

Discovery and characterization of enzymes acting on chitin

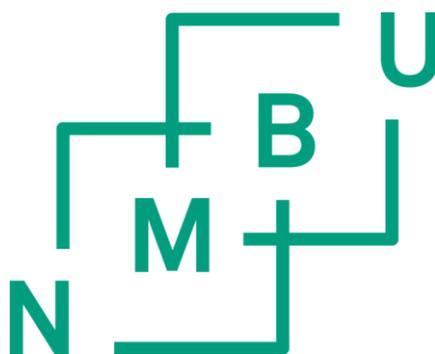
Oppdagelse og karakterisering av kitin-aktive enzymer

Philosophiae Doctor (PhD) Thesis

Tina Rise Tuveng

Norwegian University of Life Sciences
Faculty of Chemistry, Biotechnology and Food Science

Ås 2017



Thesis number 2017:71
ISSN 1894-6402
ISBN 978-82-575-1467-9

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	i
SUMMARY	iii
SAMMENDRAG	vii
ABBREVIATIONS	xi
LIST OF PAPERS	xiii
1 INTRODUCTION	1
1.1 Chitin	1
1.2 Chitosan and chitooligosaccharides	3
1.3 Carbohydrate-active enzymes	6
1.3.1 CAZymes in chitin degradation and modification	6
1.3.1.1 Chitinases	6
1.3.1.2 β -N-acetylhexosaminidases	11
1.3.1.3 Lytic polysaccharide monoxygenases	11
1.3.1.4 Carbohydrate esterases	13
1.3.1.4.1 Carbohydrate esterase family 4	13
1.3.1.5 Chitosanases	17
1.3.1.6 Carbohydrate-binding modules	17
1.3.2 Biological roles of chitin-active enzymes	18
1.4 Microbial degradation and utilization of chitin	21
1.4.1 Chitin degradation by <i>Serratia marcescens</i>	21
1.4.2 Chitin degradation by bacteria in the Bacteroidetes phylum	24
1.4.3 Chitin degradation by <i>Thermococcus kodakaraensis</i>	25
1.4.4 Chitin degradation by fungi	26
1.5 Protein secretion in Gram-negative bacteria	27
1.5.1 <i>In silico</i> prediction of secreted proteins	30
1.6 Proteomics as a tool for studying bacterial secretomes	32
2 OUTLINE AND PURPOSE OF THE RESEARCH PRESENTED IN THIS THESIS	37
3 MAIN RESULTS AND DISCUSSION	39
3.1 Paper I – Structure and function of a CE4 deacetylase isolated from a marine environment	39
3.2 Paper II – Proteomic investigation of the secretome of <i>Cellvibrio japonicus</i> during growth on chitin	46
3.3 Paper III – Chitin degradation by <i>Cellvibrio japonicus</i> is dependent on the non-redundant CjChi18D chitinase	52
3.4 Paper IV – Genomic, proteomic and biochemical analysis of the chitinolytic machinery of <i>Serratia marcescens</i> B JL200	60
4 CONCLUDING REMARKS AND PERSPECTIVES	65
5 REFERENCES	68
PAPERS I-IV	APPENDIX

ACKNOWLEDGEMENTS

The work presented in this thesis was carried out in the Protein Engineering and Proteomics (PEP) group, Faculty of Chemistry, Biotechnology and Food Science at the Norwegian University of Life Sciences in the period from 2013 to 2017. The project was financed by the Norwegian Research Council through the Marpol project (project code 221576) lead by Prof. Gudmund Skjåk-Bræk at the Norwegian University of Science and Technology.

Firstly, I would like to express my gratitude to my main supervisor **Prof. Vincent G. H. Eijsink** for offering me this Ph.D. position. Your knowledge and enthusiasm is admirable and very inspiring. Thank you for being a great supervisor; a meeting with you is always good for the motivation. I would also like to express my gratefulness to my co-supervisor **Assoc. Prof. Gustav Vaaje-Kolstad**. Thank you for answering questions, sharing ideas, and being supportive. I am especially grateful for all the work both of you have done for helping me achieve today's result.

Furthermore, I would like to thank **Dr. Magnus Ø. Arntzen** for sharing your impressive knowledge of proteomics, patiently teaching me the laboratory techniques in this field, and for answering all my questions.

A special thanks to **Zarah Forsberg** for all scientific and non-scientific discussions in the office. Thanks to **Sophanit Mekasha** for always being positive and for all the early morning drives to work, and late night drives home. I would also like to thank the rest of the PEP-group, you have made the last four years very enjoyable.

To the "**Biotech-girls**", thank you for the Tuesday evenings and annual weekend trips. If I'm worried about work, time with you always helps.

To my family, especially **Mum** and **Halvor**, thank you for always being supportive and pretending to understand what I have done at work during the last four years. Last, but not least, **Martin**, thank you for just being you and reminding me of the important things in life. As my namesake Tina Turner sings; you're simply the best!

Tina R. Tuveng
Ås, September 2017

SUMMARY

In the shift from a fossil-based to a bio-based economy, exploration of renewable resources is needed. Chitin is considered as the second most abundant polysaccharide on Earth, after cellulose, and its water-soluble derivatives chitosan and chitooligosaccharides (CHOS) have several applications, for example in medicine, agriculture, and the food industry. Today, the extraction of chitin from chitin-rich biomasses and the subsequent production of chitosan and CHOS involve harsh chemicals. It is of interest to replace the current chemical processing technology with enzyme-driven processes, since this would be more environmentally friendly. In addition, enzymes can be used to produce well-defined chitosans and CHOS, which is of interest, since the bioactivity of these compounds depends on properties such as the fraction of acetylation (F_A), the degree of polymerization (DP) and the pattern of acetylation (P_A). Investigation of proteins utilized by microorganisms during growth on chitin might provide insight into natural chitin conversion and may yield enzymes that can aid in industrial valorization of chitin-rich biomasses.

Paper I describes the characterization of a carbohydrate esterase family 4 (CE4) deacetylase, which was selected because of its potential application in the production of CHOS with defined F_A and P_A . To utilize these enzymes in an optimal way, good understanding of their substrate interactions and specificities is needed. **Paper I** includes the first enzyme-substrate complex of a CE4 deacetylase with an open active site, providing valuable insight into how the enzyme interacts with its substrate. The enzyme is able to deacetylate a variety of substrates at varying positions. This broad specificity and the presence of seemingly few subsites occupied by the substrate indicate that it may be difficult to use or develop this type of CE4 enzymes for enzymatic tailoring of the P_A of CHOS.

The genome of *Cellvibrio japonicus* encodes a large array of carbohydrate-active enzymes, including several putative chitinases and other enzymes possibly involved in chitin degradation. Whether these enzymes are actually involved in chitin utilization by this Gram-negative bacterium had not been investigated at the start of the work described in this thesis. **Paper II** describes a study of proteins that *C. japonicus* secretes during growth on chitin, using a novel, plate-based proteomics approach which yielded secretome samples with a relatively low fraction of cytoplasmic proteins. This study revealed that the four glycosyl hydrolase family 18 (GH18) chitinases encoded in the *C. japonicus* genome are produced in

high amounts, indicating that these enzymes are all involved in natural chitin turnover. Chitin degradation studies showed that *C. japonicus* has considerable chitinolytic power. The proteomics study revealed several proteins without an obvious role in chitin degradation that also are produced in high amounts during growth on chitin, thus providing a list of proteins that could be targeted in future searches for proteins that degrade chitin-rich biomass.

