10.1107/S090744491003982X

Evans PR, Murshudov GN (2013) How good are my data and what is the resolution? Acta Crystallogr Sect D Biol Crystallogr 69:1204–1214 . https://doi.org/10.1107/S0907444913000061

Eymann C, Dreisbach A, Albrecht D, Bernhardt J, Becher D, Gentner S, Tam LT, Büttner K, Buurman G, Scharf C, Venz S, Völker U, Hecker M (2004) A comprehensive proteome map of growing *Bacillus subtilis* cells. Proteomics 4:2849–2876 . https://doi.org/10.1002/pmic.200400907

Falkowski P (2000) The Global Carbon Cycle: A Test of Our Knowledge of Earth as a System. Science (80-) 290:291–296 . https://doi.org/10.1126/science.290.5490.291

Fan S, Zhang H, Chen X, Lu L, Xu L, Xiao M (2016) Cloning, characterization, and production of three α-L-fucosidases from *Clostridium perfringens* ATCC 13124. J Basic Microbiol 56:347–357 . https://doi.org/10.1002/jobm.201500582

Fernandes N, Steinberg P, Rusch D, Kjelleberg S, Thomas T (2012) Community Structure and Functional Gene Profile of Bacteria on Healthy and Diseased Thalli of the Red Seaweed Delisea pulchra. PLoS One 7:1–8 . https://doi.org/10.1371/journal.pone.0050854

Ficko-Blean E et al. (2017) Carrageenan catabolism is encoded by a complex regulon in marine heterotrophic bacteria. Nat Commun 8:1685 . https://doi.org/10.1038/s41467-017-01832-6

Ficko-Blean E, Boraston AB (2012) Insights into the recognition of the human glycome by microbial carbohydrate-binding modules. Curr Opin Struct Biol 22:570–577 . https://doi.org/10.1016/j.sbi.2012.07.009

Field CB, Behrenfeld MJ, Randerson JT, Falkowski PG (1998) Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. Science (80-) 281:237–240 . https://doi.org/10.1126/science.281.5374.237

Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. Nucleic Acids Res 39:W29–W37 . https://doi.org/10.1093/nar/gkr367

Fitton J, Stringer D, Karpiniec S (2015) Therapies from Fucoidan: An Update. Mar Drugs 13:5920–5946 . https://doi.org/10.3390/md13095920

Fletcher HR, Biller P, Ross AB, Adams JMM (2017) The seasonal variation of fucoidan within three species of brown macroalgae. Algal Res 22:79–86 . https://doi.org/10.1016/j.algal.2016.10.015

Foran E, Buravenkov V, Kopel M, Mizrahi N, Shoshani S, Helbert W, Banin E (2017) Functional characterization of a novel “ulvan utilization loci” found in *Alteromonas* sp. LOR genome. Algal Res 25:39–46 . https://doi.org/10.1016/j.algal.2017.04.036

Fuhrman JA, Azam F (1982) Thymidine incorporation as a measure of heterotrophic bacterioplankton production in marine surface waters: Evaluation and field results. Mar Biol 66:109–120 . https://doi.org/10.1007/BF00397184

Gaignard C, Laroche C, Pierre G, Dubessay P, Delattre C, Gardarin C, Gourvil P, Probert I, Dubuffet A, Michaud P (2019) Screening of marine microalgae: Investigation of new exopolysaccharide producers. Algal Res 44:101711 . https://doi.org/10.1016/j.algal.2019.101711

Garbary DJ, Francis S, Scotia N (2001) Biogeography of Marine Algae. Life Sci 1–9 . https://doi.org/10.1038/npg.els.0000312

Garcia-Vaquero M, Rajauria G, O’Doherty J V., Sweeney T (2017) Polysaccharides from macroalgae: Recent advances, innovative technologies and challenges in extraction and purification. Food Res Int 99:1011–1020 . https://doi.org/10.1016/j.foodres.2016.11.016

Garron M, Cygler M (2014) Uronic polysaccharide degrading enzymes. Curr Opin Struct Biol 28:87–95 . https://doi.org/10.1016/j.sbi.2014.07.012

Garron ML, Cygler M (2010) Structural and mechanistic classification of uronic acid-containing polysaccharide lyases. Glycobiology 20:1547–1573 . https://doi.org/10.1093/glycob/cwq122

Gati C, Bourenkov G, Klinge M, Rehders D, Stellato F, Oberthür D, Yefanov O, Sommer BP, Mogk S, Duszenko M, Betzel C, Schneider TR, Chapman HN, Redecke L (2014) Serial crystallography on in vivo grown microcrystals using synchrotron radiation. IUCrJ 1:87–94 . https://doi.org/10.1107/S2052252513033939

Gibson TS, Al Qalla H, McCleary B V. (1992) An improved enzymic method for the measurement of starch damage in wheat flour. J Cereal Sci 15:15–27 . https://doi.org/10.1016/S0733-5210(09)80053-2

Glaser F, Pupko T, Paz I, Bell RE, Bechor-Shental D, Martz E, Ben-Tal N (2003) ConSurf:

Identification of functional regions in proteins by surface-mapping of phylogenetic information. Bioinformatics 19:163–164 . https://doi.org/10.1093/bioinformatics/19.1.163

Goecke F, Labes A, Wiese J, Imhoff J (2010) Chemical interactions between marine macroalgae and bacteria. Mar Ecol Prog Ser 409:267–299 . https://doi.org/10.3354/meps08607

Gower J, King S (2019) Seaweed, seaweed everywhere. Science (80-) 365:27–27 . https://doi.org/10.1126/science.aay0989

Grant G, Morris E, Rees D, Smith P, Thom D (1973) Biological interactions between polysaccharides and divalent cations: The egg-box- model. FEBS Lett 32:195–198 . https://doi.org/10.1002/rmv.519

Gray CJ, Migas LG, Barran PE, Pagel K, Seeberger PH, Eyers CE, Boons G-J, Pohl NLB, Compagnon I, Widmalm G, Flitsch SL (2019) Advancing Solutions to the Carbohydrate Sequencing Challenge. J Am Chem Soc 141:14463–14479 . https://doi.org/10.1021/jacs.9b06406

Groisillier A, Hervé C, Jeudy A, Rebuffet E, Pluchon PF, Chevolot Y, Flament D, Geslin C, Morgado IM, Power D, Branno M, Moreau H, Michel G, Boyen C, Czjzek M (2010) MARINE-EXPRESS: taking advantage of high throughput cloning and expression strategies for the post-genomic analysis of marine organisms. Microb Cell Fact 9:45 . https://doi.org/10.1186/1475-2859-9-45

Grube M, Cernava T, Soh J, Fuchs S, Aschenbrenner I, Lassek C, Wegner U, Becher D, Riedel K, Sensen CW, Berg G (2015) Exploring functional contexts of symbiotic sustain within lichen-associated bacteria by comparative omics. ISME J 9:412–424 . https://doi.org/10.1038/ismej.2014.138

Guex N, Peitsch MC, Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis 30:162–173 . https://doi.org/10.1002/elps.200900140

Gügi B, Le Costaouec T, Burel C, Lerouge P, Helbert W, Bardor M (2015) Diatom-Specific Oligosaccharide and Polysaccharide Structures Help to Unravel Biosynthetic Capabilities in Diatoms. Mar Drugs 13:5993–6018 . https://doi.org/10.3390/md13095993

Guillotin L, Lafite P, Daniellou R (2014) Unraveling the Substrate Recognition Mechanism and Specificity of the Unusual Glycosyl Hydrolase Family 29 BT2192 from *Bacteroides thetaiotaomicron*. Biochemistry 53:1447–1455 . https://doi.org/10.1021/bi400951q

Gurpilhares D de B, Moreira TR, Bueno J da L, Cinelli LP, Mazzola PG, Pessoa A, Sette LD (2016) "Algae's sulfated polysaccharides modifications: Potential use of microbial enzymes." Process Biochem 51:989–998 . https://doi.org/10.1016/j.procbio.2016.04.020

Hanson SR, Best MD, Wong C-H (2004) Sulfatases: Structure, Mechanism, Biological Activity, Inhibition, and Synthetic Utility. Angew Chemie Int Ed 43:5736–5763 . https://doi.org/10.1002/anie.200300632

Haug A, Larsen B, Smidsrød O (1966) A Study of the Consitution of Alginic Acid by Partial Acid Hydrolysis. Acta Chem Scand 20:183–190

Haug A, Larsen B, Smidsrød O (1967) Studies on the Sequence of Uronic Acid Residues in Alginic Acid. Acta Chem. Scand. 21:691–704

Hedges JI, Baldock JA, Gélinas Y, Lee C, Peterson M, Wakeham SG (2001) Evidence for non-selective preservation of organic matter in sinking marine particles. Nature 409:801–804 . https://doi.org/10.1038/35057247

Hehemann J-H, Arevalo P, Datta MS, Yu X, Corzett CH, Henschel A, Preheim SP, Timberlake S,

Alm EJ, Polz MF (2016) Adaptive radiation by waves of gene transfer leads to fine-scale resource partitioning in marine microbes. Nat Commun 7:12860 . https://doi.org/10.1038/ncomms12860

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

Hehemann J-H, Kelly AG, Pudlo NA, Martens EC, Boraston AB (2012) Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. PNAS 109:19786–91 . https://doi.org/10.1073/pnas.1211002109

Hehemann J-H, Truong L Van, Unfried F, Welsch N, Kabisch J, Heiden SE, Junker S, Becher D, Thürmer A, Daniel R, Amann R, Schweder T (2017) Aquatic adaptation of a laterally acquired pectin degradation pathway in marine gammaproteobacteria. Environ Microbiol 19:2320–2333 . https://doi.org/10.1111/1462-2920.13726

Heinze S, Mechelke M, Kornberger P, Liebl W, Schwarz WH, Zverlov V V. (2017) Identification of endoxylanase XynE from *Clostridium thermocellum* as the first xylanase of glycoside hydrolase family GH141. Sci Rep 7:11178 . https://doi.org/10.1038/s41598-017-11598-y

Hemmingson JA, Falshaw R, Furneaux RH, Thompson K (2006) Structure and Antiviral Activity of the Galactofucan Sulfates Extracted from *Undaria pinnatifida* (Phaeophyta). J Appl Phycol 18:185–193 . https://doi.org/10.1007/s10811-006-9096-9

Hemsworth GR, Dejean G, Davies GJ, Brumer H (2016) Learning from microbial strategies for polysaccharide degradation. Biochem Soc Trans 44:94–108 . https://doi.org/10.1042/BST20150180

Hettle AG, Hobbs JK, Pluvinage B, Vickers C, Abe KT, Salama-Alber O, McGuire BE, Hehemann J-H, Hui JPM, Berrue F, Banskota A, Zhang J, Bottos EM, Van Hamme J, Boraston AB (2019) Insights into the κ/ι-carrageenan metabolism pathway of some marine *Pseudoalteromonas* species. Commun Biol 2:474 . https://doi.org/10.1038/s42003-019-0721-y

Hettle AG, Vickers C, Robb CS, Liu F, Withers SG, Hehemann J-H, Boraston AB (2018) The Molecular Basis of Polysaccharide Sulfatase Activity and a Nomenclature for Catalytic Subsites in this Class of Enzyme. Structure 26:747-758.e4 . https://doi.org/10.1016/j.str.2018.03.012

Hill R, Bellgrove A, Macreadie PI, Petrou K, Beardall J, Steven A, Ralph PJ (2015) Can macroalgae contribute to blue carbon? An Australian perspective. Limnol Oceanogr 60:1689–1706 . https://doi.org/10.1002/lno.10128

Hirokawa Y, Fujiwara S, Tsuzuki M (2005) Three Types of Acidic Polysaccharides Associated with Coccolith of *Pleurochrysis haptonemofera*: Comparison with *Pleurochrysis carterae* and Analysis Using Fluorescein-Isothiocyanate-Labeled Lectins. Mar Biotechnol 7:634–644 . https://doi.org/10.1007/s10126-004-5148-9

Hobbs JK, Lee SM, Robb M, Hof F, Barr C, Abe KT, Hehemann J-H, McLean R, Abbott DW, Boraston AB (2016) KdgF, the missing link in the microbial metabolism of uronate sugars from pectin and alginate. Proc Natl Acad Sci 113:6188–6193 . https://doi.org/10.1073/pnas.1524214113

Hobbs JK, Pluvinage B, Robb M, Smith SP, Boraston AB (2019) Two complementary α-fucosidases from *Streptococcus pneumoniae* promote complete degradation of host-derived carbohydrate antigens. J Biol Chem 294:12670–12682 . https://doi.org/10.1074/jbc.RA119.009368

Hochuli E, Bannwarth W, Döbeli H, Gentz R, Stüber D (1988) Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent. Nat Biotechnol 6:1321–1325 . https://doi.org/10.1038/nbt1188-1321

Holgersson J, Gustafsson A, Breimer ME (2005) Characteristics of protein-carbohydrate interactions as a basis for developing novel carbohydrate-based antirejection therapies. Immunol Cell Biol 83:694–708 . https://doi.org/10.1111/j.1440-1711.2005.01373.x

Holm L, Elofsson A (2019) Benchmarking fold detection by DaliLite v.5. Bioinformatics 35:5326–5327 . https://doi.org/10.1093/bioinformatics/btz536

Holtkamp AD, Kelly S, Ulber R, Lang S (2009) Fucoidans and fucoidanases—focus on techniques for molecular structure elucidation and modification of marine polysaccharides. Appl Microbiol Biotechnol 82:1–11 . https://doi.org/10.1007/s00253-008-1790-x

Hu P, Li Z, Chen M, Sun Z, Ling Y, Jiang J, Huang C (2016) Structural elucidation and protective role of a polysaccharide from *Sargassum fusiforme* on ameliorating learning and memory deficiencies in mice. Carbohydr Polym 139:150–158 . https://doi.org/10.1016/j.carbpol.2015.12.019

Hu P, Xue R, Li Z, Chen M, Sun Z, Jiang J, Huang C (2014) Structural investigation and immunological activity of a heteropolysaccharide from *Sargassum fusiforme*. Carbohydr Res 390:28–32 . https://doi.org/10.1016/j.carres.2014.02.027

Huerta-cepas J (2016) ETE 3 : Reconstruction , Analysis , and Visualization of Phylogenomic Data Brief communication. https://doi.org/10.1093/molbev/msw046

Ihua M, Guihéneuf F, Mohammed H, Margassery L, Jackson S, Stengel D, Clarke D, Dobson A (2019) Microbial Population Changes in Decaying *Ascophyllum nodosum* Result in Macroalgal-Polysaccharide-Degrading Bacteria with Potential Applicability in Enzyme-Assisted Extraction Technologies. Mar Drugs 17:200 . https://doi.org/10.3390/md17040200

IPCC (2013) Climate Change 2013. www.climatechange2013.org.

Ivanova EP, Alexeeva Y V., Flavier S, Wright JP, Zhukova N V., Gorshkova NM, Mikhailov V V., Nicolau D V., Christen R (2004) *Formosa* algae gen. nov., sp. nov., a novel member of the family Flavobacteriaceae. Int J Syst Evol Microbiol 54:705–711 . https://doi.org/10.1099/ijs.0.02763-0

Ivars-Martínez E, D'Auria G, Rodríguez-Valera F, Sánchez-Porro C, Ventosa A, Joint I, Mühling M (2008) Biogeography of the ubiquitous marine bacterium *Alteromonas macleodii* determined by multilocus sequence analysis. Mol Ecol 17:4092–4106 . https://doi.org/10.1111/j.1365-294X.2008.03883.x

Ivars-Martinez E, Martin-Cuadrado A-B, D'Auria G, Mira A, Ferriera S, Johnson J, Friedman R, Rodriguez-Valera F (2008) Comparative genomics of two ecotypes of the marine planktonic copiotroph *Alteromonas macleodii* suggests alternative lifestyles associated with different kinds of particulate organic matter. ISME J 2:1194–1212 . https://doi.org/10.1038/ismej.2008.74

Jiao N et al. (2014) Mechanisms of microbial carbon sequestration in the ocean – future research directions. Biogeosciences 11:5285–5306 . https://doi.org/10.5194/bg-11-5285-2014

Jiao N, Cai R, Zheng Q, Tang K, Liu J, Jiao F, Wallace D, Chen F, Li C, Amann R, Benner R, Azam F (2018) Unveiling the enigma of refractory carbon in the ocean. Natl Sci Rev 5:459–463 . https://doi.org/10.1093/nsr/nwy020

Jiao N, Herndl GJ, Hansell DA, Benner R, Kattner G, Wilhelm SW, Kirchman DL, Weinbauer MG, Luo T, Chen F, Azam F (2010) Microbial production of recalcitrant dissolved organic matter: Long-term carbon storage in the global ocean. Nat Rev Microbiol 8:593–599 .

https://doi.org/10.1038/nrmicro2386

Jiao N, Zheng Q (2011) The microbial carbon pump: From genes to ecosystems. Appl Environ Microbiol 77:7439–7444 . https://doi.org/10.1128/AEM.05640-11

Jones DR, Uddin MS, Gruninger RJ, Pham TTM, Thomas D, Boraston AB, Briggs J, Pluvinage B, McAllister TA, Forster RJ, Tsang A, Selinger LB, Abbott DW (2017) Discovery and characterization of family 39 glycoside hydrolases from rumen anaerobic fungi with polyspecific activity on rare arabinosyl substrates. J Biol Chem 292:12606–12620 . https://doi.org/10.1074/jbc.M117.789008

Kabsch W, T. BA, K. D, A. KP, K. D, S. M, G. RRB, P. E, S. F, K. W, W. K, W. K, W. K, W. K, P. K, S. WM (2010) *XDS*. Acta Crystallogr Sect D Biol Crystallogr 66:125–132 . https://doi.org/10.1107/S0907444909047337

Kannan RRR, Arumugam R, Anantharaman P (2013) Pharmaceutical potential of a fucoidan-like sulphated polysaccharide isolated from *Halodule pinifolia*. Int J Biol Macromol 62:30–34 . https://doi.org/10.1016/j.ijbiomac.2013.08.005

Kappelmann L, Krüger K, Hehemann J-H, Harder J, Markert S, Unfried F, Becher D, Shapiro N, Schweder T, Amann RI, Teeling H (2019) Polysaccharide utilization loci of North Sea *Flavobacteriia* as basis for using SusC/D-protein expression for predicting major phytoplankton glycans. ISME J 13:76–91 . https://doi.org/10.1038/s41396-018-0242-6

Karsenti E et al. (2011) A holistic approach to marine Eco-systems biology. PLoS Biol 9:7–11 . https://doi.org/10.1371/journal.pbio.1001177

Katayama T, Sakuma A, Kimura T, Makimura Y, Hiratake J, Sakata K, Yamanoi T, Kumagai H, Yamamoto K (2004) Molecular Cloning and Characterization of *Bifidobacterium bifidum* 1,2-α-l-Fucosidase (AfcA), a Novel Inverting Glycosidase (Glycoside Hydrolase Family 95). J Bacteriol 186:4885–4893 . https://doi.org/10.1128/JB.186.15.4885-4893.2004

Kawai S, Murata K (2016) Biofuel Production Based on Carbohydrates from Both Brown and Red Macroalgae: Recent Developments in Key Biotechnologies. Int J Mol Sci 17:145 . https://doi.org/10.3390/ijms17020145

Keegan RM, Winn MD (2007) Automated search-model discovery and preparation for structure solution by molecular replacement. Acta Crystallogr Sect D Biol Crystallogr 63:447–457 . https://doi.org/10.1107/S0907444907002661

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015) The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10:845–858 . https://doi.org/10.1038/nprot.2015.053

Khailov KM, Burlakova ZP (1969) Release of Dissolved Organic Matter By Marine Seaweads and Distribution of Their Total Organic Production To Inshore Communities. Limnol Ocean 14:

Kim JW, Brawley SH, Prochnik S, Chovatia M, Grimwood J, Jenkins J, LaButti K, Mavromatis K, Nolan M, Zane M, Schmutz J, Stiller JW, Grossman AR (2016) Genome Analysis of Planctomycetes Inhabiting Blades of the Red Alga *Porphyra umbilicalis*. PLoS One 11:e0151883 . https://doi.org/10.1371/journal.pone.0151883

Kim W, Park J, Park J, Choi D, Park Y (2015) Purification and Characterization of a Fucoidanase (FNase S) from a Marine Bacterium *Sphingomonas paucimobilis* PF-1. Mar Drugs 13:4398–4417 . https://doi.org/10.3390/md13074398

Kirkman H, Kendrick GA (1997) Ecological significance and commercial harvesting of drifting and beach-cast macro-algae and seagrasses in Australia: A review. J Appl Phycol 9:311–326 . https://doi.org/10.1023/A:1007965506873

Klemetsen T, Raknes IA, Fu J, Agafonov A, Balasundaram S V., Tartari G, Robertsen E, Willassen NP (2018) The MAR databases: Development and implementation of databases specific for marine metagenomics. Nucleic Acids Res 46:D693–D699 . https://doi.org/10.1093/nar/gkx1036

Kloareg B, Quatrano RS (1988) Structure of the Cell Walls of Marine Algae and Ecophysiological Functions of the Matrix Polysaccharides. Ocean Mar Biol 25:259–315

Koch H, Dürwald A, Schweder T, Noriega-Ortega B, Vidal-Melgosa S, Hehemann J-H, Dittmar T, Freese HM, Becher D, Simon M, Wietz M (2019a) Biphasic cellular adaptations and ecological implications of *Alteromonas macleodii* degrading a mixture of algal polysaccharides. ISME J 13:92–103 . https://doi.org/10.1038/s41396-018-0252-4

Koch H, Freese HM, Hahnke RL, Simon M, Wietz M (2019b) Adaptations of *Alteromonas* sp. 76-1 to Polysaccharide Degradation: A CAZyme Plasmid for Ulvan Degradation and Two Alginolytic Systems. Front Microbiol 10:1–12 . https://doi.org/10.3389/fmicb.2019.00504

Koopmann R et al. (2012) In vivo protein crystallization opens new routes in structural biology. Nat Methods 9:259–262 . https://doi.org/10.1038/nmeth.1859

Kopel M, Helbert W, Belnik Y, Buravenkov V, Herman A, Banin E (2016) New family of ulvan lyases identified in three isolates from the Alteromonadales order. J Biol Chem 291:5871–5878 . https://doi.org/10.1074/jbc.M115.673947

Kopel M, Helbert W, Henrissat B, Doniger T, Banin E (2014a) Draft Genome Sequence of *Nonlabens ulvanivorans*, an Ulvan-Degrading Bacterium. Genome Announc 2:3–4 . https://doi.org/10.1128/genomeA.00793-14.Copyright

Kopel M, Helbert W, Henrissat B, Doniger T, Banin E (2014b) Draft Genome Sequence of *Pseudoalteromonas* sp. Strain PLSV, an Ulvan-Degrading Bacterium. Genome Announc 2:4–6 . https://doi.org/10.1128/genomeA.01257-14

Kopel M, Helbert W, Henrissat B, Doniger T, Banin E (2014c) Draft genome sequences of two ulvan-degrading isolates, strains LTR and LOR , that belong to the *Alteromonas* genus. Genome Announc 2:13–15 . https://doi.org/10.1128/genomeA.01081-14

Kopplin G, Rokstad AM, Mélida H, Bulone V, Skjåk-Bræk G, Aachmann FL (2018) Structural Characterization of Fucoidan from *Laminaria hyperborea* : Assessment of Coagulation and Inflammatory Properties and Their Structure–Function Relationship. ACS Appl Bio Mater 1:1880–1892 . https://doi.org/10.1021/acsabm.8b00436

Korndörfer IP, Fessner WD, Matthews BW (2000) The structure of rhamnose isomerase from Escherichia coli and its relation with xylose isomerase illustrates a change between inter and intra-subunit complementation during evolution. J Mol Biol 300:917–933 . https://doi.org/10.1006/jmbi.2000.3896

Kraan S (2012) Algal polysaccharides, novel applications and outlook. In: Carbohydrates - Comprehensive Studies on Glycobiology and Glycotechnology. pp 489–532

Krause-Jensen D, Duarte CM (2016) Substantial role of macroalgae in marine carbon sequestration. Nat Geosci 9:737–742 . https://doi.org/10.1038/ngeo2790

Krissinel E, Uski V, Lebedev A, Winn M, Ballard C (2018) Distributed computing for macromolecular crystallography. Acta Crystallogr Sect D Struct Biol 74:143–151 . https://doi.org/10.1107/S2059798317014565

Krüger K, Chafee M, Ben Francis T, Glavina del Rio T, Becher D, Schweder T, Amann RI, Teeling H (2019) In marine Bacteroidetes the bulk of glycan degradation during algae blooms is mediated by few clades using a restricted set of genes. ISME J 1–17 . https://doi.org/10.1038/s41396-019-0476-y

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol 33:msw054 . https://doi.org/10.1093/molbev/msw054

Küpper FC, Kloareg B, Guern J, Potin P (2001) Oligoguluronates Elicit an Oxidative Burst in the Brown Algal Kelp *Laminaria digitata*. Plant Physiol 125:278–291 . https://doi.org/10.1104/pp.125.1.278

Kusaykin MI, Chizhov AO, Grachev AA, Alekseeva SA, Bakunina IY, Nedashkovskaya OI, Sova V V., Zvyagintseva TN (2006) A comparative study of specificity of fucoidanases from marine microorganisms and invertebrates. J Appl Phycol 18:369–373 . https://doi.org/10.1007/s10811-006-9042-x

Kusaykin MI, Silchenko AS, Zakharenko AM, Zvyagintseva TN (2015) Fucoidanases. Glycobiology 26:cwv072 . https://doi.org/10.1093/glycob/cwv072

Kylin H (1913) Zur Biochemie der Meeresalgen. Hoppe-Seyler´s Zeitschrift für Physiol Chemie 83:171–197 . https://doi.org/10.1515/bchm2.1913.83.3.171

Lachnit T, Meske D, Wahl M, Harder T, Schmitz R (2011) Epibacterial community patterns on marine macroalgae are host-specific but temporally variable. Environ Microbiol 13:655–665 . https://doi.org/10.1111/j.1462-2920.2010.02371.x

Ladner CL, Yang J, Turner RJ, Edwards RA (2004) Visible fluorescent detection of proteins in polyacrylamide gels without staining. Anal Biochem 326:13–20 . https://doi.org/10.1016/j.ab.2003.10.047

LAEMMLI UK (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature 227:680–685 . https://doi.org/10.1038/227680a0

Lahaye M (1998) NMR spectroscopic characterisation of oligosaccharides from two *Ulva rigida* ulvan samples (Ulvales, Chlorophyta) degraded by a lyase. Carbohydr Res 314:1–12 . https://doi.org/10.1016/S0008-6215(98)00293-6

Lahaye M, Brunel M, Bonnin E (1997) Fine chemical structure analysis of oligosaccharides produced by an ulvan-lyase degradation of the water-soluble cell-wall polysaccharides from Ulua sp. (Ulvales, Chlorophyta). Carbohydr Res 304:325–333 . https://doi.org/10.1016/S0008-6215(97)00270-X

Lahaye M, Robic A (2007) Structure and function properties of Ulvan, a polysaccharide from green seaweeds. Biomacromolecules 8:1765–1774 . https://doi.org/10.1021/bm061185q

Laine RA (1994) Invited Commentary: A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05 × 10 12 structures for a reducing hexasaccharide: the Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. Glycobiology 4:759–767 . https://doi.org/10.1093/glycob/4.6.759

Langkamp EJAN, Uijterlinde C (2015) Recovery and reuse of alginate from granular Nereda sludge. Reseearch Innov 48

Lara A, Rodríguez-Jasso RM, Loredo-Trevino A, Aguilar CN, Meyer AS, Ruiz HA (2020) 17. ENZYMES IN THE THIRD GENERATION BIOREFINERY FOR MACROALGAE BIOMASS. In: Biomass, Biofuels, Biochemicals. Elsevier, pp 363–396

Laskowski RA, Jabłońska J, Pravda L, Vařeková RS, Thornton JM (2018) PDBsum: Structural summaries of PDB entries. Protein Sci 27:129–134 . https://doi.org/10.1002/pro.3289

Laurienzo P (2010) Marine polysaccharides in pharmaceutical applications: An overview. Mar Drugs 8:2435–2465 . https://doi.org/10.3390/md8092435

Lee J-B, Hayashi K, Hashimoto M, Nakano T, Hayashi T (2004) Novel Antiviral Fucoidan from Sporophyll of *Undaria pinnatifida* (Mekabu). Chem Pharm Bull (Tokyo) 52:1091–1094 .

https://doi.org/10.1248/cpb.52.1091

Lee RBY, Mavridou DAI, Papadakos G, McClelland HLO, Rickaby REM (2016) The uronic acid content of coccolith-associated polysaccharides provides insight into coccolithogenesis and past climate. Nat Commun 7:1–7 . https://doi.org/10.1038/ncomms13144

Léonard R, Pabst M, Bondili JS, Chambat G, Veit C, Strasser R, Altmann F (2008) Identification of an Arabidopsis gene encoding a GH95 alpha1,2-fucosidase active on xyloglucan oligo- and polysaccharides. Phytochemistry 69:1983–1988 . https://doi.org/10.1016/j.phytochem.2008.03.024

Li C, Wen A, Shen B, Lu J, Huang Y, Chang Y (2011) FastCloning: a highly simplified, purification-free, sequence- and ligation-independent PCR cloning method. BMC Biotechnol 11:92 . https://doi.org/10.1186/1472-6750-11-92

Li J, Kisara K, Danielsson S, Lindström ME, Gellerstedt G (2007a) An improved methodology for the quantification of uronic acid units in xylans and other polysaccharides. Carbohydr Res 342:1442–1449 . https://doi.org/10.1016/j.carres.2007.03.031

Li L, Fang Y, Vreeker R, Appelqvist I, Mendes E (2007b) Reexamining the egg-box model in calcium - Alginate gels with X-ray diffraction. Biomacromolecules 8:464–468 . https://doi.org/10.1021/bm060550a

Li Q, Hu F, Zhu B, Ni F, Yao Z (2020) Insights into ulvan lyase: review of source, biochemical characteristics, structure and catalytic mechanism. Crit Rev Biotechnol 40:432–441 . https://doi.org/10.1080/07388551.2020.1723486

Lim SJ, Wan Aida WM, Schiehser S, Rosenau T, Böhmdorfer S (2019) Structural elucidation of fucoidan from *Cladosiphon okamuranus* (Okinawa mozuku). Food Chem 272:222–226 . https://doi.org/10.1016/j.foodchem.2018.08.034

Liu D, Keesing JK, He P, Wang Z, Shi Y, Wang Y (2013) The world's largest macroalgal bloom in the Yellow Sea, China: Formation and implications. Estuar Coast Shelf Sci 129:2–10 . https://doi.org/10.1016/j.ecss.2013.05.021

Liu H, Naismith JH (2008) An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. BMC Biotechnol 8:91 . https://doi.org/10.1186/1472-6750-8-91

Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014a) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:D490–D495 . https://doi.org/10.1093/nar/gkt1178

Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B (2014b) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:490–495 . https://doi.org/10.1093/nar/gkt1178

Lu J, Yang H, Hao J, Wu C, Liu L, Xu N, Linhardt RJ, Zhang Z (2015) Impact of hydrolysis conditions on the detection of mannuronic to guluronic acid ratio in alginate and its derivatives. Carbohydr Polym 122:180–188 . https://doi.org/10.1016/j.carbpol.2015.01.008

Lukatela G, Krauss N, Theis K, Selmer T, Gieseimann V, Von Figura K, Saenger W (1998) Crystal structure of human arylsulfatase A: The aldehyde function and the metal ion at the active site suggest a novel mechanism for sulfate ester hydrolysis. Biochemistry 37:3654–3664 . https://doi.org/10.1021/bi9714924

Mabeau S, Kloareg B (1987) Isolation and Analysis of the Cell Walls of Brown Algae: *Fucus spiralis*, *F. ceranoides*, *F. vesiculosus*, *F. serratus*, *Bifurcaria bifurcata* and *Laminaria digitata*. J Exp Bot 38:1573–1580 . https://doi.org/10.1093/jxb/38.9.1573

MacCoss MJ, Wu CC, Liu H, Sadygov R, Yates JR (2003) A Correlation Algorithm for the Automated Quantitative Analysis of Shotgun Proteomics Data. Anal Chem 75:6912–6921 . https://doi.org/10.1021/ac034790h

Macreadie PI, Jarvis J, Trevathan-Tackett SM, Bellgrove A (2017) Seagrasses and Macroalgae: Importance, Vulnerability and Impacts. In: Climate Change Impacts on Fisheries and Aquaculture. John Wiley & Sons, Ltd, Chichester, UK, pp 729–770

Magnusdottir A, Johansson I, Dahlgren L-G, Nordlund P, Berglund H (2009) Enabling IMAC purification of low abundance recombinant proteins from *E. coli* lysates. Nat Methods 6:477–478 . https://doi.org/10.1038/nmeth0709-477

Mahowald NM, Baker AR, Bergametti G, Brooks N, Duce RA, Jickells TD, Kubilay N, Prospero JM, Tegen I (2005) Atmospheric global dust cycle and iron inputs to the ocean. Global Biogeochem Cycles 19: . https://doi.org/10.1029/2004GB002402

Maita N, Tsukimura T, Taniguchi T, Saito S, Ohno K, Taniguchi H, Sakuraba H (2013) Human α-L-iduronidase uses its own N-glycan as a substrate-binding and catalytic module. Proc Natl Acad Sci U S A 110:14628–14633 . https://doi.org/10.1073/pnas.1306939110

Mangan D, McCleary B V., Culleton H, Cornaggia C, Ivory R, McKie VA, Delaney E, Kargelis T (2019) A novel enzymatic method for the measurement of lactose in lactose-free products. J Sci Food Agric 99:947–956 . https://doi.org/10.1002/jsfa.9317

Mann AJ, Hahnke RL, Huang S, Werner J, Xing P, Barbeyron T, Huettel B, Stüber K, Reinhardt R, Harder J, Glöckner FO, Amann RI, Teeling H (2013) The Genome of the Alga-Associated Marine Flavobacterium *Formosa agariphila* KMM3901$_T$ Reveals a Broad Potential for Degradation of Algal Polysaccharides. Appl Environ Microbiol 79:6813–6822 . https://doi.org/10.1128/AEM.01937-13

Manns D, Deutschle AL, Saake B, Meyer AS (2014) Methodology for quantitative determination of the carbohydrate composition of brown seaweeds (Laminariaceae). RSC Adv 4:25736–25746 . https://doi.org/10.1039/c4ra03537b

Martin M, Barbeyron T, Martin R, Portetelle D, Michel G, Vandenbol M (2015) The Cultivable Surface Microbiota of the Brown Alga *Ascophyllum nodosum* is Enriched in Macroalgal-Polysaccharide-Degrading Bacteria. Front Microbiol 6: . https://doi.org/10.3389/fmicb.2015.01487

Martin M, Portetelle D, Michel G, Vandenbol M (2014) Microorganisms living on macroalgae: diversity, interactions, and biotechnological applications. Appl Microbiol Biotechnol 98:2917–2935 . https://doi.org/10.1007/s00253-014-5557-2

McCarren J, Becker JW, Repeta DJ, Shi Y, Young CR, Malmstrom RR, Chisholm SW, DeLong EF (2010) Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. Proc Natl Acad Sci U S A 107:16420–7 . https://doi.org/10.1073/pnas.1010732107

McCarter JD, Stephen Withers G (1994) Mechanisms of enzymatic glycoside hydrolysis. Curr Opin Struct Biol 4:885–892 . https://doi.org/10.1016/0959-440X(94)90271-2

McCleary BV, Blakeney A. (1999) Measurement of inulin and oligofructan. Cereal Foods World 44:398–406

McCleary B V., Gibson TS, Mugford DC (1997) Measurement of Total Starch in Cereal Products by Amyloglucosidase-α-Amylase Method: Collaborative Study. J AOAC Int 80:571–579 . https://doi.org/10.1093/jaoac/80.3.571

McCleary B V (1981) An Enzymic Technique for Jhe Quantitation of Galactomannan in Guar Seeds. Leb und-Technologie 14:56–59

McCleary B V, Charnock SJ, Rossiter PC, O'Shea MF, Power AM, Lloyd RM (2007) Measurement of carbohydrates in grain, feed and food. J Sci Food Agric 1243:1237–1243 . https://doi.org/10.1002/jsfa

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. J Appl Crystallogr 40:658–674 . https://doi.org/10.1107/S0021889807021206

Meents A, Reime B, Stuebe N, Fischer P, Warmer M, Goeries D, Roever J, Meyer J, Fischer J, Burkhardt A, Vartiainen I, Karvinen P, David C (2013) Development of an in-vacuum X-ray microscope with cryogenic sample cooling for beamline P11 at PETRA III. X-Ray Nanoimaging: Instruments and Methods 8851:1–7 . https://doi.org/88510k\r10.1117/12.2027303

Melcher RLJ, Neumann M, Fuenzalida Werner JP, Gröhn F, Moerschbacher BM (2017) Revised domain structure of ulvan lyase and characterization of the first ulvan binding domain. Sci Rep 7:44115 . https://doi.org/10.1038/srep44115

Ménesguen A, Piriou JY (1995) Nitrogen loadings and macroalgal (*Ulva* sp.) mass accumulation in brittany (France). Ophelia 42:227–237 . https://doi.org/10.1080/00785326.1995.10431506

Mewis K, Lenfant N, Lombard V, Henrissat B (2016) Dividing the large glycoside hydrolase family 43 into subfamilies: A motivation for detailed enzyme characterization. Appl Environ Microbiol 82:1686–1692 . https://doi.org/10.1128/AEM.03453-15

Mitulla M, Dinasquet J, Guillemette R, Simon M, Azam F, Wietz M (2016) Response of bacterial communities from California coastal waters to alginate particles and an alginolytic *Alteromonas macleodii* strain. Environ Microbiol 18:4369–4377 . https://doi.org/10.1111/1462-2920.13314

Mopper K, Kieber DJ (2002) Photochemistry and the Cycling of Carbon, Sulfur, Nitrogen and Phosphorus. Biogeochem Mar Dissolved Org Matter 455–507 . https://doi.org/10.1016/b978-012323841-2/50011-7