Paper III describes an in-depth investigation of the GH18 chitinases encoded by *C. japonicus*. Knockout studies showed that one of the chitinases, *CjChi18D*, is crucial for the bacterium's ability to utilize chitin as a carbon source. Biochemical characterization showed that *CjChi18D* is the most efficient chitin degrader, which could explain its crucial role. Comparative studies of the four enzymes indicated different and putatively complementary functions, as exemplified by *CjChi18C* having the by far highest activity against chitohexaose. Indeed, when combining enzymes, synergistic effects on chitin degradation efficiency were observed. Transcriptomic analysis showed that the four GH18 chitinases and a chitin-active LPMO, *CjLPMO10A*, are strongly up-regulated when *C. japonicus* grows on chitin, along with several other putatively chitin-active enzymes as well as a few proteins of unknown function, which are up-regulated to a lesser extent.

Serratia marcescens produces one the best studied chitinolytic machineries, involving three chitinases, a lytic polysaccharide monooxygenase, and a chitobiase. However, the genome sequence of one of the most frequently studied *S. marcescens* strains was not available at the start of this thesis work, and possible involvement of other proteins in chitin utilization had not been investigated. **Paper IV** describes the genome sequence of *S. marcescens* BJL200 and a proteomics investigation of proteins secreted during growth on chitin. The genome sequence showed that *S. marcescens* encodes a fourth chitinase, *SmChiD*, but the proteomics data indicated that this chitinase is not important in chitin utilization. Indeed, biochemical characterization of *SmChiD* supported the notion that this enzyme is not important for chitin conversion and, thus, likely has another, yet unknown, biological role.

Taken together, the results presented in this thesis provide novel insight into chitin-active enzymes encoded by bacteria. **Paper I** provides insights into the substrate binding of CE4 deacetylases with an open active site. **Papers II-IV** reveal chitin-active enzymes, in particular hydrolases, that play key roles in natural chitin conversion. Additionally, **Papers II-IV** yield a list with proteins without an obvious role in chitin degradation, which may be targeted in

future studies of the degradation of chitin-rich biomasses. Further studies on tailoring CE4 deacetylases for modification of chitosan and CHOS and on more efficient chitin conversion using enzymes derived from *S. marcescens* and *C. japonicus* are currently in progress.

SAMMENDRAG

I overgangen fra en fossil-basert til en bio-basert økonomi må bruken av fornybare ressurser utforskes. Kitin er, etter cellulose, ansett som den biomassen det fins mest av på jorden, og de vannløselige kitin-derivatene kitosan og kitooligosakkarider har mange applikasjoner innen eksempelvis medisin, jordbruk og matindustri. Ekstraksjonsprosessen av kitin fra kitin-rik biomasse og videre produksjon av kitosan og kitooligosakkarider involverer i dag farlige kjemikalier. Det er derfor ønskelig å erstatte dagens kjemiske prosess med en enzymdrevet prosess da dette vil være mer miljøvennlig. I tillegg kan enzymer brukes til å produsere godt definerte kitosaner og kitooligosakkarider, noe som er av interesse siden bioaktiviteten til disse forbindelsene er avhengig av egenskaper slik som fraksjon av acetylering, grad av polymerisering og acetyleringsmønster. Å undersøke hvilke proteiner mikroorganismer bruker når de vokser på kitin kan gi innsikt i naturlig kitin-nedbrytning og kan gi relevante enzymer som trengs for industriell valorisering av kitin-rik biomasse.

Artikkel I beskriver karakteriseringen av en karbohydrat esterase familie 4 (CE4) deacetylase, som ble valgt på grunn av sitt potensiale for bruk i produksjon av kitooligosakkarider med definert fraksjon av acetylering og acetyleringsmønster. En god forståelse av hvordan disse enzymene interagerer med sitt substrat og enzymenes spesifisitet er viktig for å utnytte enzymene på en best mulig måte. **Artikkel I** inkluderer det første enzym-substrat komplekset for en CE4 med et åpent aktivt sete, og gir verdifull informasjon om hvordan dette enzymet interagerer med sitt substrat. Enzymet kan deacetylere flere ulike substrater på ulike posisjoner. Den brede substratspesifisiteten og at substratet okkuperer tilsynelatende få substeder indikerer at det kan bli vanskelig å utvikle denne typen CE4 enzymer for å skreddersy acetyleringsmønsteret i kitooligosakkarider.

Et vidt spekter av karbohydrataktive enzymer kodes av genomet til *Cellvibrio japonicus*, inkludert flere kitinaser og andre enzymer som muligens er involvert i kitin nedbrytning. Når arbeidet med denne avhandlingen startet hadde det ikke blitt undersøkt om noen av disse enzymene faktisk er involvert i denne Gram-negative bakterien sin utnyttelse av kitin. **Artikkel II** utforsker de proteinene som *C. japonicus* sekreterer når den vokser på kitin, ved bruk av en ny plate-basert proteomikkmetode som ga sekretomprøver med relativt lav fraksjon av cytoplasmiske proteiner. Resultatene viser at de fire glykosyl hydrolase familie 18 (GH18) kitinasene som finnes i genomet til *C. japonicus* produseres i store mengder, noe

som indikerer at disse enzymene er involvert i nedbrytning av naturlig kitin. Forsøk med kitin-nedbrytning viste at *C. japonicus* har betydelig kitinolytisk kraft. Flere proteiner uten en åpenbar rolle i kitin-nedbrytning ble produsert i store mengder under vekst på kitin, og danner en liste med proteiner som i fremtiden kan utforskers for en mulig rolle i nedbrytning av kitinrik biomasse.

Artikkel III gir en fyldigere beskrivelse av GH18 kitinasene produsert av *C. japonicus*. Ved å slå ut kitinasegenene, ble det vist at en av kitinasene, *CjChi18D*, er helt avgjørende for at bakterien skal kunne utnytte kitin som karbonkilde. Biokjemisk karakterisering viste at *CjChi18D* er den mest effektive i nedbrytning av kitin, som kan forklare dens avgjørende rolle. Sammenligning av de fire enzymene indikerte at de har ulike og antageligvis komplementære funksjoner, eksemplifisert med *CjChi18C* som har høyest aktivitet mot kitoheksaose. Ved å kombinere enzymene så ble synergistiske effekter i kitin-nedbrytning observert. Transkriptomikk analyser viser at de fire GH18 kitinasene og en kitin-aktiv lytisk polysakkarid monooxygenase, *CjLPMO10A*, er oppregulert når *C. japonicus* vokser på kitin. Det samme er mange andre antatte kitinaktive enzymer, i tillegg til noen få proteiner med ukjent funksjon, som er oppregulert i noe mindre grad.