Mulloy B, Ribeiro AC, Alves AP, Vieira RP, Mourão PAS (1994) Sulfated fucans from echinoderms have a regular tetrasaccharide repeating unit defined by specific patterns of sulfation at the 0-2 and 0-4 positions. J Biol Chem 269:22113–23

Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr Sect D Biol Crystallogr 67:355–367 . https://doi.org/10.1107/S0907444911001314

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr Sect D Biol Crystallogr 53:240–255 . https://doi.org/10.1107/S0907444996012255

Mystkowska AA, Robb C, Vidal-Melgosa S, Vanni C, Fernandez-Guerra A, Höhne M, Hehemann J-H (2018) Molecular recognition of the beta-glucans laminarin and pustulan by a SusD-like glycan-binding protein of a marine Bacteroidetes. FEBS J 285:4465–4481 . https://doi.org/10.1111/febs.14674

Nagae M, Tsuchiya A, Katayama T, Yamamoto K, Wakatsuki S, Kato R (2007) Structural Basis of the Catalytic Reaction Mechanism of Novel 1,2-α-L-Fucosidase from *Bifidobacterium bifidum*. J Biol Chem 282:18497–18509 . https://doi.org/10.1074/jbc.M702246200

Nagao T, Kumabe A, Komatsu F, Yagi H, Suzuki H, Ohshiro T (2017) Gene identification and characterization of fucoidan deacetylase for potential application to fucoidan degradation and diversification. J Biosci Bioeng 124:277–282 . https://doi.org/10.1016/j.jbiosc.2017.04.002

Nagaoka M, Shibata H, Kimura-Takagi I, Hashimoto S, Kimura K, Makino T, Aiyama R, Ueyama S, Yokokura T (1999) Structural study of fucoidan from Cladosiphon okamuranus TOKIDA. Glycoconj J 16:19–26 . https://doi.org/10.1023/a:1006945618657

Nakamura T, Yamada KD, Tomii K, Katoh K (2018) Sequence analysis Parallelization of MAFFT for large-scale multiple sequence alignments. 34:2490–2492 . https://doi.org/10.1093/bioinformatics/bty121

Ndeh D et al. (2017) Complex pectin metabolism by gut bacteria reveals novel catalytic functions. Nature 544:65–70 . https://doi.org/10.1038/nature21725

Ndeh D, Baslé A, Strahl H, Yates EA, McClurgg UL, Henrissat B, Terrapon N, Cartmell A (2020) Metabolism of multiple glycosaminoglycans by *Bacteroides thetaiotaomicron* is orchestrated by a versatile core genetic locus. Nat Commun 11:646 . https://doi.org/10.1038/s41467-020-14509-4

Nedashkovskaya OI, Kim SB, Vancanneyt M, Snauwaert C, Lysenko AM, Rohde M, Frolova GM, Zhukova N V., Mikhailov V V., Bae KS, Oh HW, Swings J (2006) *Formosa agariphila* sp. nov., a budding bacterium of the family Flavobacteriaceae isolated from marine environments, and emended description of the genus Formosa. Int J Syst Evol Microbiol 56:161–167 . https://doi.org/10.1099/ijs.0.63875-0

Neumann AM, Balmonte JP, Berger M, Giebel H, Arnosti C, Voget S, Simon M, Brinkhoff T, Wietz M (2015) Different utilization of alginate and other algal polysaccharides by marine *Alteromonas macleodii* ecotypes. Environ Microbiol 17:3857–3868 . https://doi.org/10.1111/1462-2920.12862

Nguyen TT, Mikkelsen MD, Tran VHN, Trang VTD, Rhein-Knudsen N, Holck J, Rasin AB, Cao HTT, Van TTT, Meyer AS (2020) Enzyme-Assisted Fucoidan Extraction from Brown Macroalgae *Fucus distichus* subsp. *evanescens* and *Saccharina latissima*. Mar Drugs 18:296 . https://doi.org/10.3390/md18060296

Nielsen H (2017) Predicting Secretory Proteins with SignalP. In: Kihara D (ed) Protein Function Prediction: Methods and Protocols. Springer New York, New York, NY, pp 59–73

Nishino T, Nishioka C, Ura H, Nagumo T (1994) Isolation and partial characterization of a novel amino sugar-containing fucan sulfate from commercial *Fucus vesiculosus* fucoidan. Carbohydr Res 255:213–224

Otto A, Bernhardt J, Meyer H, Schaffer M, Herbst F-A, Siebourg J, Mäder U, Lalk M, Hecker M, Becher D (2010) Systems-wide temporal proteomic profiling in glucose-starved Bacillus subtilis. Nat Commun 1:137 . https://doi.org/10.1038/ncomms1137

Panagiotopoulos C, Sempéré R (2005) Analytical methods for the determination of sugars in marine samples: A historical perspective and future directions. Limnol Oceanogr Methods 3:419–454 . https://doi.org/10.4319/lom.2005.3.419

Passow U (2002) Transparent exopolymer particles (TEP) in aquatic environments. Prog Oceanogr 55:287–333 . https://doi.org/10.1016/S0079-6611(02)00138-6

Passow U (2000) Formation of transparent exopolymer particles, TEP, from dissolved precursor material. Mar Ecol Prog Ser 192:1–11 . https://doi.org/10.3354/meps192001

Passow U, Alldredge AL, Logan BE (1994) The role of particulate carbohydrate exudates in the flocculation of diatom blooms. Deep Sea Res 41:335–357

PDB Statistics PDB Statistics: Growth of Structures from X-ray Crystallography Experiments Released per Yea. In: Protein Data Bank. https://www.rcsb.org/stats/growth/growth-xray. Accessed 23 Jun 2020

Pflugrath JW (2015) Practical macromolecular cryocrystallography. Acta Crystallogr Sect F Struct Biol Commun 71:622–642 . https://doi.org/10.1107/S2053230X15008304

Popper ZA, Michel G, Herve C, Domozych DS, Willats WG, Tuohy MG, Kloareg B, Stengel DB (2011) Evolution and diversity of plant cell walls: from algae to flowering plants. Annu Rev Plant Biol 62:567–590 . https://doi.org/10.1146/annurev-arplant-042110-103809

Porcal P, Dillon PJ, Molot LA (2015) Temperature dependence of photodegradation of dissolved organic matter to dissolved inorganic carbon and particulate organic carbon. PLoS One 10:1–15 . https://doi.org/10.1371/journal.pone.0128884

Poretsky RS, Sun S, Mou X, Moran MA (2010) Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon. Environ Microbiol 12:616–627 . https://doi.org/10.1111/j.1462-2920.2009.02102.x

Rabillé H, Torode TA, Tesson B, Le Bail A, Billoud B, Rolland E, Le Panse S, Jam M, Charrier B (2019) Alginates along the filament of the brown alga Ectocarpus help cells cope with stress. Sci Rep 9:1–17 . https://doi.org/10.1038/s41598-019-49427-z

Rahikainen JL, Evans JD, Mikander S, Kalliola A, Puranen T, Tamminen T, Marjamaa K, Kruus K (2013) Cellulase-lignin interactions-The role of carbohydrate-binding module and pH in non-productive binding. Enzyme Microb Technol 53:315–321 . https://doi.org/10.1016/j.enzmictec.2013.07.003

Ramos Santos C, Polo CC, Corrêa JM, Simão RDCG, Seixas FAV, Murakami MT (2012) The accessory domain changes the accessibility and molecular topography of the catalytic interface in monomeric GH39 β-xylosidases. Acta Crystallogr Sect D Biol Crystallogr 68:1339–1345 . https://doi.org/10.1107/S0907444912028491

Rehm BHA, Valla S (1997) Bacterial alginates: Biosynthesis and applications. Appl Microbiol Biotechnol 48:281–288 . https://doi.org/10.1007/s002530051051

Reintjes G, Arnosti C, Fuchs B, Amann R (2018) Selfish, sharing and scavenging bacteria in the Atlantic Ocean: a biogeographical study of bacterial substrate utilisation. ISME J 13:1119–1132 . https://doi.org/10.1038/s41396-018-0326-3

Reintjes G, Arnosti C, Fuchs BM, Amann R (2017) An alternative polysaccharide uptake mechanism of marine bacteria. ISME J 11:1640–1650 . https://doi.org/10.1038/ismej.2017.26

Reisky L et al. (2019) A marine bacterial enzymatic cascade degrades the algal polysaccharide ulvan. Nat Chem Biol 15:803–812 . https://doi.org/10.1038/s41589-019-0311-9

Reisky L (2019) Discovery and Characterization of Novel Carbohydrate-Active Enzymes from Marine Bacteria

Reisky L, Stanetty C, Mihovilovic MD, Schweder T, Hehemann J-H, Bornscheuer UT (2018) Biochemical characterization of an ulvan lyase from the marine flavobacterium *Formosa agariphila* KMM 3901_T. Appl Microbiol Biotechnol 102:6987–6996 . https://doi.org/10.1007/s00253-018-9142-y

Ribeiro A-C, Vieira RP, Mourão PAS, Mulloy B (1994) A sulfated α-L-fucan from sea cucumber. Carbohydr Res 255:225–240 . https://doi.org/10.1016/S0008-6215(00)90981-9

Richardson JS, Carpena X, Switala J, Perez-Luque R, Donald LJ, Loewen PC, Oresnik IJ (2008) RhaU of Rhizobium leguminosarum Is a Rhamnose Mutarotase. J Bacteriol 190:2903–2910 . https://doi.org/10.1128/JB.01120-07

Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res 42:W320–W324 . https://doi.org/10.1093/nar/gku316

Robic A, Gaillard C, Sassi JF, Leral Y, Lahaye M (2009) Ultrastructure of Ulvan: A polysaccharide

from green seaweeds. Biopolymers 91:652–664 . https://doi.org/10.1002/bip.21195

Rogowski A et al. (2015) Glycan complexity dictates microbial resource allocation in the large intestine. Nat Commun 6:7481 . https://doi.org/10.1038/ncomms8481

Ryu K-S, Kim J-I, Cho S-J, Park D, Park C, Cheong H-K, Lee J-O, Choi B-S (2005) Structural Insights into the Monosaccharide Specificity of Escherichia coli Rhamnose Mutarotase. J Mol Biol 349:153–162 . https://doi.org/10.1016/j.jmb.2005.03.047

Sabra W, Zeng AP, Deckwer WD (2001) Bacterial alginate: Physiology, product quality and process aspects. Appl Microbiol Biotechnol 56:315–325 . https://doi.org/10.1007/s002530100699

Sakai T, Ishizuka K, Kato I (2003) Isolation and Characterization of a Fucoidan-Degrading Marine Bacterium. Mar Biotechnol 5:409–416 . https://doi.org/10.1007/s10126-002-0118-6

Sakurama H, Fushinobu S, Hidaka M, Yoshida E, Honda Y, Ashida H, Kitaoka M, Kumagai H, Yamamoto K, Katayama T (2012a) 1,3-1,4-α-L-Fucosynthase that specifically introduces Lewis a/x antigens into type-1/2 chains. J Biol Chem 287:16709–16719 . https://doi.org/10.1074/jbc.M111.333781

Sakurama H, Tsutsumi E, Ashida H, Katayama T, Yamamoto K, Kumagai H (2012b) Differences in the substrate specificities and active-site structures of two α-L-fucosidases (glycoside hydrolase family 29) from *Bacteroides thetaiotaomicron*. Biosci Biotechnol Biochem 76:1022–4 . https://doi.org/10.1271/bbb.111004

Salinas A, French CE (2017) The enzymatic ulvan depolymerisation system from the alga-associated marine flavobacterium *Formosa agariphila*. Algal Res 27:335–344 . https://doi.org/10.1016/j.algal.2017.09.025

Sato K, Naito M, Yukitake H, Hirakawa H, Shoji M, McBride MJ, Rhodes RG, Nakayama K (2010) A protein secretion system linked to bacteroidete gliding motility and pathogenesis. Proc Natl Acad Sci U S A 107:276–281 . https://doi.org/10.1073/pnas.0912010107

Scheschonk L, Becker S, Hehemann JH, Diehl N, Karsten U, Bischof K (2019) Arctic kelp eco-physiology during the polar night in the face of global warming: A crucial role for laminarin. Mar Ecol Prog Ser 611:59–74 . https://doi.org/10.3354/meps12860

Schmidt B, Selmer T, Ingendoh A, Figurat K von (1995) A novel amino acid modification in sulfatases that is defective in multiple sulfatase deficiency. Cell 82:271–278 . https://doi.org/10.1016/0092-8674(95)90314-3

Schmitt C, Sanchez C, Desobry-Banon S, Hardy J (1998) Structure and Technofunctional Properties of Protein-Polysaccharide Complexes: A Review. Crit Rev Food Sci Nutr 38:689–753 . https://doi.org/10.1080/10408699891274354

Schultz-Johansen M, Cueff M, Hardouin K, Jam M, Larocque R, Glaring MA, Hervé C, Czjzek M, Stougaard P (2018) Discovery and screening of novel metagenome-derived GH107 enzymes targeting sulfated fucans from brown algae. FEBS J 285:4281–4295 . https://doi.org/10.1111/febs.14662

Schut F, de Vries EJ, Gottschal JC, Robertson BR, Harder W, Prins RA, Button DK (1993) Isolation of Typical Marine Bacteria by Dilution Culture: Growth, Maintenance, and Characteristics of Isolates under Laboratory Conditions. Appl Environ Microbiol 59:2150–2160 . https://doi.org/10.1128/AEM.59.7.2150-2160.1993

Sela DA, Garrido D, Lerno L, Wu S, Tan K, Eom H-J, Joachimiak A, Lebrilla CB, Mills DA (2012) *Bifidobacterium longum* subsp. infantis ATCC 15697 α-Fucosidases Are Active on Fucosylated Human Milk Oligosaccharides. Appl Environ Microbiol 78:795–803 . https://doi.org/10.1128/AEM.06762-11

Shen Y, Benner R (2018) Mixing it up in the ocean carbon cycle and the removal of refractory dissolved organic carbon. Sci Rep 8:1–9 . https://doi.org/10.1038/s41598-018-20857-5

Shepherd VA, Beilby MJ, Heslop DJ (1999) Ecophysiology of the hypotonic response in the salt-tolerant charophyte alga Lamprothamnium papulosum. Plant, Cell Environ 22:333–346 . https://doi.org/10.1046/j.1365-3040.1999.00414.x

Shevchenko NM, Anastyuk SD, Menshova R V., Vishchuk OS, Isakov VI, Zadorozhny PA, Sikorskaya T V., Zvyagintseva TN (2015) Further studies on structure of fucoidan from brown alga Saccharina gurjanovae. Carbohydr Polym 121:207–216 . https://doi.org/10.1016/j.carbpol.2014.12.042

Shoseyov O, Shani Z, Levy I (2006) Carbohydrate Binding Modules: Biochemical Properties and Novel Applications. Microbiol Mol Biol Rev 70:283–295 . https://doi.org/10.1128/MMBR.00028-05

Sichert A, Corzett CH, Schechter MS, Unfried F, Markert S, Becher D, Fernandez-Guerra A, Liebeke M, Schweder T, Polz MF, Hehemann J (2020) Verrucomicrobia use hundreds of enzymes to digest the algal polysaccharide fucoidan. Nat Microbiol 5:1026–1039 . https://doi.org/10.1038/s41564-020-0720-2

Silchenko A, Kusaykin M, Kurilenko V, Zakharenko A, Isakov V, Zaporozhets T, Gazha A, Zvyagintseva T (2013) Hydrolysis of Fucoidan by Fucoidanase Isolated from the Marine Bacterium, *Formosa algae*. Mar Drugs 11:2413–2430 . https://doi.org/10.3390/md11072413

Silchenko AS, Kusaykin MI, Zakharenko AM, Menshova R V., Khanh HHN, Dmitrenok PS, Isakov V V., Zvyagintseva TN (2014) Endo-1,4-fucoidanase from Vietnamese marine mollusk *Lambis* sp. which producing sulphated fucooligosaccharides. J Mol Catal B Enzym 102:154–160 . https://doi.org/10.1016/j.molcatb.2014.02.007

Silchenko AS, Rasin AB, Kusaykin MI, Kalinovsky AI, Miansong Z, Changheng L, Malyarenko O, Zueva AO, Zvyagintseva TN, Ermakova SP (2017) Structure, enzymatic transformation, anticancer activity of fucoidan and sulphated fucooligosaccharides from *Sargassum horneri*. Carbohydr Polym 175:654–660 . https://doi.org/10.1016/j.carbpol.2017.08.043

Silchenko AS, Rasin AB, Zueva AO, Kusaykin MI, Zvyagintseva TN, Kalinovsky AI, Kurilenko V V., Ermakova SP (2018) Fucoidan Sulfatases from Marine Bacterium *Wenyingzhuangia fucanilytica* CZ1127T. Biomolecules 8:98 . https://doi.org/10.3390/biom8040098

Silchenko AS, Ustyuzhanina NE, Kusaykin MI, Krylov VB, Shashkov AS, Dmitrenok AS, Usoltseva R V, Zueva AO, Nifantiev NE, Zvyagintseva TN (2016) Expression and biochemical characterization and substrate specificity of the fucoidanase from *Formosa algae*. Glycobiology 27:254–263 . https://doi.org/10.1093/glycob/cww138

Sim JS, Im AR, Cho SM, Jang HJ, Jo JH, Kim YS (2007) Evaluation of chondroitin sulfate in shark cartilage powder as a dietary supplement: Raw materials and finished products. Food Chem 101:532–539 . https://doi.org/10.1016/j.foodchem.2006.02.011

Simpkin AJ, Thomas JMH, Simkovic F, Keegan RM, Rigden DJ (2019) Molecular replacement using structure predictions from databases. Acta Crystallogr Sect D Struct Biol 75:1051–1062 . https://doi.org/10.1107/S2059798319013962

Skriptsova A V. (2015) Fucoidans of brown algae: Biosynthesis, localization, and physiological role in thallus. Russ J Mar Biol 41:145–156 . https://doi.org/10.1134/S1063074015030098

Skriptsova A V., Shevchenko NM, Zvyagintseva TN, Imbs TI (2010) Monthly changes in the content and monosaccharide composition of fucoidan from *Undaria pinnatifida* (Laminariales, Phaeophyta). J Appl Phycol 22:79–86 . https://doi.org/10.1007/s10811-009-9438-5

Smart OS, Womack TO, Flensburg C, Keller P, Paciorek W, Sharff A, Vonrhein C, Bricogne G (2012) Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. Acta Crystallogr Sect D Biol Crystallogr 68:368–380 . https://doi.org/10.1107/S0907444911056058

Smetacek V, Zingone A (2013) Green and golden seaweed tides on the rise. Nature 504:84–88 . https://doi.org/10.1038/nature12860

Smyth MS (2000) x Ray crystallography. Mol Pathol 53:8–14 . https://doi.org/10.1136/mp.53.1.8

St John FJ, González JM, Pozharski E (2010) Consolidation of glycosyl hydrolase family 30: A dual domain 4/7 hydrolase family consisting of two structurally distinct groups. FEBS Lett 584:4435–4441 . https://doi.org/10.1016/j.febslet.2010.09.051

Stam MR, Danchin EGJ, Rancurel C, Coutinho PM, Henrissat B (2006) Dividing the large glycoside hydrolase family 13 into subfamilies: Towards improved functional annotations of α-amylase-related proteins. Protein Eng Des Sel 19:555–562 . https://doi.org/10.1093/protein/gzl044

Stewart PDS, Kolek SA, Briggs RA, Chayen NE, Baldock PFM (2011) Random Microseeding : A Theoretical and Practical Exploration of Seed Stability and Seeding Techniques for Successful Protein Crystallization. 3432–3441

Studier FW (2005) Protein production by auto-induction in high-density shaking cultures. Protein Expr Purif 41:207–234 . https://doi.org/10.1016/j.pep.2005.01.016

Stura EA, Wilson IA (1991) Applications of the streak seeding technique in protein crystallization. J Cryst Growth 110:270–282 . https://doi.org/10.1016/0022-0248(91)90896-D

Sulzenbacher G, Bignon C, Nishimura T, Tarling CA, Withers SG, Henrissat B, Bourne Y (2004) Crystal structure of Thermotoga maritima α-L-fucosidase: Insights into the catalytic mechanism and the molecular basis for fucosidosis. J Biol Chem 279:13119–13128 . https://doi.org/10.1074/jbc.M313783200

Summers EL, Moon CD, Atua R, Arcus VL (2016) The structure of a glycoside hydrolase 29 family member from a rumen bacterium reveals unique, dual carbohydrate-binding domains. Acta Crystallogr Sect F Struct Biol Commun 72:750–761 . https://doi.org/10.1107/S2053230X16014072

Suttle CA, Chan AM, Cottrell MT (1990) Infection of phytoplankton by viruses and reduction of primary productivity. Nature 34:467–469 . https://doi.org/https://doi.org/10.1038/347467a0

Synytsya A, Kim W-J, Kim S-M, Pohl R, Synytsya A, Kvasnička F, Čopíková J, Il Park Y (2010) Structure and antitumour activity of fucoidan isolated from sporophyll of Korean brown seaweed *Undaria pinnatifida*. Carbohydr Polym 81:41–48 . https://doi.org/10.1016/j.carbpol.2010.01.052

Tanaka K, Sorai S (1970) Hydrolysis of fucoidan by abalone liver α-L-fucosidase. FEBS Lett 9:45–48 . https://doi.org/10.1016/0014-5793(70)80307-6

Tarling CA, He S, Sulzenbacher G, Bignon C, Bourne Y, Henrissat B, Withers SG (2003) Identification of the Catalytic Nucleophile of the Family 29 α-L-Fucosidase from *Thermotoga maritima* through Trapping of a Covalent Glycosyl-Enzyme Intermediate and Mutagenesis. J Biol Chem 278:47394–47399 . https://doi.org/10.1074/jbc.M306610200

Teeling H et al. (2012) Substrate-Controlled Succession of Marine Bacterioplankton Populations Induced by a Phytoplankton Bloom. Science (80-) 336:608–611 . https://doi.org/10.1126/science.1218344

Teeling H, Fuchs BM, Bennke CM, Krüger K, Chafee M, Kappelmann L, Reintjes G, Waldmann

J, Quast C, Glöckner FO, Lucas J, Wichels A, Gerdts G, Wiltshire KH, Amann RI (2016) Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. Elife 5:447–450 . https://doi.org/10.7554/eLife.11888

Thirumurugan G, Dhanaraju MD (2017) Marine Polysaccharides as Multifunctional Pharmaceutical Excipients. In: Shalaby E (ed) Biological Activities and Application of Marine Polysaccharides. Intech, p 129pp

Thomas F, Barbeyron T, Tonon T, Génicot S, Czjzek M, Michel G (2012) Characterization of the first alginolytic operons in a marine bacterium: from their emergence in marine Flavobacteriia to their independent transfers to marine Proteobacteria and human gut Bacteroides. Environ Microbiol 14:2379–2394 . https://doi.org/10.1111/j.1462-2920.2012.02751.x

Thomas F, Lundqvist LCE, Jam M, Jeudy A, Barbeyron T, Sandström C, Michel G, 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. J Biol Chem 288:23021–23037 . https://doi.org/10.1074/jbc.M113.467217

Thornton DCO (2009) Spatiotemporal distribution of dissolved acid polysaccharides (dAPS) in a tidal estuary. Limnol Oceanogr 54:1449–1460 . https://doi.org/10.4319/lo.2009.54.5.1449

Thrash JC, Cho J-C, Vergin KL, Morris RM, Giovannoni SJ (2010) Genome Sequence of *Lentisphaera araneosa* HTCC2155T, the Type Species of the Order Lentisphaerales in the Phylum Lentisphaerae. J Bacteriol 192:2938–2939 . https://doi.org/10.1128/JB.00208-10

Tohru Y, Shintaro O, Tomochika K, Koji I, Hyakuji Y (1982) Galactose 6-sulfate sulfatase activity in Morquio syndrome. Clin Chim Acta 122:169–180 . https://doi.org/10.1016/0009-8981(82)90276-5

Tomme P, Warren RAJ, Miller RC, Kilburn DG, Gilkes NR (1996) Cellulose-Binding Domains : Classification and Properties. Am Chem Soc Chapter 10:142–163

Torode TA, Siméon A, Marcus SE, Jam M, Le Moigne M-A, Duffieux D, Knox JP, Hervé C (2016) Dynamics of cell wall assembly during early embryogenesis in the brown alga *Fucus*. J Exp Bot 67:6089–6100 . https://doi.org/10.1093/jxb/erw369

Turgeon S (2003) Protein–polysaccharide interactions: phase-ordering kinetics, thermodynamic and structural aspects. Curr Opin Colloid Interface Sci 8:401–414 . https://doi.org/10.1016/j.cocis.2003.08.013

Turgeon SL, Schmitt C, Sanchez C (2007) Protein–polysaccharide complexes and coacervates. Curr Opin Colloid Interface Sci 12:166–178 . https://doi.org/10.1016/j.cocis.2007.07.007

Tyanova S, Cox J (2018) Perseus: A Bioinformatics Platform for Integrative Analysis of Proteomics Data in Cancer Research. In: Methods in Molecular Biology. pp 133–148

Ulaganathan T, Banin E, Helbert W, Cygler M (2018a) Structural and functional characterization of PL28 family ulvan lyase NLR48 from *Nonlabens ulvanivorans*. J Biol Chem 293:11564–11573 . https://doi.org/10.1074/jbc.RA118.003659

Ulaganathan T, Boniecki MT, Foran E, Buravenkov V, Mizrachi N, Banin E, Helbert W, Cygler M (2017) New ulvan-degrading polysaccharide lyase family: structure and catalytic mechanism suggests convergent evolution of active site architecture. ACS Chem Biol acschembio.7b00126 . https://doi.org/10.1021/acschembio.7b00126

Ulaganathan T, Helbert W, Kopel M, Banin E, Cygler M (2018b) Structure–function analyses of a PL24 family ulvan lyase reveal key features and suggest its catalytic mechanism. J Biol Chem 293:4026–4036 . https://doi.org/10.1074/jbc.RA117.001642

Underwood GJC, Fietz S, Papadimitriou S, Thomas DN, Dieckmann GS (2010) Distribution and

composition of dissolved extracellular polymeric substances (EPS) in Antarctic sea ice. Mar Ecol Prog Ser 404:1–19 . https://doi.org/10.3354/meps08557

Unfried F et al. (2018) Adaptive mechanisms that provide competitive advantages to marine bacteroidetes during microalgal blooms. ISME J. https://doi.org/10.1038/s41396-018-0243-5

Ustyuzhanina NE, Bilan MI, Ushakova NA, Usov AI, Kiselevskiy M V., Nifantiev NE (2014) Fucoidans: Pro- or antiangiogenic agents? Glycobiology 24:1265–1274 . https://doi.org/10.1093/glycob/cwu063

van Vliet DM, Palakawong Na Ayudthaya S, Diop S, Villanueva L, Stams AJM, Sánchez-Andrea I (2019) Anaerobic Degradation of Sulfated Polysaccharides by Two Novel *Kiritimatiellales* Strains Isolated From Black Sea Sediment. Front Microbiol 10:1–16 . https://doi.org/10.3389/fmicb.2019.00253

Verdugo P, Alldredge AL, Azam F, Kirchman DL, Passow U, Santschi PH (2004) The oceanic gel phase: a bridge in the DOM–POM continuum. Mar Chem 92:67–85 . https://doi.org/10.1016/j.marchem.2004.06.017

Viborg AH, Terrapon N, Lombard V, Michel G, Czjzek M, Henrissat B, Brumer H (2019) A subfamily roadmap of the evolutionarily diverse glycoside hydrolase family 16 (GH16). J Biol Chem 294:15973–15986 . https://doi.org/10.1074/jbc.RA119.010619

Vickers C, Liu F, Abe K, Salama-Alber O, Jenkins ML, Springate CM, Burke JE, Withers SG, Boraston AB (2018) Endo-fucoidan hydrolases from glycoside hydrolase family 107 (GH107) display structural and mechanistic similarities to α-L-fucosidases from GH29. J Biol Chem 107:jbc.RA118.005134 . https://doi.org/10.1074/JBC.RA118.005134

Vidal-Melgosa S, Hehemann J-H (2019) Fucoidans from diatoms

Vidal-Melgosa S, Pedersen HL, Schückel J, Arnal G, Dumon C, Amby DB, Monrad RN, Westereng B, Willats WGT (2015) A New Versatile Microarray-based Method for High Throughput Screening of Carbohydrate-active Enzymes. J Biol Chem 290:9020–9036 . https://doi.org/10.1074/jbc.M114.630673

Vizcaíno JA, Csordas A, Del-Toro N, Dianes JA, Griss J, Lavidas I, Mayer G, Perez-Riverol Y, Reisinger F, Ternent T, Xu QW, Wang R, Hermjakob H (2016) 2016 update of the PRIDE database and its related tools. Nucleic Acids Res 44:D447–D456 . https://doi.org/10.1093/nar/gkv1145

Vo T-S, Kim S-K (2013) Fucoidans as a natural bioactive ingredient for functional foods. J Funct Foods 5:16–27 . https://doi.org/10.1016/j.jff.2012.08.007

Von Bülow R, Schmidt B, Dierks T, Von Figura K, Usón I (2001) Crystal structure of an enzyme-substrate complex provides insight into the interaction between human arylsulfatase A and its substrates during catalysis. J Mol Biol 305:269–277 . https://doi.org/10.1006/jmbi.2000.4297

Vreeburg RAM, Airianah OB, Fry SC (2014) Fingerprinting of hydroxyl radical-attacked polysaccharides by N-isopropyl-2-aminoacridone labelling. Biochem J 463:225–237 . https://doi.org/10.1042/BJ20140678

Wallmann K, Aloisi G (1992) The Global Carbon Cycle. Int Geophys 50:239–262 . https://doi.org/10.1016/S0074-6142(08)62694-7

Wang M, Hu C, Barnes BB, Mitchum G, Lapointe B, Montoya JP (2019a) The great Atlantic Sargassum belt. Science (80-) 364:83–87 . https://doi.org/10.1126/science.aaw7912

Wang S, Su T, Zhang Q, Guan J, He J, Gu L, Li F (2019b) Comparative study of two chondroitin sulfate/dermatan sulfate 4-O-sulfatases with high identity. Front Microbiol 10:1–17 . https://doi.org/10.3389/fmicb.2019.01309

Wargacki AJ, Leonard E, Win MN, Regitsky DD, Santos CNS, Kim PB, Cooper SR, Raisner RM, Herman A, Sivitz AB, Lakshmanaswamy A, Kashiyama Y (2012a) An Engineered Microbial Platform for Direct Biofuel Production from Brown Macroalgae. Science (80-) 335:308–314 . https://doi.org/10.1126/science.1214547

Wargacki AJ, Leonard E, Win MN, Regitsky DD, Santos CNS, Kim PB, Cooper SR, Raisner RM, Herman A, Sivitz AB, Lakshmanaswamy A, Kashiyama Y, Baker D, Yoshikuni Y (2012b) An Engineered Microbial Platform for Direct Biofuel Production from Brown Macroalgae. Science (80-) 335:308–313 . https://doi.org/10.1126/science.1214547

Wattam AR et al. (2017) Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. Nucleic Acids Res 45:D535–D542 . https://doi.org/10.1093/nar/gkw1017

Wegner C-E, Richter-Heitmann T, Klindworth A, Klockow C, Richter M, Achstetter T, Glöckner FO, Harder J (2013) Expression of sulfatases in *Rhodopirellula baltica* and the diversity of sulfatases in the genus *Rhodopirellula*. Mar Genomics 9:51–61 . https://doi.org/10.1016/j.margen.2012.12.001

Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18:691–699 . https://doi.org/10.1093/oxfordjournals.molbev.a003851

White Jr. WJ, Schray KJ, Legler G, Alhadeff JA (1987) Further studies on the catalytic mechanism of human liver α-l-fucosidase. Biochim Biophys Acta - Protein Struct Mol Enzymol 912:132–138 . https://doi.org/10.1016/0167-4838(87)90256-1

Wiegand S et al. (2020) Cultivation and functional characterization of 79 planctomycetes uncovers their unique biology. Nat Microbiol 5:126–140 . https://doi.org/10.1038/s41564-019-0588-1

Wietz M, Wemheuer B, Simon H, Giebel H, Seibt MA, Daniel R, Brinkhoff T, Simon M (2015) Bacterial community dynamics during polysaccharide degradation at contrasting sites in the Southern and Atlantic Oceans. Environ Microbiol 17:3822–3831 . https://doi.org/10.1111/1462-2920.12842

Wilkins MR, Gasteiger E, Bairoch A, Sanchez J-C, Williams KL, Appel RD, Hochstrasser DF (2005) Protein Identification and Analysis Tools in the ExPASy Server. In: 2-D Proteome Analysis Protocols, John M. Wa. Humana Press, New Jersey, pp 531–552

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr Sect D Biol Crystallogr 67:235–242 . https://doi.org/10.1107/S0907444910045749

Winter G, Waterman DG, Parkhurst JM, Brewster AS, Gildea RJ, Gerstel M, Fuentes-Montero L, Vollmar M, Michels-Clark T, Young ID, Sauter NK, Evans G (2018) DIALS: Implementation and evaluation of a new integration package. Acta Crystallogr Sect D Struct Biol 74:85–97 . https://doi.org/10.1107/S2059798317017235

Wong TY, Preston LA, Schiller NL (2000) Alginate Lyase: Review of Major Sources and Enzyme Characteristics, Structure-Function Analysis, Biological Roles, and Applications. Annu Rev Microbiol 54:289–340 . https://doi.org/10.1146/annurev.micro.54.1.289

Yabur R, Bashan Y, Hernández-Carmona G (2007) Alginate from the macroalgae *Sargassum sinicola* as a novel source for microbial immobilization material in wastewater treatment and plant growth promotion. J Appl Phycol 19:43–53 . https://doi.org/10.1007/s10811-006-9109-8

Yang JK, Yoon H-J, Ahn HJ, Il Lee B, Pedelacq J-D, Liong EC, Berendzen J, Laivenieks M, Vieille C, Zeikus GJ, Vocadlo DJ, Withers SG, Suh SW (2004) Crystal Structure of β-d-Xylosidase from *Thermoanaerobacterium saccharolyticum*, a Family 39 Glycoside Hydrolase. J Mol Biol 335:155–165 . https://doi.org/10.1016/j.jmb.2003.10.026

Yip VLY, Withers SG (2006) Breakdown of oligosaccharides by the process of elimination. Curr Opin Chem Biol 10:147–155 . https://doi.org/10.1016/j.cbpa.2006.02.005

Yip VLY, Withers SG (2004) Nature's many mechanisms for the degradation of oligosaccharides. Org Biomol Chem 2:2707–2713 . https://doi.org/10.1039/b408880h

Yu L, Ge L, Xue C, Chang Y, Zhang C, Xu X, Wang Y (2014a) Structural study of fucoidan from sea cucumber *Acaudina molpadioides*: A fucoidan containing novel tetrafucose repeating unit. Food Chem 142:197–200 . https://doi.org/10.1016/j.foodchem.2013.06.079

Yu L, Xue C, Chang Y, Xu X, Ge L, Liu G, Wang Y (2014b) Structure elucidation of fucoidan composed of a novel tetrafucose repeating unit from sea cucumber *Thelenota ananas*. Food Chem 146:113–119 . https://doi.org/10.1016/j.foodchem.2013.09.033

Yu S, Su T, Wu H, Liu S, Wang D, Zhao T, Jin Z, Du W, Zhu MJ, Chua SL, Yang L, Zhu D, Gu L, Ma LZ (2015) PslG, a self-produced glycosyl hydrolase, triggers biofilm disassembly by disrupting exopolysaccharide matrix. Cell Res 25:1352–1367 . https://doi.org/10.1038/cr.2015.129

Zeuner B, Muschiol J, Holck J, Lezyk M, Gedde MR, Jers C, Mikkelsen JD, Meyer AS (2018) Substrate specificity and transfucosylation activity of GH29 α-L-fucosidases for enzymatic production of human milk oligosaccharides. N Biotechnol 41:34–45 . https://doi.org/10.1016/j.nbt.2017.12.002

Zeuner B, Teze D, Muschiol J, Meyer AS (2019) Synthesis of Human Milk Oligosaccharides: Protein Engineering Strategies for Improved Enzymatic Transglycosylation. Molecules 24:2033 . https://doi.org/10.3390/molecules24112033

Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y (2018) DbCAN2: A meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res 46:W95–W101 . https://doi.org/10.1093/nar/gky418

Zhao Y, Zheng Y, Wang J, Ma S, Yu Y, White W, Yang S, Yang F, Lu J (2018) Fucoidan Extracted from *Undaria pinnatifida*: Source for Nutraceuticals/Functional Foods. Mar Drugs 16:321 . https://doi.org/10.3390/md16090321

Zhou H-X, Pang X (2018) Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. Chem Rev 118:1691–1741 . https://doi.org/10.1021/acs.chemrev.7b00305

Zybailov B, Mosley AL, Sardiu ME, Coleman MK, Florens L, Washburn MP (2006) Statistical Analysis of Membrane Proteome Expression Changes in Saccharomyces c erevisiae. J Proteome Res 5:2339–2347 . https://doi.org/10.1021/pr060161n

Appendix

Supplementary information – Discussion and future perspective

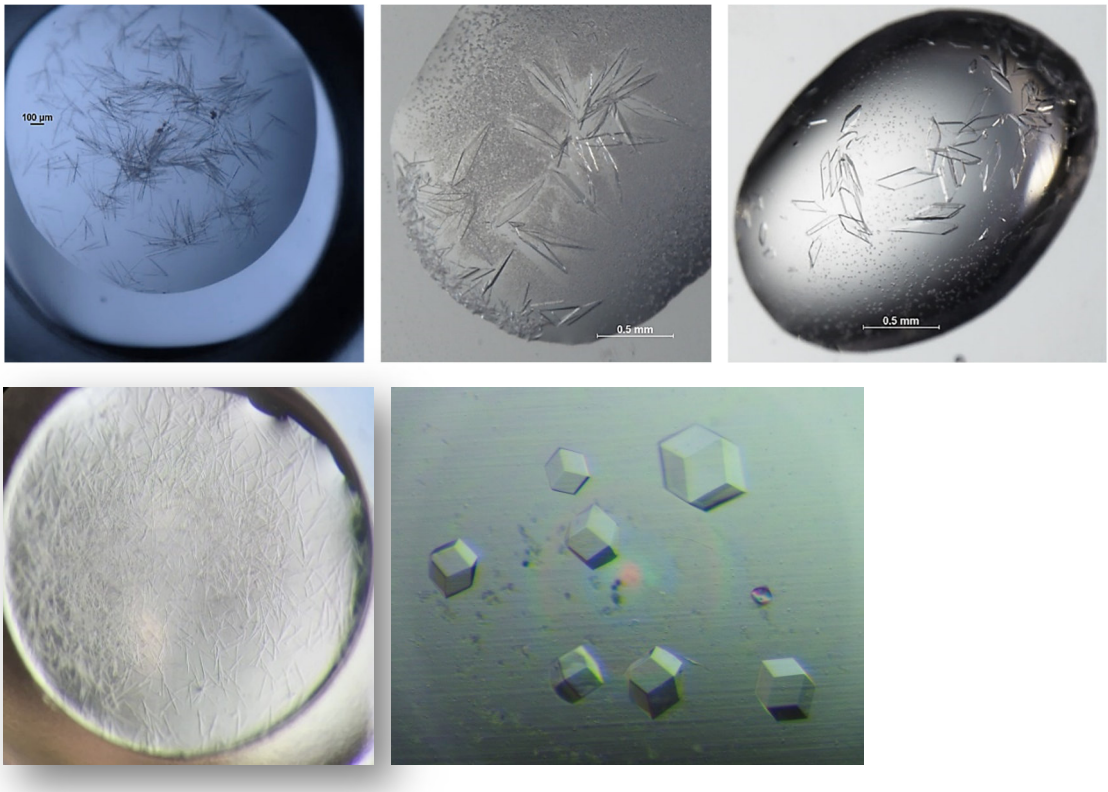


Figure S1. Protein crystals of ulvan lyase PL28 and GH39 from *F. agariphila* KMM3901^T. A-C: Optimization of crystal growth of 30_PL28. C+D: Variations in crystal morphology of 31_GH39.