Serratia marcescens innehar ett av de mest studerte kitinolytiske maskineriene, som involverer tre kitinaser, en lytisk polysakkarid monooxygenase og en kitobiase. Genomet til en av de best studerte *S. marcescens* familiene hadde imidlertid ikke blitt sekvensert da arbeidet med denne avhandlingen startet, og viktigheten av eventuelt andre proteiner i kitin-nedbrytningen hadde ikke blitt undersøkt. Dataene presentert i **Artikkel IV** inkluderer genomsekvensen til *S. marcescens* BJJ200 og en proteomikk-basert undersøkelse av hvilke proteiner som sekreteres når bakterien vokser på kitin. Genomsekvensen viste at genomet inneholder en fjerde kitinase, *SmChiD*, men proteomikken antyder at denne kitinasen ikke er viktig i utnyttelsen av kitin. En biokjemisk karakterisering av *SmChiD* støttet antydningene om at dette enzymet ikke er viktig i kitin-nedbrytning, og en annen, hittil ukjent, biologisk rolle for denne kitinasen antas.

Oppsummert så gir dataene presentert i denne avhandlingen innsikt i kitinaktive enzymer produsert av bakterier. **Artikkel I** gir verdifull innsikt i hvordan CE4 deacetylaser med et åpent aktivt sete binder sitt substrat. **Artikkel II-IV** bekrefter at kitinaktive enzymer, spesielt hydrolaser, har viktige roller i naturlig nedbrytning av kitin. I tillegg gir **Artikkel II-IV** en

liste med proteiner uten en åpenbar rolle i kitin-nedbrytning. Disse proteinene blir dratt frem som mulig mål for videre undersøkelse av deres potensielle rolle i nedbrytning av kitin-rik biomasse. Videre studier på tilpasning av CE4 deacetylasen for modifisering av kitosan og kitooligosakkarider, og på mer effektiv kitin-nedbrytning ved bruk av enzymer fra *S. marcescens* og *C. japonicus* pågår.

ABBREVIATIONS

AA	Auxiliary activity
<i>Ar</i>	<i>Arthrobacter</i>
CAZyme	Carbohydrate-active enzyme
CBM	Carbohydrate-binding module
CE	Carbohydrate esterase
Chi	Chitinase
CHOS	Chitooligosaccharides
<i>Cj</i>	<i>Cellvibrio japonicus</i>
<i>Cl</i>	<i>Colletotrichum lindemuthianum</i>
DP	Degree of polymerization
e.g.	For example
F _A	Fraction of acetylation
GH	Glycosyl hydrolase
GlcN	Glucosamine
GlcNAc	<i>N</i> -Acetylglucosamine
i.e.	That is
LFQ	Label free quantification
LPMO	Lytic polysaccharide monooxygenase
MS	Mass spectroscopy
P _A	Pattern of acetylation
PUL	Polysaccharide utilization loci
Sec	SecYEG translocon
<i>Sm</i>	<i>Serratia marcescens</i>
<i>Sp</i>	<i>Serratia proteamaculans</i>
SpI	Signal peptidase I signal peptide
SpII	Signal peptidase II signal peptide
TxSS	Type x secretion system (where x ranges from 1 to 9)
Tat	Twin-arginine translocon
<i>Tk</i>	<i>Thermococcus kodakaraensis</i>
<i>Vc</i>	<i>Vibrio cholerae</i>

LIST OF PAPERS**Paper I**

Tuveng, T. R., Rothwiler, U., Udata, G., Vaaje-Kolstad, G., Smalås, A. & Eijsink, V. G. H. 2017. Structure and function of a CE4 deacetylase isolated from a marine environment. *Submitted to PlosOne*.

Paper II

Tuveng, T. R., Arntzen, M. Ø., Bengtsson, O., Gardner, J. G., Vaaje-Kolstad, G. & Eijsink, V. G. H. 2016. Proteomic investigation of the secretome of *Cellvibrio japonicus* during growth on chitin. *Proteomics*, 16, 1904-1914.

Paper III

Monge, E., **Tuveng, T. R.**, Vaaje-Kolstad, G., Eijsink, V. G. H. & Gardner, J. G. 2017. Chitin degradation by *Cellvibrio japonicus* is dependent on the non-redundant CjChi18D chitinase. *Manuscript*.

Paper IV

Tuveng, T. R., Hagen, L. H., Mekasha, S., Frank, J., Arntzen, M. Ø., Vaaje-Kolstad, G. & Eijsink, V. G. H. 2017. Genomic, proteomic and biochemical analysis of the chitinolytic machinery of *Serratia marcescens* BJL200. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1865, 414-421.

Other publications by the author

Larsbrink, J., **Tuveng, T. R.**, Pope, P. B., Bulone, V., Eijsink, V. G. H., Brumer, H. & Mckee, L. S. 2017. Proteomic insights into mannan degradation and protein secretion by the forest floor bacterium *Chitinophaga pinensis*. *Journal of Proteomics*, 156, 63-74.

Liu, Z., Gay, L. M., **Tuveng, T. R.**, Agger, J. W., Westereng, B., Mathiesen, G., Horn, S. J., Vaaje-Kolstad, G., Van Aalten, D. M. F. & Eijsink, V. G. H. 2017. Structure and function of a broad-specificity chitin deacetylase from *Aspergillus nidulans* FGSC A4. *Scientific Reports*, 7, 1746.

1 INTRODUCTION

Today there is a worldwide focus on shifting from a fossil-based economy towards a bio-based economy. This implies a need to replace products derived from fossil resources, such as oil, with products produced using renewable resources. Biomasses of different origins represent important renewable resources, which can be used to produce fuels, chemicals and other products. Cellulose is the most abundant biomass on Earth, being a main component in the plant cell wall. So called “first generation feedstocks”, including corn and sugar cane used for ethanol production, and rapeseed oil used for biodiesel production, are the main resources in the bio-based economy today. Although renewable, the use of first generation feedstocks raises some issues, as their use for producing e.g. fuels competes with their use as food or feed (Williams, 2008). Therefore, in recent years, research focus has shifted towards non-edible feedstocks, often referred to as “second generation feedstocks”. The most common of these feedstocks is lignocellulosic biomass derived from plants and trees. Other non-edible feedstocks are also in focus, such as algae (Bibi et al., 2016) and chitin-rich biomasses (Aranda-Martinez et al., 2017), but their exploration has not yet advanced to the level reached for lignocellulose.

Chitin is the second most abundant polysaccharide on Earth and is, like cellulose, a recalcitrant, insoluble polysaccharide found in the exoskeletons of arthropods and in fungal cell walls, with an estimated annual production of around 10^{11} tons (Kurita, 2006). Although crab and shrimp shell waste from the seafood industry is used for commercial production of chitin and its derivative chitosan (see below), much of this waste is deposited, meaning that valuable biomass is wasted (Kandra et al., 2012). The current process for chitin extraction from crustaceans and production of down-stream products involves harsh and not environmentally friendly chemicals, which is needed to remove minerals and proteins. One way to achieve more environmentally friendly and sustainable utilization of chitin-rich biomass is to replace one or more of the chemical processing steps with enzyme-based processes. This thesis deals with the identification and characterization of enzymes for chitin processing.

1.1 Chitin

Chitin is a linear homopolymer consisting of β -1,4 linked *N*-acetylglucosamine [GlcNAc (Fig. 1a)]. Chitin chains assemble into crystalline structures and different polymorphic forms exist

depending on the arrangement of individual chains in the chitin fiber. The different polymorphic forms are named α -chitin, β -chitin, and γ -chitin, [Fig. 1b, (Carlström, 1957, Rudall, 1963)], with α -chitin being most recalcitrant. In Nature α -chitin is the most abundant form, found in fungal cell walls and in the exoskeletons of crustaceans and insects (Rinaudo, 2006). Beta-chitin is found in the gladius of squids (also known as the squid pen) (Blackwell, 1969), while γ -chitin has been identified in cocoon fibers of the *Ptinus* beetle and in the stomach of the squid *Loligo* (Jang et al., 2004). The chitin chains are organized in sheets, which in α -chitin are held together by inter- and intramolecular hydrogen bonds making it more rigid than β -chitin, which lacks inter-sheet hydrogen bonds (Rudall, 1963, Rinaudo, 2006). Due to the absence of intermolecular hydrogen bonds, β -chitin is more loosely packed and is more susceptible to swelling by accommodating various polar molecules (Saito et al., 2000, Saito et al., 2002, Rinaudo, 2006).

In the exoskeleton of crustaceans, chitin exists in complex with proteins and minerals (mainly calcium carbonate; Fig. 1c). This composite material consist of 15-40 % chitin, 20-40 % protein and 20-50 % calcium carbonate. Extraction of chitin from this composite material requires the removal of proteins and calcium carbonate. This is commonly achieved by the use of concentrated sodium hydroxide and hydrochloric acid, respectively. There are several drawbacks in these chemical methods: the chemicals are hazardous, the process destroys the protein fraction, which represents a potentially valuable co-product, and the chemicals come with a cost. Alternative and more environmentally friendly extraction processes have been reported, using enzymes (Younes et al., 2014, Younes et al., 2016) or microbial fermentation (Bajaj et al., 2015). These latter processes may preserve both the protein and calcium carbonate, but have so far only been explored in laboratory scale (Kaur and Dhillon, 2015). Although enzymatic deproteinization using proteases is promising, this process is inferior to the chemical process, as the enzymatic process leaves 5-10 % residual protein in the chitin. For many applications this is not a huge problem, but for biomedical applications complete protein removal is crucial, as shellfish allergy in humans is caused by the protein components (Younes and Rinaudo, 2015).

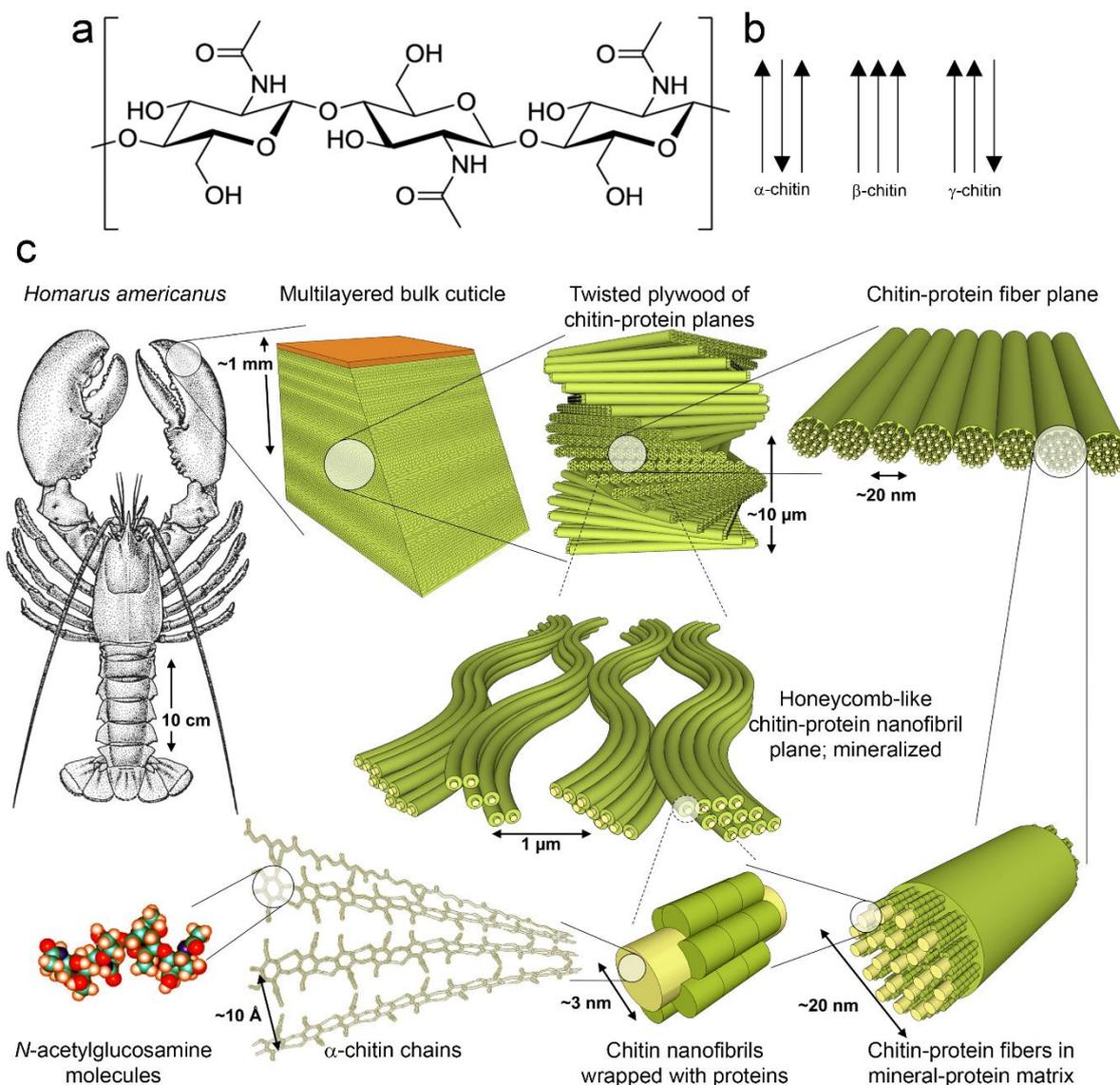


Figure 1. Chitin structure and organization. (a) Chemical structure of chitin showing β -1,4 linked *N*-acetylglucosamine units. (b) The different polymorphic forms of chitin. (c) Example of how chitin is packed in the exoskeleton of the lobster *Homarus americanus*. Figure adapted from Nikolov et al. (2011).