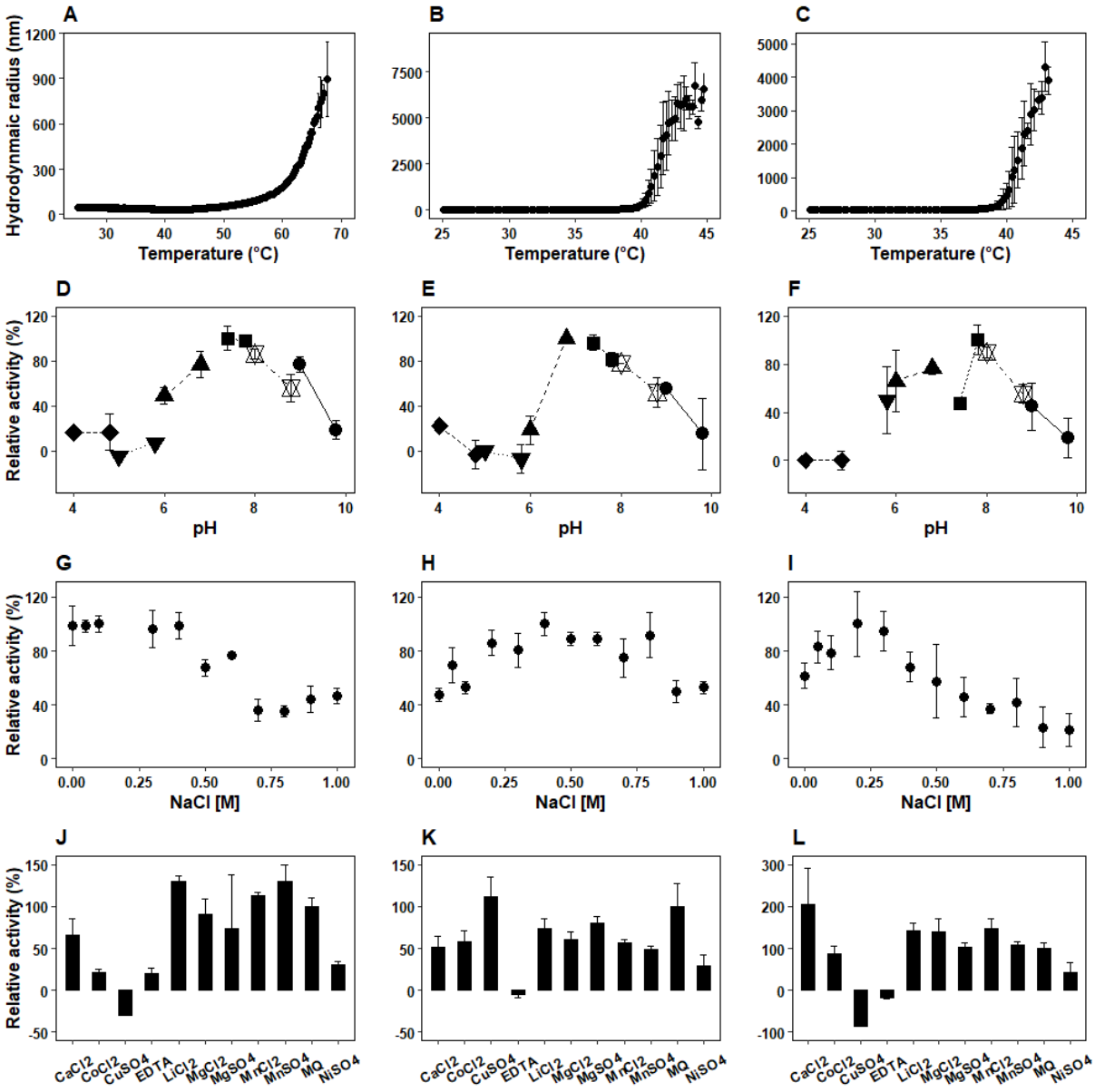


Figure S2. Biochemical spectra of Alys PL6-1, PL7-1CD and PL17 from *Alteromonas macleodii* 83-1. (A-C) Thermal stability. Thermal stability of the protein was indicated by a constant radius (nm). Increasing hydrodynamic radius indicates unfolding and aggregation of the protein. (D-F) Influence of pH on 0.1 % poly G using 0.1 M ♦:Na-acetate, ▼:Na-citrate, ▲:Na-cacodylate ■:HEPES, ◊:TRIS, ●:CAPSO, (G-I) Influence of sodium chloride concentration on 0.1 % poly G and 0.1 M HEPES pH 7.8 (J-L) Influence of cations on 0.1 % poly G, 0.1 M HEPES pH 7.8 and 0.2 M NaCl. Error bars indicate the standard deviation of replicates.

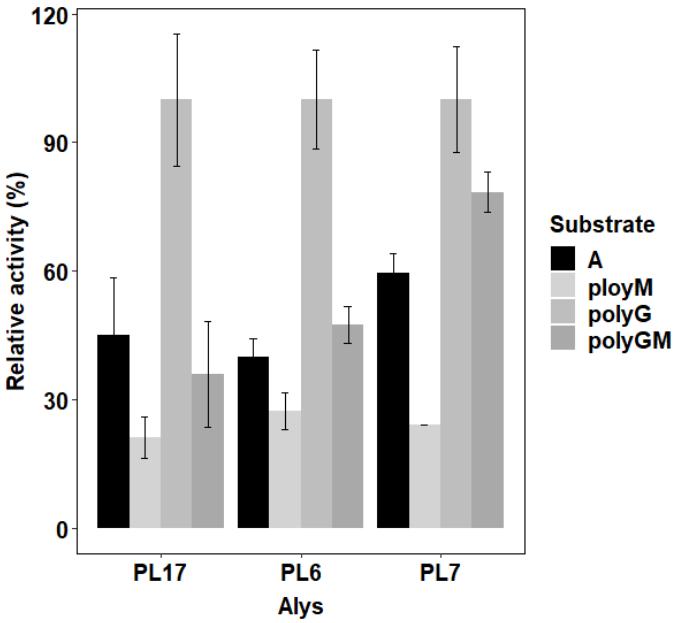


Figure S3. Alys from *A. macleodii* 83-1 are L-guluronate specific. Enzymatic activity was determined on 0.1% substrate in 250 mM NaCl, 20 mM TRIS pH 8 at 25 °C. Error bars indicate measurement of triplicates.

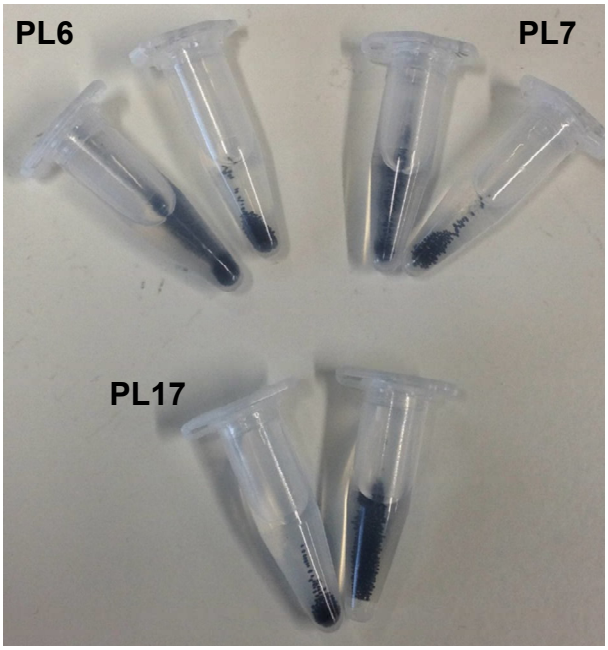


Figure S4. Degradation of alginate particles by Alys. Enzyme digestion at room temperature overnight (left): Endo-acting Alys PL6 and PL7 degraded alginate particles, whereas no activity was found for exo-acting PL17. Enzyme buffer was used as negative control (right) to demonstrate particle stability.

Supplementary information – Chapter 1.2.

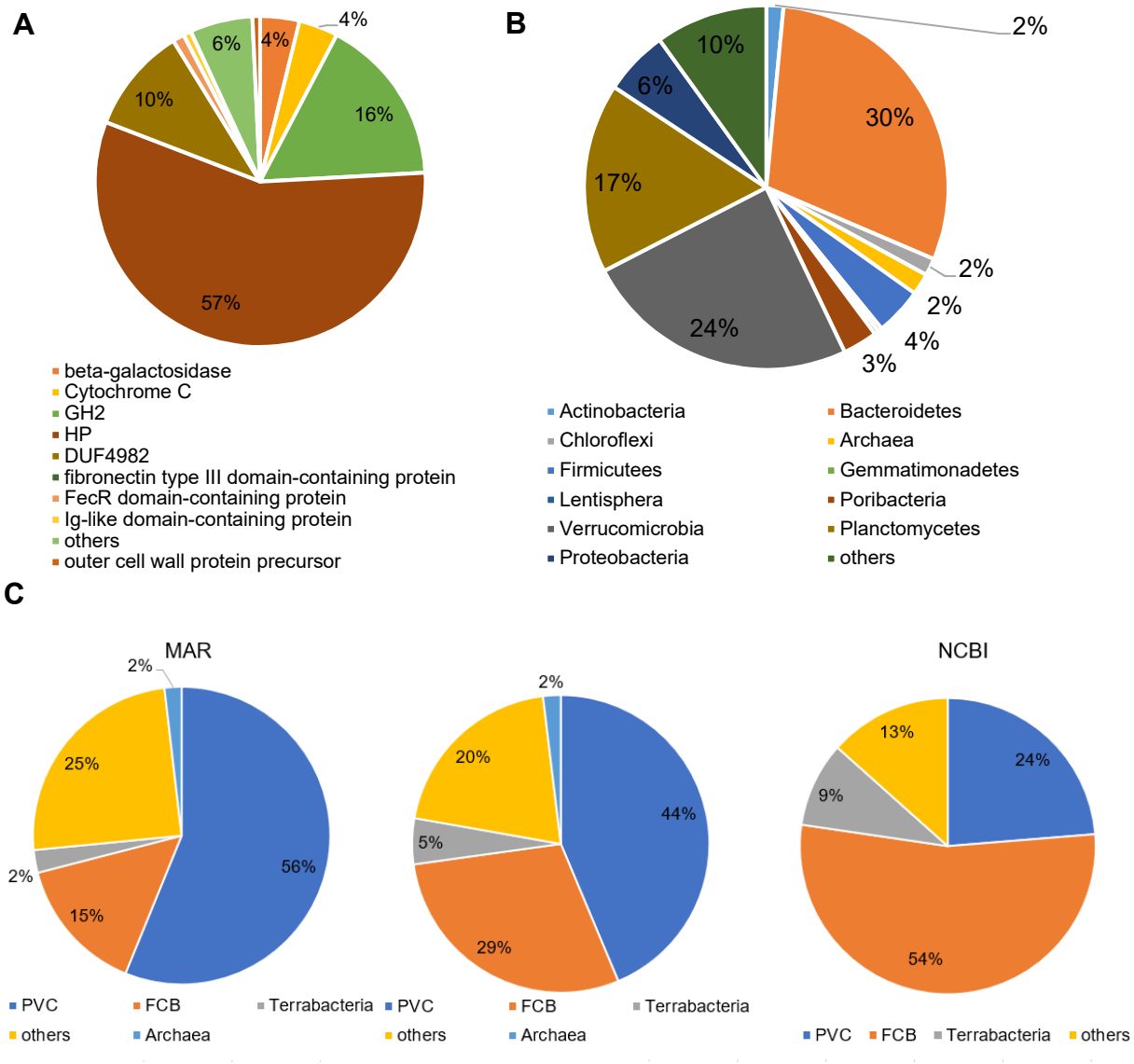


Figure S1. (A) Annotation and relative abundance in (B) phyla. (C) Comparison between marine metagenomes (left) and NCBI (right) with total abundance (middle).

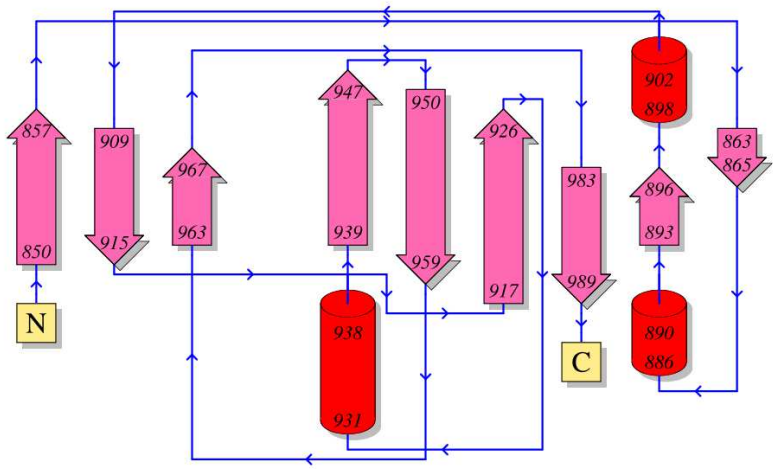
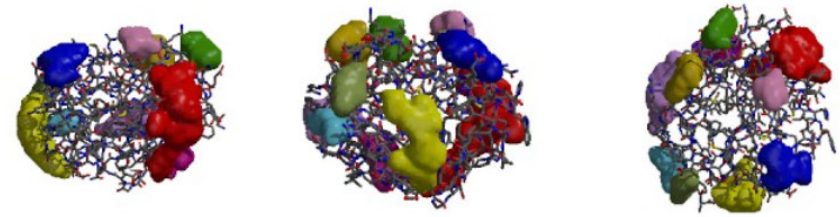


Figure S2. Secondary structure of FaGH2CBM. Image produced by PDBSum(Laskowski et al. 2018) showing α-helices as red barrels and β-strands as pink arrows. Blue arrows indicate the direction from N- to C-terminus.



Clefts

	Volume	R1 ratio	Accessible vertices		Buried vertices		Average depth		Residue type						
1	889.73	1.82	58.68	6	6.78	1	10.30	1	6	6	5	2	3	0	0
2	488.95	0.00	66.06	2	5.00	5	6.81	2	3	3	1	2	3	0	0
3	467.02	0.00	63.71	3	6.33	2	6.48	3	1	3	6	2	1	2	1
4	299.95	0.00	57.82	7	5.17	4	5.66	5	2	2	2	2	1	0	0
5	192.80	0.00	51.21	9	3.64	8	5.60	6	2	1	0	2	0	2	0
6	233.30	0.00	59.15	4	3.73	7	5.29	7	1	1	3	2	1	0	0
7	167.06	0.00	57.53	8	2.97	9	0.00	9	1	1	2	1	1	0	0
8	158.20	0.00	48.18	10	1.82	10	5.68	4	1	2	2	0	1	0	0
9	213.05	0.00	58.94	5	5.79	3	4.25	8	3	3	0	1	1	0	0
10	157.78	0.00	73.50	1	4.38	6	0.00	10	1	2	1	1	2	1	0

Protein structure

Figure S3. Clefts identified by PDBSum(Laskowski et al. 2018)**.** Volume is given in $Å^3$. Colour code for residue types: positive (blue), negative (red), neutral (green), aliphatic (grey), aromatic (purple) and others (yellow and orange).

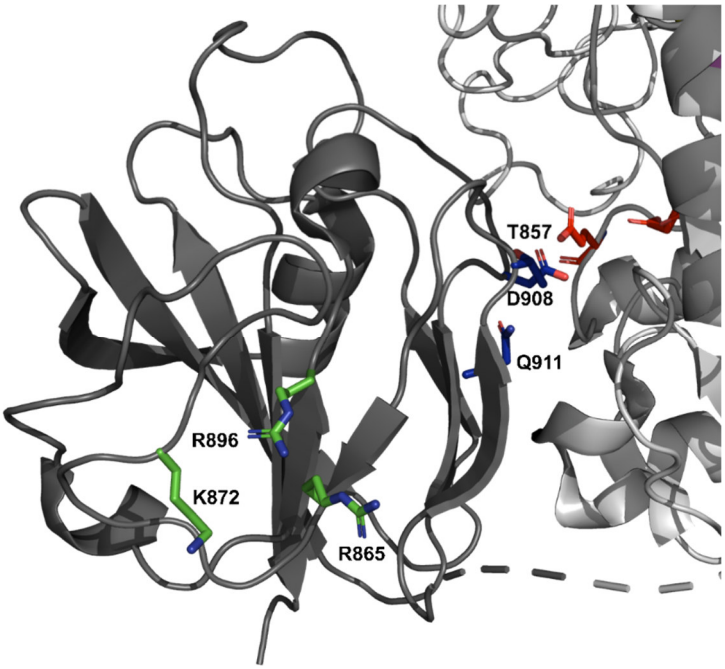


Figure S4. Potential binding regions identified by multiple sequence and structural alignment. Catalytic (light grey) and putative binding domain (dark grey) of P17_GH2 are shown in cartoon representation. Proposed catalytic residues of P17_GH2(Reisky et al. 2019) are shown in red and proposed binding residues are shown in green and blue.

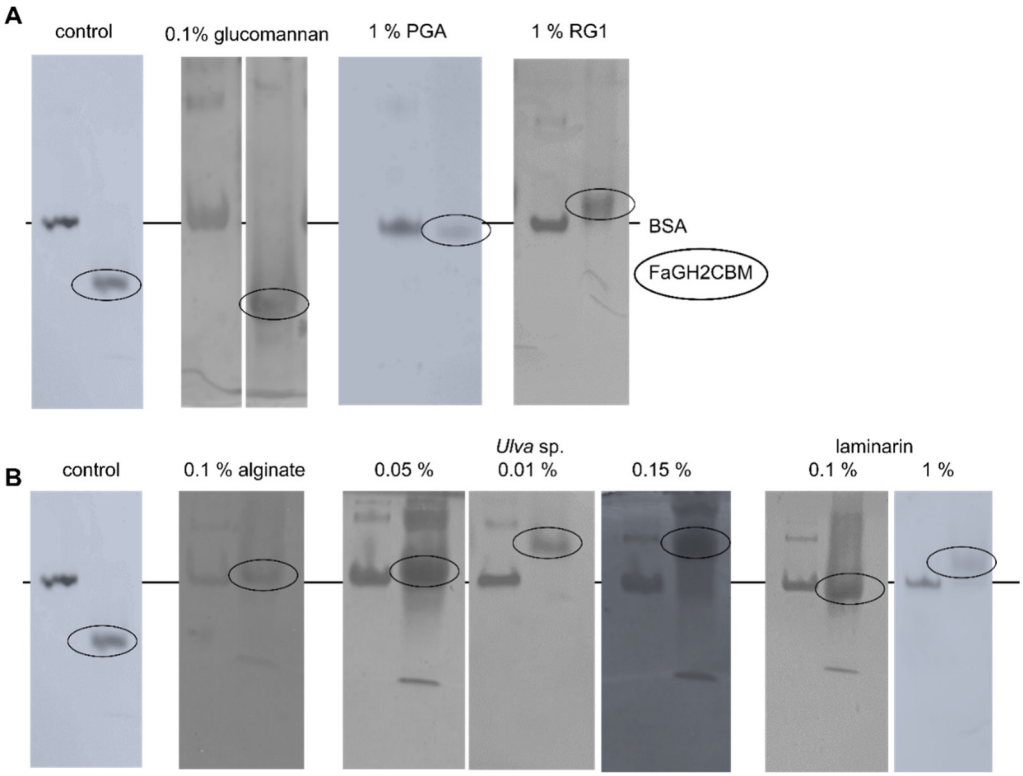


Figure S5. Retention of FaGH2CBM on (A) terrestrial and (B) marine polysaccharides. BSA was used as negative control and shows no retention to any of the tested glycans during affinity gel electrophoresis. The control gels contain no glycan.

Table S1. Identified β-turns by PDBSum (Laskowski et al. 2018)**.**

No	Turn	Sequence	Turn type	H-bond
1	Ser848-Pe851	SELF	IV	Yes
2	Glu849-Arg852	ELFR	IV	Yes
3	His869-Lys872	HWGK	II	No
4	Tyr875-Leu878	YTDL	I	Yes
5	Leu889-Ser892	LNES	II	No
6	Asp901-Tyr904	DNRY	I	No
7	Ala906-Gln909	ARDQ	IV	No
8	Asp908-Gln911	DQLQ	VIII	Yes
9	Ala914-Lys917	AGKK	IV	Yes
10	Asp925-Val928	DDTV	I	No
11	Asn946-Gly949	NVVG	IV	Yes
12	Val947-Ala950	VVGA	I	No
13	Glu969-Glu963	EEGE	II	No
14	Gly969-Asp972	GNSD	VIII	Yes
15	Ser971-Asp974	SDGD	IV	Yes
16	Pro976-Cys979	PENC	I	no

Table S2. Identified β-bulges by PDBSum (Laskowski et al. 2018)**.**

No.	Bulge type	Res X	Res 1	Res 2	Res 3
1	Antiparallel wide	Phe912A	Lys853A	Phe854A	
2	Antiparallel G1	Ala914A	Gly962A	Glu963A	
3	Antiparallel special	Val954A	Thr941A	Gly942A	Asp943A
4	Antiparallel classical	Met967A	Gln909A	Leu910A	

Table S3. Amino acid composition.

Residue	**abundance**	
	total	%
Al	8	5.6
Arg	10	7
Asn	9	6.3
Asp	11	7.7
Cys	1	0.7
Gln	2	1.4
Glu	11	7.7
Gly	9	6.3
His	4	5.6
Ile	4	4.9
Leu	9	4.2
Lys	8	3.5
Met	7	5.6
Phe	6	5.6
Pro	5	3.5
Ser	8	5.6
Thr	8	5.6
Trp	2	1.4
Tyr	10	7
Val	11	7.7

Table S1. Protein sequence and primer sequences of FaGH2CBM.

Amino acid sequence	SELFRKFSYTGDGKAMLRTNMHWGKKAYTDLEYNYTVLPRYLNESEYVRTP NSDNRYWARDQLQFIAGKKMHIYVLHDDTVPRPEFLLRDYEDTG DNVNVVGASMSVFHRVAEEGESIIMAGNSDGDAPENCRMYTVMVKEFK*
Molar extinction coefficient	25900
F primer (5'→'3)	CTG GTG CCG CGC GGC AGC CAT ATG GCT AGC CCA AAT GTT ATC TAT ATC CTC GCC
R primer (5'→'3)	ATC TCA GTG GTG GTG GTG GTG GTG CTC GAG CTA TTT GAA GCG TTT CTT GAA ATT TAC ATT

Table S2. Statistics of data collection and refinement.

Data collection	FaGH2CBM
PDB	Not submitted yet
Data set	3964 pos12
X-ray source	PETRA III, DESY, P11
Molecules in the ASU	3
Wavelength (Å)	1.0332
Resolution range (Å)	39.16-1.22
Space group	P1 21 1
Unit cell (Å) Unit cell (°)	34.91 37.10 157.06 90.00 94.09 90.00
R_{sym}(%)	0.113 (1.22)
Completeness (%)	96 (86.3)
Redundancy	5.7 (4.7)
$\langle I/\sigma(I)\rangle$	4.8 (0.6)
No. of Reflections	686684 (24845)
No. Unique	121078 (5306)
Mosaicity	0.21

Table S3. Polysaccharides used in binding analyses.

Polysaccharide	Source	Origin	Supplier	AGE	ELISA	CM
Polygalacturonic acid (α-(1→4)-D-galacturonic acid)	Citrus pectin	Land plants	Megazyme	+	+	+
Rhamnogalacturonan I	Soy bean	Land plants	Megazyme	+	+	+
Glucomannan (β-(1→4)-D-mannose and D-glucose backbone)	Konjac	Land plants	Megazyme	+	+	+
Laminarin (β-(1→3)-D-glucan)	*Eisenia bicyclis*	Macroalgae	Carbosynth	+	+	+
Alginate	Brown algae	Macroalgae	Sigma Aldrich	+	+	+
Ulvan	*Ulva armoricana*	Macroalgae	Carbosynth	+	-	+
α-(1→5)-L-arabinan	Sugar beet	Land plants	Megazyme	-	+	+
β-(1→4)-D-galactan	Potato	Land plants	Megazyme	-	+	+
Galactomannan (β-(1→4)-D-mannan backbone)	Carob	Land plants	Megazyme	-	+	+
Arabinoxylan	Wheat flour	Land plants	Megazyme	-	+	+
Carboxymethyl cellulose (β-(1→4)-D-glucan)	not available	Land plants	Megazyme	-	+	+
Xyloglucan	Tamarind seed	Land plants	Megazyme	-	+	+
Partially acetylated alginate	*Azotobacter vinelandii*	Bacteria	Megazyme	-	-	+
Glucurono-Xylomannan (α-(1→3)-D-mannan backbone)	*Tremella fuciformis*	Fungi	Elicityl	-	-	+
Pullulan (α-(1→4)(1→6)-D-glucan)	*Pullularia pullulans*	Fungi	Megazyme	-	-	+
Pachyman (β-(1→3)-D-glucan)	*Poria cocos*	Fungi	Megazyme	-	-	+
Scleroglucan (β-(1→3)-D-glucan backbone)	*Sclerotium rolfsii*	Fungi	Carbosynth	-	-	+
Mannan (β-(1→4)-D-mannan)	Ivory nut	Land plants	Megazyme	-	-	+
β-(1→4)-D-xylan	Beechwood	Land plants	Megazyme	-	-	+
β-glucan (β-(1→3)(1→4)-D-glucan)	Barley	Land plants	Megazyme	-	-	+
Lichenan (β-(1→3)(1→4)-D-glucan)	Icelandic moss	Lichen (Fungi)	Megazyme	-	-	+
Pustulan (β-(1→6)-D-glucan)	*Lasallia pustulata*	Lichen (Fungi)	Elicityl	-	-	+
Laminarin (β-(1→3)-D-glucan)	*Laminaria digitata*	Macroalgae		-	-	+
Fucoidan	*Fucus vesiculosus*	Macroalgae	Sigma Aldrich	-	-	+
Fucoidan	Sargassum fusiforme	Macroalgae	Glycomix	-	-	+
Fucoidan	Laminaria	Macroalgae	Glycomix	-	-	+
Galactofucan	*Undaria pinnatifida*	Macroalgae	Elicityl	-	-	+

Fucoidan	*Cladosiphon okamurans*	Macroalgae	Carbosynth	-	-	+
Fucoidan	*Ascophyllum nodosum*	Macroalgae	Carbosynth	-	-	+
Fucoidan	*Fucus serratus*	Macroalgae	Carbosynth	-	-	+
Porphyran	*Porphyra umbilicalis*	Macroalgae	-	-	-	+
Mannan (α-(1→6)-D-mannan backbone)	*Saccharomyces cerevisiae*	Yeast (Fungi)	Sigma Aldrich	-	-	+
Ulvan	*Enteromorpha sp.*	Macroalgae	Elicityl	-	+	-
Ulvan	*Ulva sp.*	Macroalgae	Elicityl	-	+	-
purified fucoidan	*Fucus serratus*	Macroalgae	Carbosynth	-	+	-
purified fucoidan	*Fucus vesiculosus*	Macroalgae	Carbosynth	-	+	-
purified fucoidan	*Undaria pinnatifida*	Macroalgae	Carbosynth	-	+	-
purified fucoidan	*Macrocystis pyrifera*	Macroalgae	Carbosynth	-	+	-
purified fucoidan	*Ecklonia maxima*	Macroalgae	Carbosynth	-	+	-
purified fucoidan	*Sargassum fusiforme*	Macroalgae	Carbosynth	-	+	-
purified fucoidan	*Cladosiphon okamurans*	Macroalgae	Carbosynth	-	+	-
purified fucoidan	*Lessoia nigresecens*	Macroalgae	Carbosynth	-	+	-
purified fucoidan	*Durvilaea potatorum*	Macroalgae	Carbosynth	-	+	-
GM-block alginate	Brown algae	Macroalgae	Sigma Aldrich	-	+	-
G-block alignate	Brown algae	Macroalgae	Sigma Aldrich	-	+	-
M-block alginate	Brown algae	Macroalgae	Sigma Aldrich	-	+	-
λ-carrageenan		Macroalgae	Sigma Aldrich	-	+	-
κ-carrageenan		Macroalgae	Sigma Aldrich	-	+	-
ι-carrageenan		Macroalgae	Sigma Aldrich	-	+	-
hyaluronic acid			Sigma Aldrich	-	+	-
chondritin sulfate	marine	-	Sigma Aldrich	-	+	-
total	**-**	**-**	**-**	**6**	**30**	**32**

Supplementary information – Chapter 2.1.

Table S1. Data collection and refinement statistics.

Data collection	21_S1_15
PDB	Not submitted yet
Structural homologue	PDB: 1E2S
Data set	21S1_15/21S1_15_011
X-ray source	PETRA III, DESY P11
Molecules in the ASU	1
Wavelength (Å)	1.0332
Resolution range, Å	48.94 - 1.70
Space group	P212121
Unit cell (Å) Unit cell (°)	61.72 80.31 97.25 90.00 90.00 90.00
R_{sym}(%)	0.072 (0.486)
Completeness (%)	99.4 (92.6)
Multiplicity	12.8 (12.6)
$\langle I/\sigma(I)\rangle$	20.0 (4.2)
CC(1/2)	0.999 (0.952)
No. of Reflections	687391 (33048)
No. Unique	53780 (2627)
Mosaicity	0.16
Wilson B.-factor	20.6

Table S2. Overview over cloned constructs.

Locus tag	LCC4_2_21
Signal peptide	0-19
Molar extinction coefficient	1.364
F primer (5'→'3)	CTG GTG CCG CGC GGC AGC CAT ATG GCT AGC CCA AAT GTT ATC TAT ATC CTC GCC
R primer (5'→'3)	ATC TCA GTG GTG GTG GTG GTG GTG CTC GAG CTA TTT GAA GCG TTT CTT GAA ATT TAC ATT

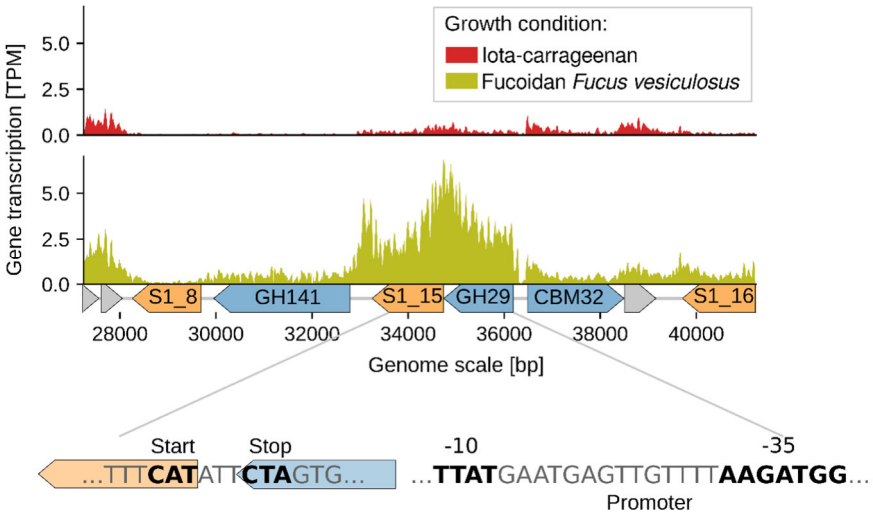


Figure S1. Operon for sulfated fucan degradation on the plasmid of '*Lentimonas*' sp. CC4. The genomic region surrounding the 21_S1_15 sulfatase and the 22_GH29 fucosidase located on the plasmid of '*Lentimonas*' sp. CC4 is shown. Sulfatases, CAZymes, and hypothetical proteins are shown in orange, blue and grey. Gene transcription from cells grown on iota-carrageenan or fucoidan from *Fucus vesiculosus* is represented as transcript per million (TPM) averaged over 10 bp.

Figure S2. Dendrogram and sequence alignment of subfamily S1_15. The Arylsulfatase of subfamily S1_3 from *Homo sapiens* was used as outgroup. Bar charts show relative abundance of phyla and habitat.

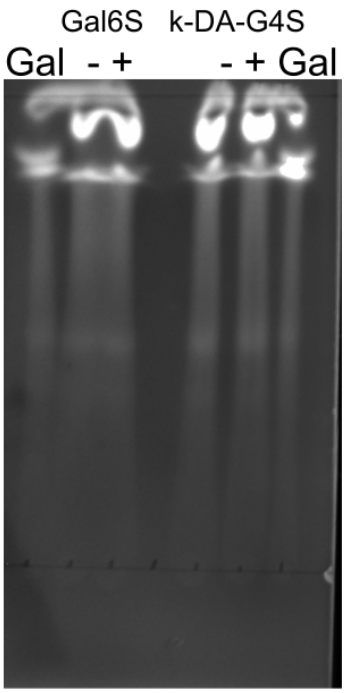


Figure S3. TLC of AMAC-labeled carrageenan oligosaccharides. Gal: D-galactose, Gal6S: D-Galactose-6-O-sulfate sodium salt, k-DA-G4S: Neocarrabiose-4-O-sulfate sodium salt DP2, -: no enzyme, +: with enzyme.

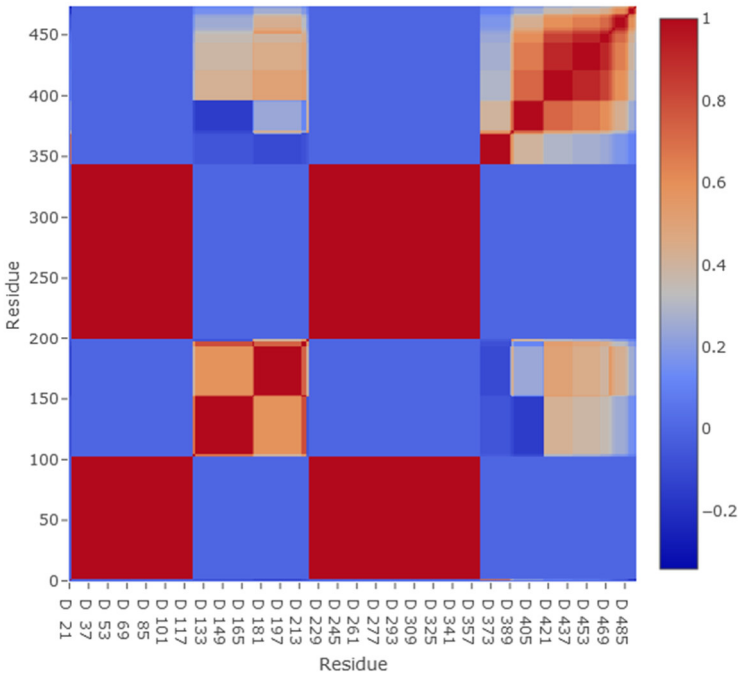


Figure S4. Correlation matrix of PDB matched structures along positions of 21_S1_15. Correlation is undefined if a position is invariantly occupied or empty; in these cases, correlation is set to one (two invariantly occupied columns) or zero (otherwise). X axis labels are PDB residue numbers and Y axis labels are sequential residue numbers (Holm and Elofsson 2019).