1.2 Chitosan and chito oligosaccharides

Chitin is commonly used to produce chitosan and chito oligosaccharides (CHOS). Chitosan is the deacetylated form of chitin, i.e. a form where the acetyl group on the C2 amino group has been removed. Thus, chitosan contains glucosamine (GlcN) rather than *N*-acetylglucosamine (Fig. 2). The term “chitosan” is used for a collection of chitin-derivatives, ranging from fully deacetylated to approximately 35 % deacetylated. In other words, the fraction of acetylation (F_A) may vary from 0 to approximately 0.65. The key property that defines a chitin-based polymer as chitosan is solubility in mildly acidic solutions. Chitosan is not commonly found

in Nature, but is produced by a few fungal species (Pochanavanich and Suntornsuk, 2002). Production of chitosan from chitin for industrial purposes is done using homogeneous or heterogeneous deacetylation processes (Vårum et al., 1991a, Vårum et al., 1991b). In the homogeneous process, the chitin is dissolved in an alkaline solution at low temperature under extensive stirring, while in the heterogeneous process the chitin is kept insoluble by adding a hot alkali solution (Aam et al., 2010).

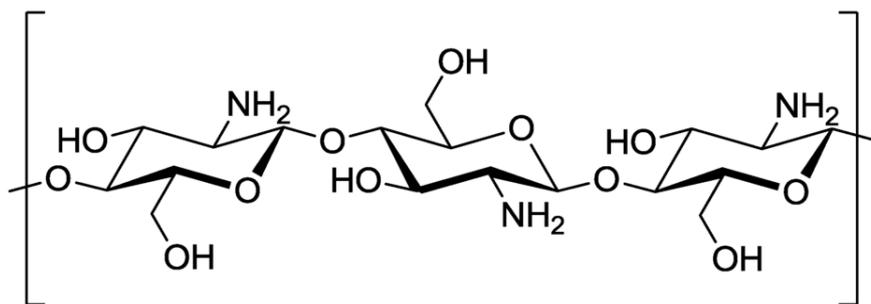


Figure 2. Chemical structure of chitosan. The figure shows β -1,4 linked glucosamine units, but it should be noted that chitosans normally also contain *N*-acetylglucosamine and that the fraction of acetylation (F_A) can be as high as 0.65.

Chitosan is a unique polymer since glucosamine has a pK_a value around 6.5, thus adapting a positive charge at mildly acidic pH and thereby yielding a poly-cationic macromolecule. Chitosan is produced industrially in the thousands of tons per year scale and has numerous applications (Kim, 2010), for example in medicine and agriculture (Table 1). This is due to the biocompatible and biodegradable nature of the chitosan [and CHOS (Aam et al., 2010, Kong et al., 2010, Anitha et al., 2014)]. Chitosan has a proved antimicrobial effect that is better compared to other disinfectants. The mechanism for this antimicrobial effect is not fully understood, but one hypothesis for the antibacterial effect is that chitosan change the permeability of the cell membrane due to interactions between the positively charged chitosan and the negatively charged components in the cell membrane (Younes and Rinaudo, 2015). The antifungal effect of chitosan towards *Rhizopus stolonifer* (Ehrenb.:Fr.) Vuill, causing rhizopus rots disease of various fruits and vegetables, was shown to be dependent on the molecular weight of the chitosan. Low molecular weight chitosan inhibited the mycelial growth, while high molecular weight chitosan had more effect on the development of spores (Hernández-Lauzardo et al., 2008). Chitin, chitosan, and CHOS are known to improve the wound healing, and Minagawa et al. (2007) found that the effect in the wound healing process

was affected by the F_A and the degree of polymerization (DP, also referred to as chain length). This shows that there is a need to control these parameters when making chitosan and CHOS.

Table 1. Applications of chitin and its derivatives chitosan and CHOS.

Field	Examples of usage	Selected references
Pharmaceuticals	Drug delivery	Agnihotri et al. (2004)
	Antimicrobial	El Hadrami et al. (2010)
Agriculture	Plant promotor	Winkler et al. (2017)
	Wastewater treatment	Raval et al. (2016) Wawrzkievicz et al. (2017)
Food	Stabilizer	Dickinson (2017)
	Packing	Marsh and Bugusu (2007)
Biomedical	Wound healing	Cho et al. (1999) Minagawa et al. (2007)
	Tissue engineering	Freier et al. (2005)
Textiles	Coating	Ye et al. (2005)
Cosmetics	Skin care	Gautier et al. (2008)

Table adapted from Hamed et al. (2016).

Enzymatic depolymerization of chitin and chitosan generates CHOS, which, depending on the starting material will be homo- or hetero-oligosaccharides of GlcNAc (A) and GlcN (D). Such CHOS are thought to have a wide variety of bio-activities, such as antifungal and antitumor activity (Nam et al., 2007, El Hadrami et al., 2010). The biochemical properties of chitosans and CHOS are not only affected by the F_A , but also by the DP, the pattern of acetylation (P_A ; also referred to as “sequence”), and the molecular weight distribution (PD; PolyDispersity). The chemical methods commonly used for production of chitosan give a random distribution of the GlcN and GlcNAc units, however for some applications it is desirable to have chitosans and CHOS with a well-defined P_A (Aam et al., 2010). It has been proposed that chitin deacetylases could be used for converting chitin and fully acetylated CHOS to chitosan and CHOS with varying F_A . These deacetylases specifically remove the acetyl-groups by a hydrolytic reaction (Tsigos et al., 2000, Hamer et al., 2015, Hamed et al., 2016) and if deacetylases with sequence specificities could be found or developed, one might even be able to generate defined P_A 's. By combining deacetylases with chitinases and chitosanases having different substrate specificities it is, at least in theory, possible to produce CHOS with specific DP, F_A and P_A . It should be noted though that, so far, enzymatic

conversion of chitin to chitosan has not been accomplished and that there are only very few examples of sequence specific deacetylation of CHOS (Andrés et al., 2014, Hamer et al., 2015).

1.3 Carbohydrate-active enzymes

Nature is full of carbohydrates, in monomeric, oligomeric and polymeric forms, and different enzymes have evolved to synthesize, modify or degrade complex carbohydrate structures. Carbohydrate-active enzymes (CAZymes) are organized and classified in the CAZy database, where they are divided into different classes, according to function, and families, according to amino acid sequence similarity [www.cazy.org, (Lombard et al., 2014)]. Currently the CAZy database contains five enzyme classes and one class of non-catalytic modules that are associated with carbohydrate-active enzymes. The glycosyl hydrolase (GH) class contains enzymes that hydrolyze glycosidic linkages. Some GH enzymes also have transglycosylating activity, where a sugar (instead of water, as in hydrolysis) acts as an acceptor resulting in the formation of a new glycosidic bond (Cantarel et al., 2009). Glycosyl transferases (GT) synthesize glycosidic linkages using activated sugars, while polysaccharide lyases (PL) perform non-hydrolytic cleavage of glycosidic bonds. Carbohydrate esterases (CE) remove ester modifications on carbohydrates (Cantarel et al., 2009). The fifth class is referred to as auxiliary activities (AA) and contains a variety of redox enzymes acting in conjunction with other CAZymes (Levasseur et al., 2013). The AA class includes the so-called lytic polysaccharide monooxygenases (LPMOs; see Section 1.3.1.3), that were discovered in 2010 (Vaaje-Kolstad et al., 2010) and that play a pivotal role in polysaccharide degradation. The AA class also contains redox enzymes acting on lignin, since lignin is found together with polysaccharides in plant cell walls. In addition, lignin facilitates the activity of LPMOs (Levasseur et al., 2013). Next to these five classes of catalytic domains, the carbohydrate-binding module (CBM) class contains proteins with no enzymatic activity. CBMs are normally covalently attached to other enzymes, and their primary function is to promote substrate binding (Cantarel et al., 2009).