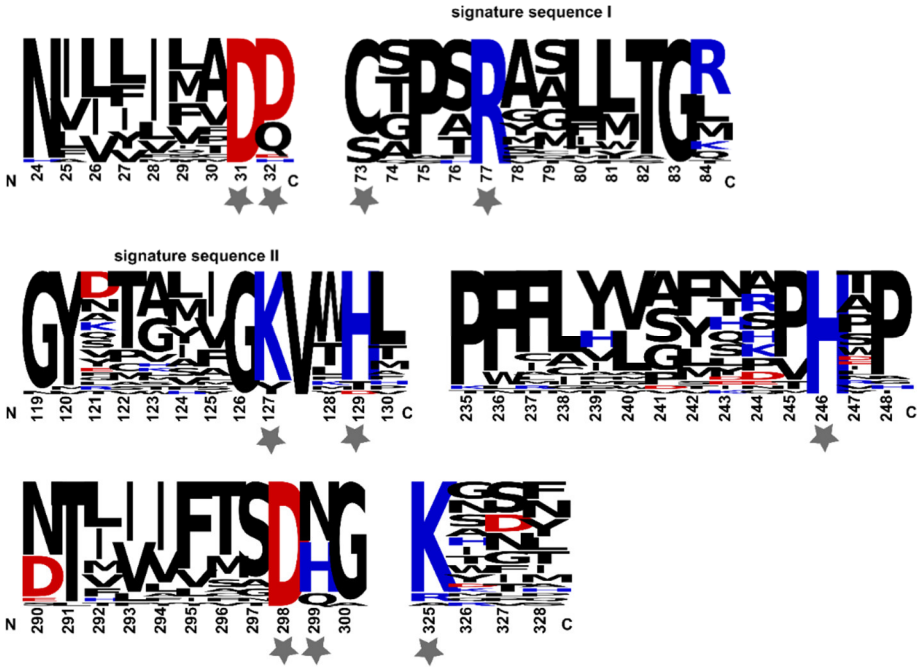


Figure S5. Regions and signature sequences of arylsulfatases active site residues. Sequences derive from structures of different S1 subfamilies. Degree of conservation is displayed as size of the amino acids. Positively charged residues: blue, hydrophobic residues: black, negatively residues: red. Stars indicated conserved residues of the active site.

Supplementary information – Chapter 2.2.

Table S1. Primer sequences of 30_GH95.

LCC4_2_30_GH95

MKITSSLALLALSLVGLNASDERYVLQYDSPAANDLIKAKGKSKEPKFIKTALPLGNGRLGAMFSG
GVDEEHLCINDITLWMNTRRGQDEVAQSAARSVTPEDYEKVREAYREERMGSGPESMESISTKYL
STKQKVGNYAPFTNVVIATGHDAASVSNYQRSLDTRTGLGRVSYSANGGQFTREYFCSYPNDVM
GARYTAEGAKLNLTIQTITYHKVNVLKATGNRILLEAETNMTQDNMHYAQIIEVDAADGKVTAQA
DGSIKVSGATDVSIYLVGYTDYLPNYPTFKGRDFRGDAEKAINAASKLGYHALKQAHVADFTALM
DRCRLELDFKPSGKTTDELTKNPSLELENLYFHHSRYLQISCSRGAPVPSNLQGLWNASLKPAWNA
DYHWDINVAMNYWMVETANLPESFAPYMEFMKVMAVSGQHTARETFGINQGWSMGLNGNIFG
WTAQNEHGRRMQQSGHWAAQHLLEHYAFGGDKEYLEEAYPIMKGAAEFFVEHLSPWKDGTLVV
YPTWSPENNYAKGKLNKQTYGAAWDQQLILGLFTDCIEASYILNRDAEFRATLQKMIPQLSPQKIG
KHGQLQEWIDDYDDPKNKHRHISHLIALHPGRDISPLTTPELHQAALVTMKHRGDASTGWSTGWK
TNFWSRLHDGDHAHKIYQFLTSKRAYPNLFDFHPPFQIDGNFGGGAGVCEMLLQSHLRSVNPNAA
TVSEAAFIAYQQSPENPKHFVPVVPGESLVSAPFILHLLPALPSEWATGKVKGLRARGGFEVDIEW
VNGQLVEATIHATRDSACRVYSNGKISDTITLKKGESKVWK*

Forward primer:

CTGGTGCCGCGCGGCAGCCATATGGCTAGC GATGAGCGTTATGTCCTACAA

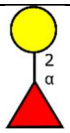
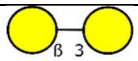
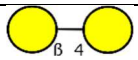
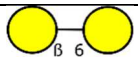
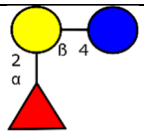
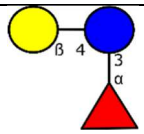
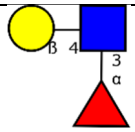
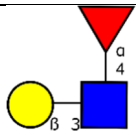
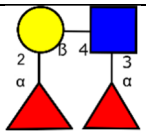
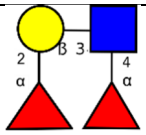
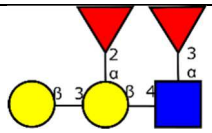
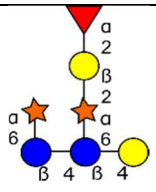
R primer:

ATCTCAGTGGTGGTGGTGGTGGTGCTCGAG TTACTTCCAAACCTTCGATTC

Table S2. IEX-purified fucoidans from macroalgae. The fucose content is given in % and mg ml^{-1} in 9 mg ml^{-1} fucoidans used for enzymatic digests.

Species	Fucose (%)	Fucose (mg)	Supplier
Fucus vesiculosus	52	4.68	Carbosynth
Fucus serratus	13.62	1.23	Carbosynth
Durvillaea potatorum	22.82	2.05	Carbosynth
Sargassum fusiforme	29.08	2.62	Carbosynth
Ecklonia maxmima	42	3.76	Carbosynth
Lessonia nigrescens	51.72	4.65	Carbosynth
Undaria pinnatifida	35.7	3.21	Carbosynth
Macrocystis pyrifera	49.8	4.48	Carbosynth
Cladosiphon okamuranus	32.87	2.96	Carbosynth

Table S3. Activity screening using human milk and blood oligosaccharides.

Substrate	Structure	Fucose	Galactose
Blood group H		✓	✓
1,3-galactobiose		n.d.	n.d.
1,4-galactobiose		n.d.	n.d.
1,6-galactobiose		n.d.	n.d.
2-fusosyllactose		✓	✓
3'-fucosyllatose		n.d.	n.d.
Lewis$_x$		✓	n.d.
Lewis$_a$		n.d.	n.d.
Lewis$_y$		n.d.	n.d.
Lewis$_b$		n.d.	n.d.
Blood group B type 5		n.d.	n.d.
Reduced XFG xylogucan		✓	n.d.

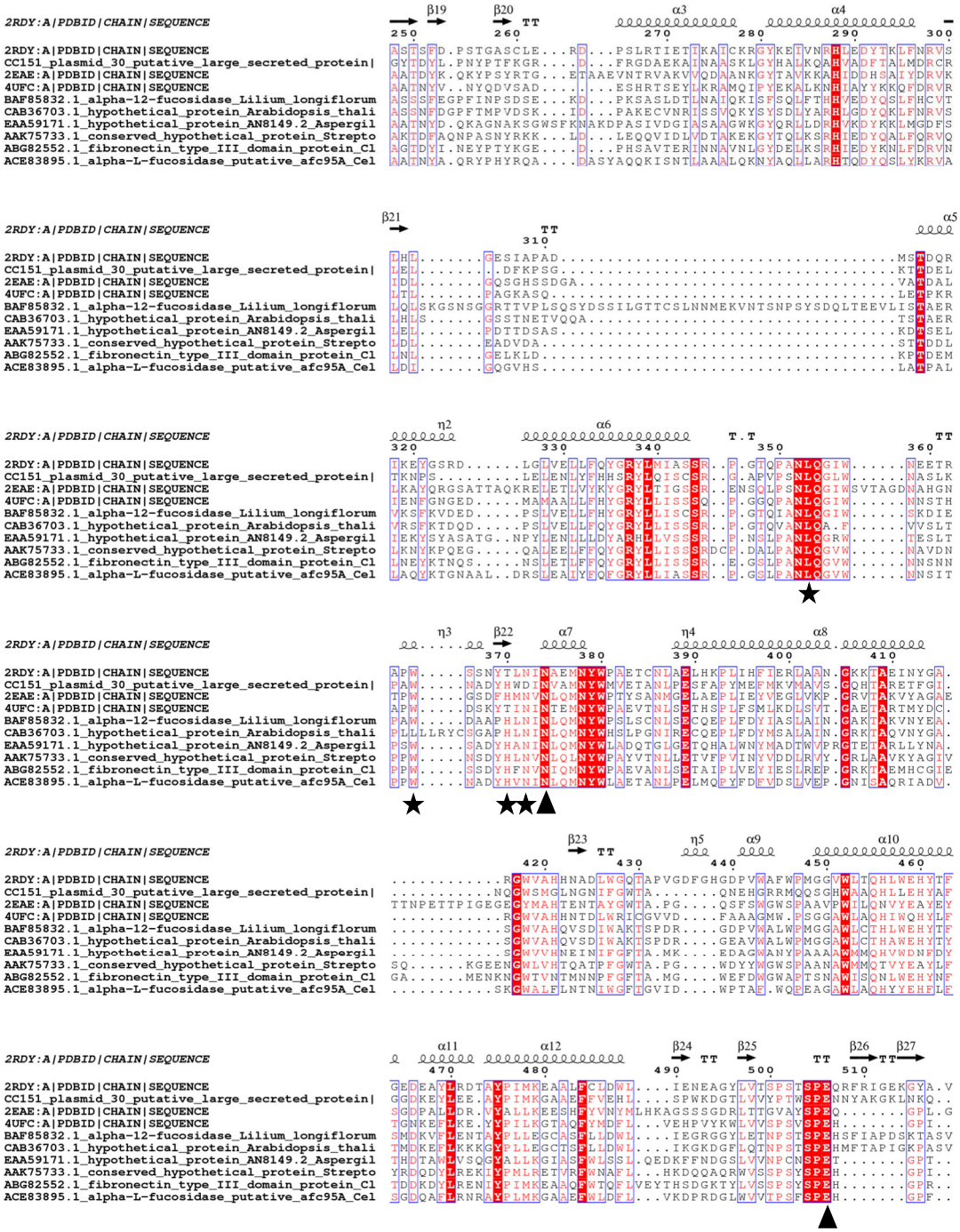


Figure S1. Multiple sequence alignment of 30_GH95 identifies putative catalytic residues. Catalytic residues are marked with a triangle (▲), whereas other highly conserved active site residues are marked with a star (★).

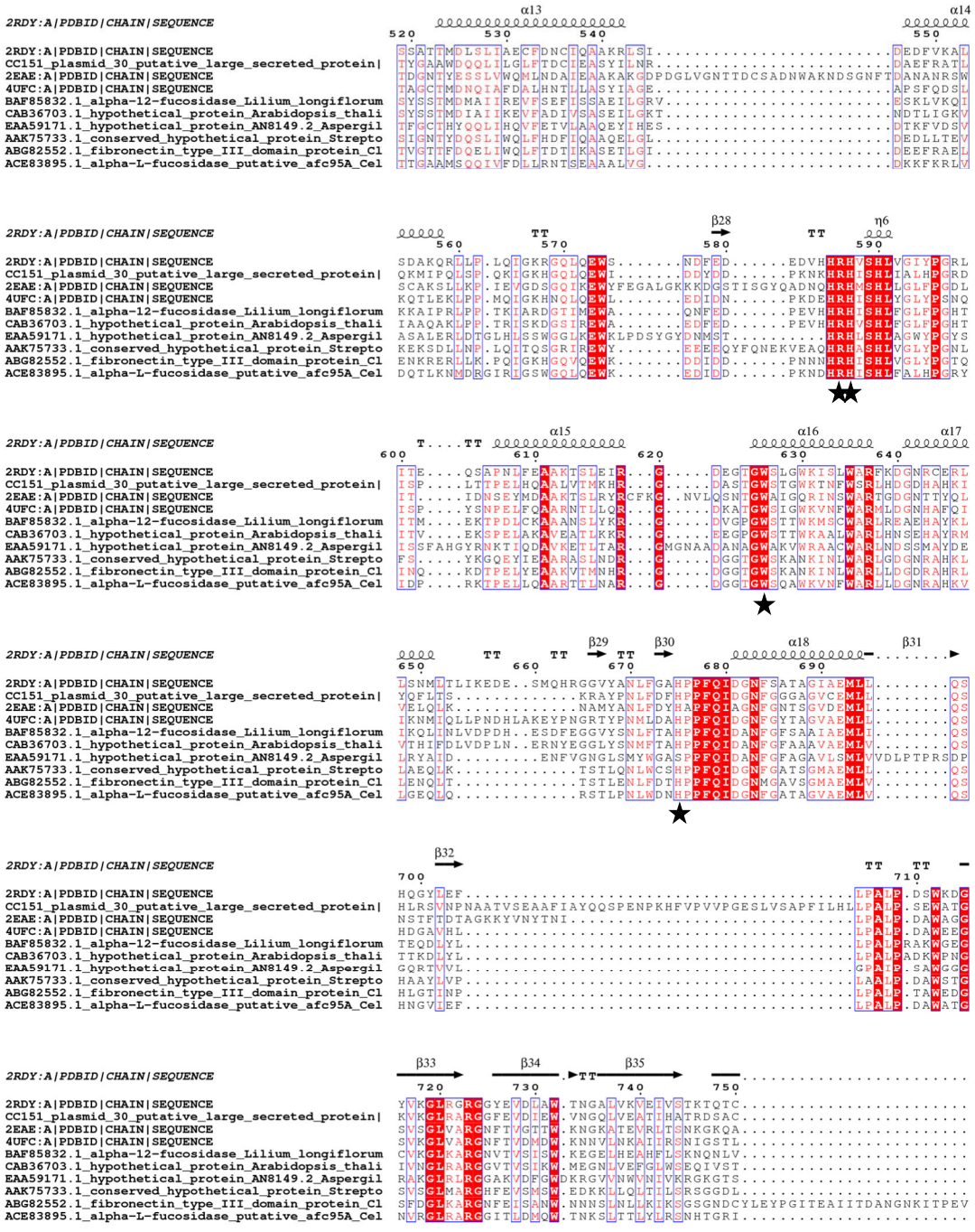


Figure S1. Multiple sequence alignment of 30_GH95 identifies putative catalytic residues. Continued.

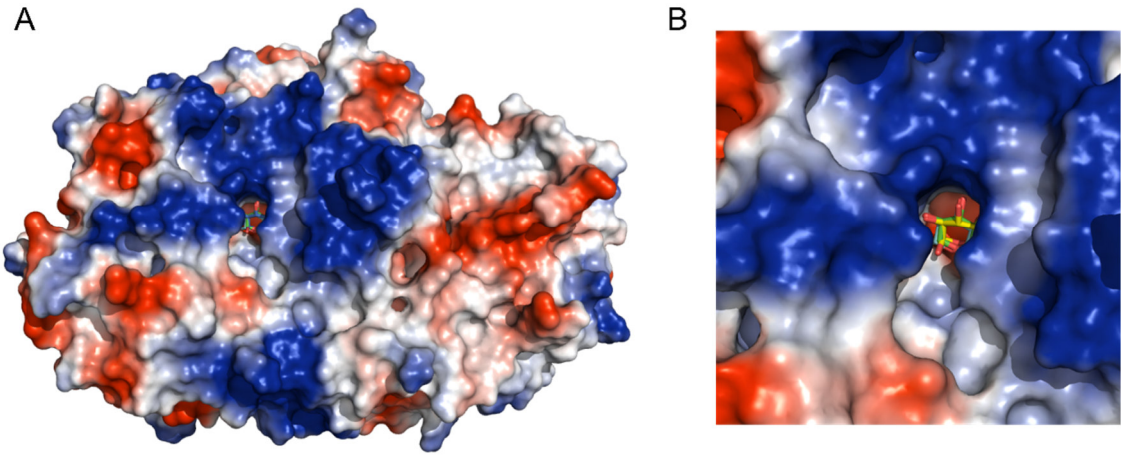


Figure S2. Structure model of 30_GH95. Surface representation of (A) 30_GH95 and (B) closeup of active site. Negatively charged small pocket-structure of the active site indicates exo-acting mechanism of desulfated substrate. Blue: negatively charged amino acids, red: positively charged amino acids. Superimposition of α-Fuc (yellow), β-Fuc (green) and Gal (turquois) from PDB: 4UFC and 2EAE into the active site. This model was generated by Phyre2. Confidence in the model: 769 residues (94 % coverage) modelled at > 90 % accuracy.

Supplementary information – Chapter 2.3.

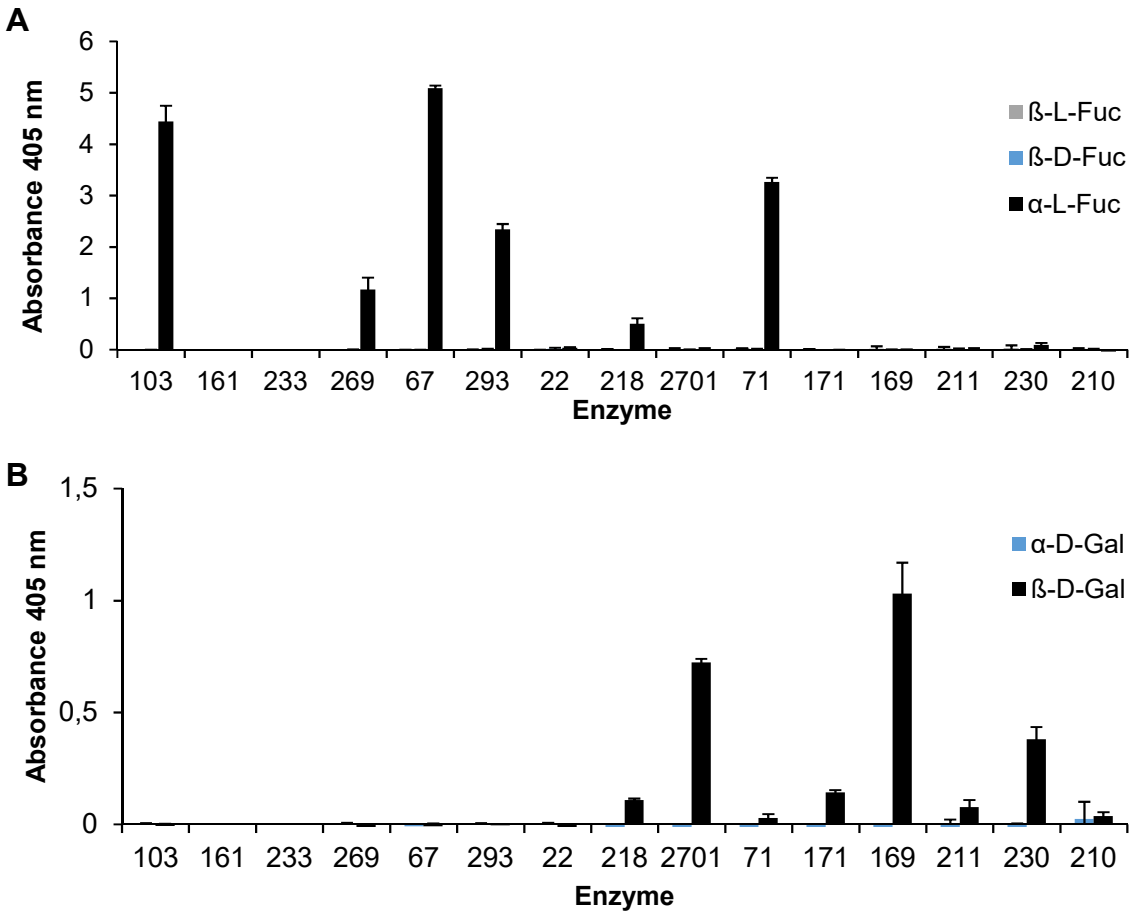


Fig. S1. Activity screening using chromogenic substrates. 0.05 mg ml^{-1} GH29s were incubated at 30°C with 1 mM substrate (A) fucopyranose or (B) galactopyranose. The reaction was stopped with the same volume of 250 mM NaOH.