1.3.1 CAZymes in chitin degradation and modification

1.3.1.1 Chitinases

Chitinases, enzymes that hydrolyze the β -1,4 glycosidic bonds in chitin chains, are found in GH families 18 and 19 (Henrissat and Bairoch, 1993). GH18s have a $(\beta/\alpha)_8$ -barrel fold (Perrakis et al., 1994, van Scheltinga et al., 1994), while GH19 chitinases mainly comprise α -

helices (Hart et al., 1995). The two enzyme types use different catalytic mechanisms; GH18s use a substrate-assisted mechanism that is retaining (Tews et al., 1997), the latter meaning that the anomeric configuration is retained, while GH19s use a classical inverting mechanism, i.e. the anomeric configuration is inverted (Davies and Henrissat, 1995). Bond cleavage by inverting enzymes proceeds through a single-displacement mechanism, contrary to the double displacement mechanism in retaining enzymes. Both mechanisms use a general acid catalysis and require a pair of carboxylic acids. One carboxylic acid acts as the catalytic acid, whereas the other acts as a water-activating base in the inverting mechanism or as a nucleophile in the retaining mechanism. The retaining mechanism entails the formation of an intermediate that is hydrolyzed by a water, activated by the deprotonated catalytic acid. Both reactions proceed through oxacarbenium-ion-like transition states (Rye and Withers, 2000). GH18 family chitinases use a special version of the retaining mechanism, namely a substrate-assisted mechanism (Fig. 3), where the oxygen of the *N*-acetyl group in the substrate acts as a nucleophile, leading to the formation of an oxazolinium ion intermediate. In GH18 enzymes the catalytic acid is a glutamate, which acts as a base in the second half of the reaction. This glutamate is located in a diagnostic DXDXE sequence motif that occurs in all active GH18 chitinases (van Aalten et al., 2001, Gloster and Davies, 2010, Vaaje-Kolstad et al., 2013).

Chitinases degrade chitin chains from one of the ends (exo-mechanism) or from a random point on the chain (endo-mechanism). In addition, the endo- or exo-activity can be combined with processivity, meaning that the enzyme stays attached to the substrate after a successful cleavage. The enzyme thus slides along the chitin chain making several successive cleavages before it detaches from the substrate (Davies and Henrissat, 1995, Horn et al., 2006b). Processive and non-processive enzymes have been studied in detail revealing both sequence and structural differences that underlay processivity and its direction (Horn et al., 2006b, Zakariassen et al., 2009, Payne et al., 2012)

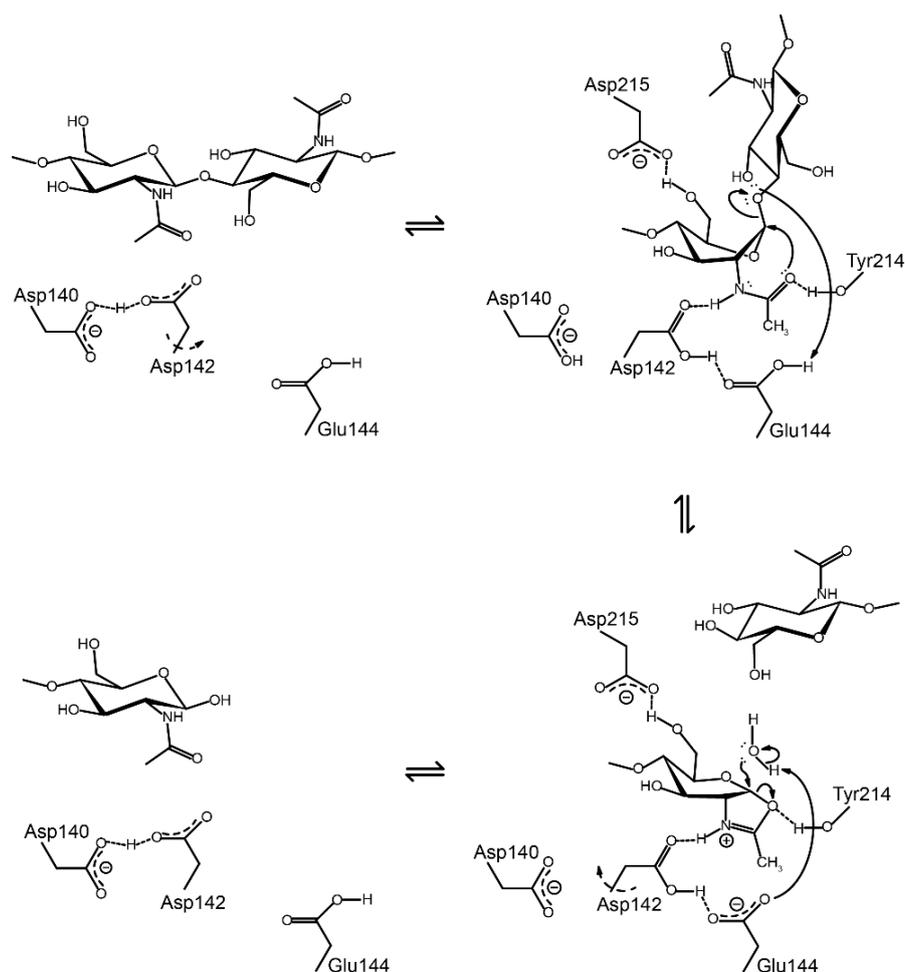


Figure 3. The substrate-assisted mechanism used by GH18 chitinases. Binding of the substrate leads to distortion of the sugar bound in subsite -1 towards a boat conformation. Simultaneously, Asp142 rotates and forms hydrogen bonds with the catalytic Glu144 and the acetamido group of the sugar in the -1 subsite. At this point, Glu144 acts as a general acid by protonating the glycosidic oxygen, which supports leaving group departure (i.e. cleavage of the glycosidic bond) that is further promoted by nucleophilic attack of the acetamido group on the anomeric carbon, forming an oxazolinium ion intermediate. Glu144 then acts as a general base, activating an incoming water molecule that hydrolyzes the oxazolinium ion. The product is released from the active site and Asp142 rotates back to its original conformation. Notably, if the water molecule hydrolyzing the oxazolinium ion is outcompeted by another acceptor, transglycosylation will occur instead of hydrolysis (Williams and Withers, 2000). Amino acid numbering is based on chitinase B (*SmChiB*) from *Serratia marcescens*. The figure was taken from Vaaje-Kolstad et al. (2013).