Tab S1. Biochemical characterization of GH29s using α-L-pNP-fucopyranoside. ND: not determined. -: absence of activity.

Locus tag	Molecular weight (kDa)	Thermal stability (°C)	pH and buffer optimum	NaCl optimum [mM]	Cofactors
LCC4_2_22	54.5	38	pH 5.7 Na-citrate	0.2 - 1	↓ $Cu(II)SO_4$
LCC4_2_71	62.7	33.5	pH 5.4 - 8	0.4	↑ $MnCl_2$
LCC4_2_169	51.1	41	pH 6.6 Na-cacodylate	-	↓ $CuSO_4^{2-}$ ↓ $CuCl_2$ ↓ $ZnCl_2$
LCC4_2_171	56.8	35	pH 5.7 – 6.6	0.7	↓ $MnCl_2$ ↓ $ZnCl_2$ ↓ $CuSO_4^{2-}$
LCC4_2_210	56	40	-	-	-
LCC4_2_211	56.5	37	-	-	-
LCC4_2_230	54.6	36.5	pH 5.4 Na-citrate	1	-
LCC4_2_67	53.6	46	pH 5 Na-citrate	-	↓ $MnCl_2$
LCC4_2_233	65.4	30	ND	ND	ND
LCC4_2_269	68.2	40	pH 5 Na-citrate	-	-
LCC4_2_293	73	40	pH 5.7 Na-citrate	0.2	-
LCC4_2_103	66.9	46	pH 5.7 Na-citrate	-	↓ $MnCl_2$
LCC4_2_161	68.7	ND	ND	ND	ND
LCC4_1_2701	60.6	ND	pH 5.7 Na-citrate	0.7	↓ $CoCl_2$ ↓ $CuCl_2$ ↓ $ZnCl_2$
LCC4_2_218	60	37	pH 5.4 - 6	ND	↓ $CuSO_4^{2-}$ ↓ $MnCl_2$ ↓ $ZnCl_2$

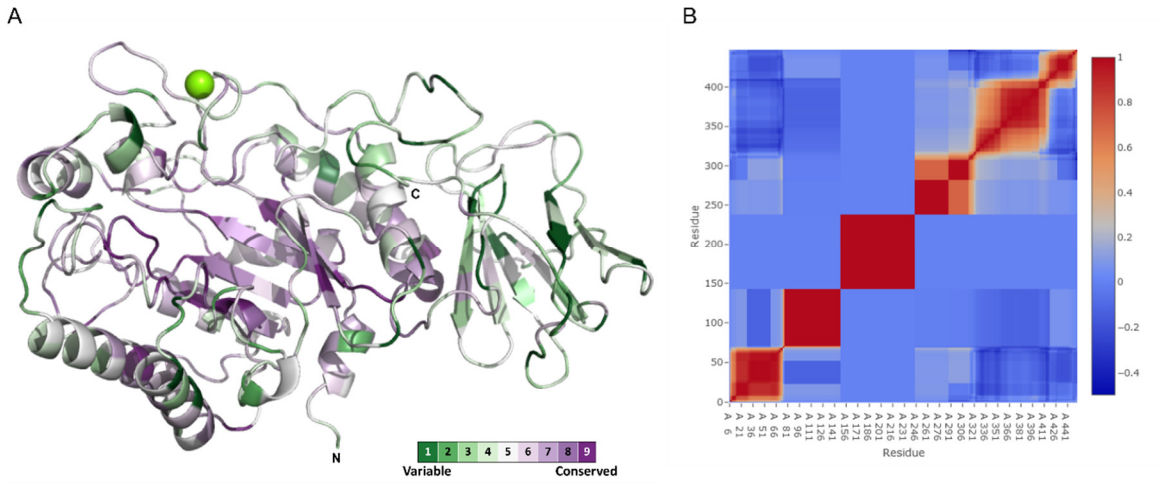


Fig. S2. Marine GH29s have a structurally diverse C-terminus. (A) Cartoon representation of LCC4_2_22: Conserved residues among GH29s from '*Lentimonas'* sp. CC4 are color-coded from variable (green) to conserved (purple). Magnesium ion is shown as sphere. (B) Match correlation matrix: Correlation of matched structures from PDB (x-axis) along positions of 22_GH29 (y-axis). Structural domains show strong positive correlation (red) within domains and negative correlation (blue) between domains.

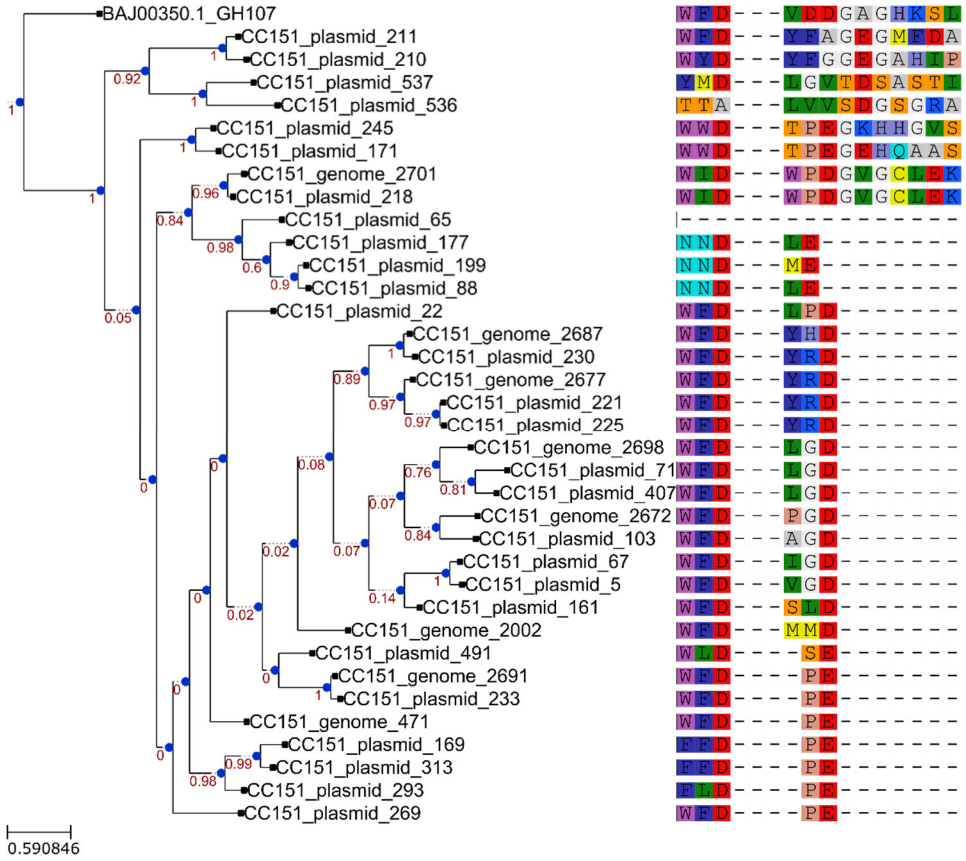


Fig. S3. GH29s from '*Lentimonas*' sp. CC4. Bootstrap values are shown at branch node of dendrogram. Structural alignment revealing conserved aspartic acid (D) as nucleophile and aspartic or glutamic acid (D/E) as acid/base. Family GH107 was used as outgroup.

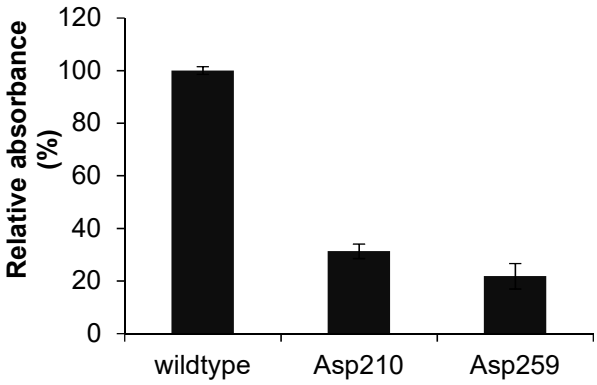


Fig. S4. **Single-site mutagenesis of catalytic residue pair of 22_GH29.** Relative activity of 22_GH29$_{D210A}$ and 22_GH29$_{D259A}$ on pNP-α-L-Fuc in comparison with wildtype. Error bars indicate standard deviation of technical triplicates.

Tab. S2. Primer list of cloned GH29s.

Locus tag	Signal peptide	Molecular Extinction coefficient	F primer (5'→'3)	R primer (3'→'5)
LCC4_2_22	0-22	2.021	CTG GTG CCG CGC GGC AGC CAT ATG GCT AGC GCC GAC CAG ACG GAC GAT	ATC TCA GTG GTG GTG GTG GTG GTG CTC GAG CTA GTG ACG TGC GGC TTT AAC
LCC4_2_22 Asp210mutant	0-22	2.021	TGGTTCGCTGTCGGTGGT AATAATGTCGTCGAAGCC	CGACAGCGAACCACAACACGG CCATATCGC
LCC4_2_22 Asp259mutant	0-22	2.021	CTTCCAGCCCGCATGCTA CCAGCGAAGCGTAT	GCGGGCTGGAAGGGAGTTAAA GTCAGCATATTTTTGCTG
LCC4_2_71	/	2.344	CTGGTGCCGCGCGGCAGCCATA TGGCTAGCATGGACAAGATGTG GGGCG	ATCTCAGTGGTGGTGGTGGTGGTGCT CGAGTTATTAGAGCTCGTTGACGGGT G
LCC4_2_169	0-26	1.548	CTGGTGCCGCGCGGCAGCCATA TGGCTAGCCGCATCGGGCCACT ATTGAA	ATCTCAGTGGTGGTGGTGGTGGTGCT CGAGTTATTATTTCACAAACTCTACA TGATC
LCC4_2_171	0-29	2.564	CTGGTGCCGCGCGGCAGCCATA TGGCTAGCGACGAACTCGGCGA AGACC	ATCTCAGTGGTGGTGGTGGTGGTGCT CGAGTTACTAGAACACAATCTTCACC ATGTC
LCC4_2_210	0-20	2.064	CTGGTGCCGCGCGGCAGCCATA TGGCTAGCGAGGAGTTTGACTA TGCAGCG	ATCTCAGTGGTGGTGGTGGTGGTGCT CGAGTTATTATTTCGATCGAACGGTA TCGC
LCC4_2_211	0-26	1.919	CTGGTGCCGCGCGGCAGCCATA TGGCTAGCAAGCAAGAACTCAC GCAAGAG	ATCTCAGTGGTGGTGGTGGTGGTGCT CGAGTTATTATTTGGATCGAATCGCA TCCC
LCC4_2_230	0-24	2.024	CTGGTGCCGCGCGGCAGCCATA TGGCTAGCGAGCGTTCAGAATG GCAAGTG	ATCTCAGTGGTGGTGGTGGTGGTGCT CGAGTTATTAAAATTCTAAACAAACA ACGGT
LCC4_2_67	0-22	1.465	CTGGTGCCGCGCGGCAGCCATA TGGCTAGCAATACCAAAAAGGC GCCCAA	ATCTCAGTGGTGGTGGTGGTGGTGCT CGAGTTAGAGGGTAACTTTAATCACC
LCC4_2_233	0-21	2.387	CTGGTGCCGCGCGGCAGCCATA TGGCTAGCGCAAAGCCTGAGAC CGAAGT	ATCTCAGTGGTGGTGGTGGTGGTGCT CGAGTTACTAATCTGAAACGGGAAC
LCC4_2_269	0-21	1.859	CTGGTGCCGCGCGGCAGCCATA TGGCTAGCAATAGCGGTTCAGA GGCGC	ATCTCAGTGGTGGTGGTGGTGGTGCT CGAGTCAATTGAATGGAATCAGCCT
LCC4_2_293	0-30	1.322	CTGGTGCCGCGCGGCAGCCATA TG GCTAGCCAAACCGATTCAAAAC CGAAA	CTGGTGCCGCGCGGCAGCCATA TGGCTAGCCAAACCGATTCAAAACCG AAA
LCC4_2_103	0-29	1.966	CTGGTGCCGCGCGGCAGCCATA TGGCTAGCGTGTCTCAGCGAGC AGAC	ATCTCAGTGGTGGTGGTGGTGGTGCT CGAGCTACTGACTTAGTATTACTTCC
LCC4_2_161	0-21	1.508	CTGGTGCCGCGCGGCAGCCATA TGGCTAGCGCCGGTAAACCTTC AACGA	ATCTCAGTGGTGGTGGTGGTGGTGCT CGAGTCAATCGCGTTCGCTGGAT
LCC4_2_218	0-23	2.342	CTGGTGCCGCGCGGCAGCCATA TGGCTAGCGACAGCTCGAATCA ACCGT	ATCTCAGTGGTGGTGGTGGTGGTGCT CGAGTTATTTTTTCTTCTGTTTCTTGG
LCC4_1_2701	0-23	2.249	CTG GTG CCG CGC GGC AGC CAT ATG GCT AGC TCT GGA ACA CCT TAT GAC GG	ATC TCA GTG GTG GTG GTG GTG GTG CTC GAG TTA ACG ATA CGG TTC GAG CGT

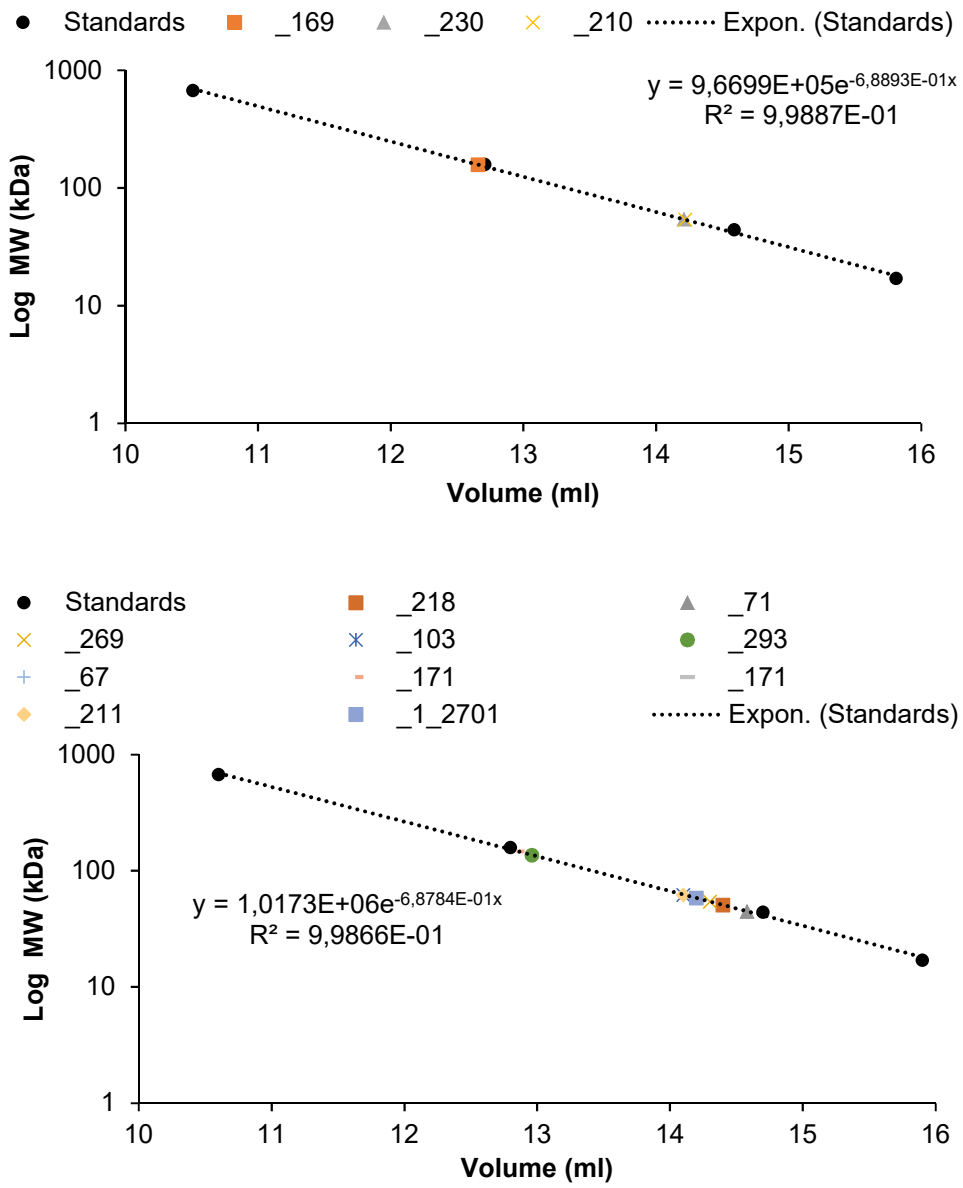


Fig. S5. Calibration curves of analytical SEC. Logarithmic scale of the molecular mass (kDa) in dependence of the elution volume (ml).

Tab. S3. X-ray data collection and structure statistics.

Data collection	LCC4_2_22	LCC4_2_71	LCC4_2_169	LCC4_1_2701
Structural homologue	2WVS	Ensemble	6GN6	Ensemble
Data set	GH29 1018	3967 pos6	3967 pos14	3692 pos11
X-ray source	PETRA III, EMBL P13	DLS MX	DLS MX	PETRA III, DESY P11
PDB	Not submitted yet	Not submitted yet	Not submitted yet	Not submitted yet
Molecules in the ASU	1	4	2	
Wavelength (Å)	1.033	1.0332	1.0332	1.0332
Resolution range, Å	59.63-1.778	58.59-2.11	62.54-1.63	69.55-2.74
Space group	P21 21 21	P212121	P 63 2 2	P1 21 1
Unit cell (Å) Unit cell (°)	74.99 76.03 96.09 90 90 90	95.94 112.86 391.98 90 90 90	144.43 144.43 61.25 90 90 120	67.86 139.11 93.94 90 102.12 90
R_{sym}(%)	0.167 (0.652)	0.158 (1.668)	0.25 (6.348)	0.192 (3.538)
Completeness (%)	98.5 (98.9)	97.4 (76.5)	93.5 (57.0)	99.9 (99.1)
Redundancy	7.2 (7.5)	11 (4.7)	37.2 (29.9)	6.8 (7.0)
$\langle I/\sigma(I)\rangle$	7.9 (3.3)	8.3 (0.6)	11 (0.4)	6.4 (0.5)
No. of Reflections	377820 (22268)	2747462 (44574)	1733159 (40508)	331924 (31089)
No. Unique	52571 (2959)	249131 (9572)	46579 (1355)	49074 (4462)
Mosaicity	0.44	0.0	0.0	0.0

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

gem. § 5 Abs. 5 der Promotionsordnung vom 18.06.2018 /
according to § 5 (5) of the Doctoral Degree Rules and Regulations of 18 June, 2018

Ich / I, Nadine Gerlach, Herbststraße 74 28215 Bremen, 2714809
(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, 10.07.2020

Ort / *Place*, Datum / *Date*

Unterschrift / *Signature*