During growth on chitin, the well-known chitinolytic Gram-negative bacterium *Serratia marcescens* produces mainly three chitinases (*SmChiA*, *SmChiB*, and *SmChiC*), of which *SmChiA* and *SmChiB* are processive exo-enzymes, while *SmChiC* is a non-processive endo-enzyme. The structures (Fig. 4) of these proteins show that *SmChiA* and *SmChiB* have deep

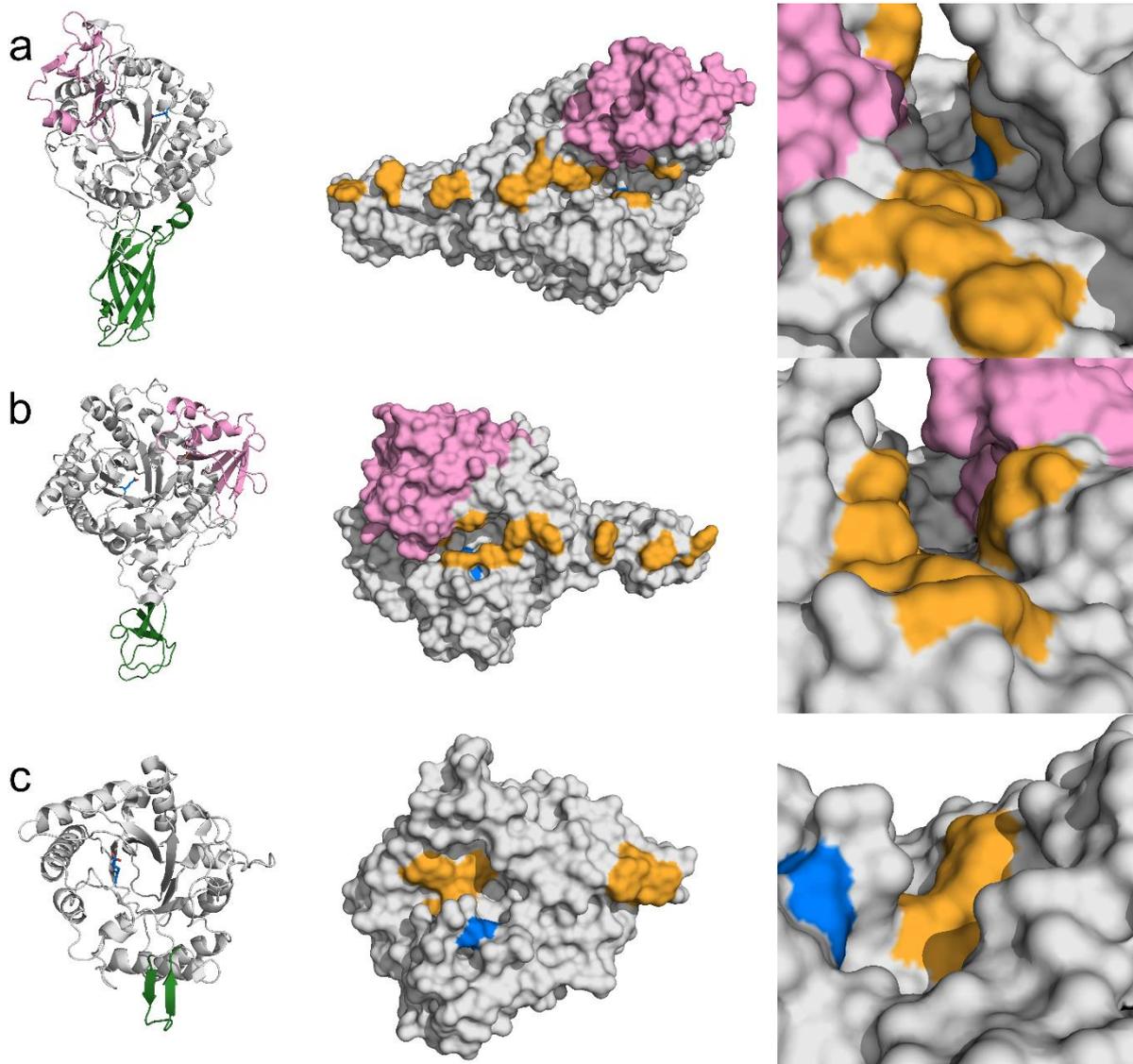


Figure 4. Structures of *S. marcescens* chitinases. The left figures show (a) *SmChiA* (PDB id 1CTN), (b) *SmChiB* (PDB id 1E15), and (c) *SmChiC* (PDB id 4AXN) in cartoon representation. The $\alpha+\beta$ domain present in *SmChiA* and *SmChiB* is shown in pink. Extra domains that promote substrate binding in *SmChiA* (FnIII) and *SmChiB* (CBM5) are shown in green, as is a small β -hairpin domain in *SmChiC* possibly aiding substrate binding. Note that *SmChiC* has an FnIII and a CBM5/12 domain, but structural data for these domains are not available. The catalytic Glu is shown as sticks with blue carbons. The figures in the middle show the chitinases in surface representation, with aromatic residues lining the active site cleft (Vaaje-Kolstad et al., 2013) highlighted in orange, and the catalytic Glu in blue. The right figures show the differences in the depth of the active site. The figure was made using PyMol (Schrödinger, 2015).

active site clefts, contrary to the shallow and open active site in *SmChiC*. The deep clefts of the two processive enzymes are defined by several loops, and a small sub-domain that occurs

only in a subset of GH18 enzymes (including *SmChiA* and *SmChiB*, but not *SmChiC*) and that has been named the $\alpha+\beta$ domain (Perrakis et al., 1994). Aromatic residues lining the surface of the active site cleft form another characteristic feature of processive enzymes. One of these aromates is next to the highly conserved SXGG sequence, which is followed by a Trp in processive enzymes (Payne et al., 2012). These aromatic residues likely help the enzyme to stay attached to the substrate as it moves along the chain. Horn et al. (2006a) and Zakariassen et al. (2009) mutated several of the aromatic residues in *SmChiB* and *SmChiA*, respectively, to non-aromatic residues, and showed that some of these mutations almost abolished processivity while having no detrimental effect on catalytic activity as such.

It is worth noting that chitinases in families GH18 and GH19 also act on chitosan, as documented in various studies (Sørbotten et al., 2005, Horn et al., 2006a, Zakariassen et al., 2009). Expectedly, the efficiency of the enzymes decreases as the F_A decreases. Furthermore, due to the substrate-assisted nature of the catalytic mechanism, GH18 enzymes only cleave after an acetylated sugar.

Experimental determination of the processivity and the endo- or exo-nature of a chitinase is challenging. For GH18 chitinases, insight can be obtained from studies with highly acetylated chitosan as shown by Horn et al. (2006b). When using water-soluble chitosans, processive *SmChiA* and *SmChiB* mainly produce even-numbered oligomers, while *SmChiC* produces equal amounts of even- and odd-numbered CHOS (Sørbotten et al., 2005, Horn et al., 2006b, Sikorski et al., 2006). The production of even-numbered oligomers by processive chitinases is expected, as an *N*-acetyl group in subsite -1 is essential for catalysis and the repetitive unit in chitin and chitosan is a dimer [Fig. 1 and 2 (Vaaje-Kolstad et al., 2013)]. Initial productive binding of the substrate to the enzyme will yield products of any length. If the enzyme is processive, it will move by two sugars at the time, until a new productive complex is formed, meaning that any further products resulting from the same initial enzyme-substrate association will be even-numbered. Non-processive enzymes will detach and rebind in between each reaction, thus yielding a continuum of product lengths. The exo- or endo-activity of a chitinase can be determined by measuring the reduction of viscosity during reactions with chitosan. An endo-enzyme, cutting randomly along the chitosan chain, will lead to fast reduction of viscosity. To the contrary, an exo-enzyme, cutting from the chain ends, will lead to slow reduction of viscosity (Sikorski et al., 2006).

1.3.1.2 *β -N-acetylhexosaminidases*

The most dominant product arising from chitin degradation by GH18 and GH19 chitinases is (GlcNAc)₂. To further convert (GlcNAc)₂ to GlcNAc, most chitinolytic enzyme systems contain a GH20 β -N-acetylhexosaminidase (also known as chitobiase), but other enzymes performing similar reactions are also known (e.g. in families GH3 and GH84). The chitobiase cleaves off the non-reducing end sugar of CHOS, using a catalytic mechanism similar to the substrate-assisted mechanism used by chitinases (Drouillard et al., 1997, Tews et al., 1997). Since most chitinases yield (GlcNAc)₂ as their primary product, this also represents the primary substrate for chitobiases. However, it is well known that chitobiases are capable of efficiently removing GlcNAc residues from the non-reducing end of longer CHOS (Drouillard et al., 1997).

1.3.1.3 *Lytic polysaccharide monooxygenases*

In 2005 Vaaje-Kolstad and colleagues showed that a chitin-binding protein produced by *S. marcescens*, named *SmCBP21* and originally classified as a CBM33, (see Section 1.3.1.6) contributed to chitin degradation by strongly boosting chitin solubilization by chitinases (Vaaje-Kolstad et al., 2005). At the time, *SmCBP21* was assumed to have no catalytic activity, but rather act as a “helper protein” (e.g. “substrate-disrupting protein”) for chitinases in the chitin degradation process. However, in 2010 it was shown that *SmCBP21* is a member of a family of enzymes capable of cleaving chitin and other polysaccharides, including cellulose, by an oxidative mechanism (Vaaje-Kolstad et al., 2010). The reaction products of these enzymes contain a single oxygen obtained from the dioxygen co-substrate (Vaaje-Kolstad et al., 2010), and, therefore, the enzymes were named lytic polysaccharide monooxygenases (Horn et al., 2012b).

The most remarkable feature of these oxidative enzymes is their ability to cleave polysaccharide chains that are embedded in a crystalline environment, something that is both sterically and energetically difficult for the canonical hydrolytic enzymes such as chitinases and cellulases. By making nicks on the surface of the polysaccharide crystals, LPMOs likely disrupt the crystal surfaces and provide attachment points for the hydrolytic enzymes, which explains the synergistic effects that are observed when combining LPMOs and hydrolytic enzymes (Vaaje-Kolstad et al., 2010, Nakagawa et al., 2013, Paspaliari et al., 2015, Nakagawa et al., 2015). LPMOs cannot use the catalytic power of substrate distortion, as e.g. GH18 chitinases do (see Section 1.3.1.1), but use instead powerful redox chemistry facilitated by a

catalytic centre that contains a copper ion coordinated by two fully conserved histidine residues and the N-terminal amino group of one of these histidines (Figure 5).

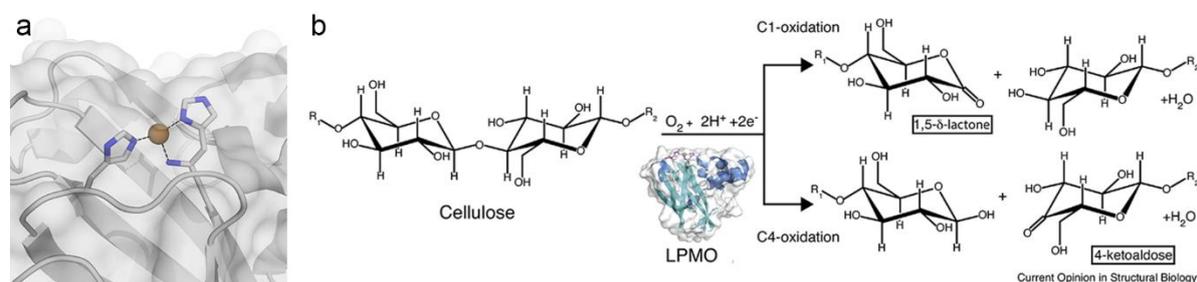


Figure 5. Catalytic center and reaction mechanism of LPMOs. (a) The catalytic center in LPMOs, exemplified by CjLPMO10A from *Cellvibrio japonicus* (PDB id 5FJQ) showing the T-shaped coordination (called the histidine brace) of the Cu-ion (brown sphere). The two conserved histidines, coordinating the Cu-ion, are shown as sticks. The figure was made using PyMol (Schrödinger, 2015). (b) Reaction mechanism of LPMOs, exemplified with cellulose as substrate, showing the possible C1 and C4 oxidized products (see text for details). Figure taken from Vaaje-Kolstad et al. (2017).

The discovery of LPMOs led the CAZY team to create the auxiliary activity class (see Section 1.3), which today contains 13 families. LPMOs are categorized in AA families 9, 10, 11, and 13, and are found in several organisms, including bacteria, fungi, viruses and higher eukaryotes like insects (Levasseur et al., 2013). Chitin-active LPMOs are found in AA families 10 and 11, and so far, chitin-active fungal LPMOs are only described in AA family 11.

LPMO activity was originally discovered for chitin but it was immediately obvious that similar enzymes acting on cellulose would exist, in particular enzymes that were at the time erroneously classified as GH61 (Vaaje-Kolstad et al., 2010). Since the discovery of the first chitin-active LPMOs, LPMOs acting on cellulose have gained much attention due to their industrial relevance (Forsberg et al., 2011, Phillips et al., 2011, Quinlan et al., 2011, Westereng et al., 2011, Beeson et al., 2015). LPMOs show different oxidative regioselectivities and these differ between chitin-active and cellulose-active LPMOs. While only C1 oxidizing chitin-active LPMOs have been described, cellulose-active LPMOs can be strictly C1 oxidizing, strictly C4 oxidizing, or be able to oxidize both the C1 and C4 position (Vaaje-Kolstad et al., 2017). Independent of the regioselectivity, LPMOs depend on copper (Phillips et al., 2011, Quinlan et al., 2011), reducing equivalents to reduce this copper, and dissolved molecular dioxygen in order to perform catalysis. However, in a recent publication

Bissaro et al. (2017) question the validity of dioxygen being the LPMO co-substrate and provide compelling evidence indicating that hydrogen peroxide (H₂O₂) is the true co-substrate of LPMOs. Notably, LPMOs can themselves generate H₂O₂ from O₂ (Kittl et al., 2012).

1.3.1.4 Carbohydrate esterases

Carbohydrate esterases (CEs) are enzymes catalyzing the *O*- or *N*-deacetylation of substituted saccharides, i.e. esters or amides in which sugars play the role of alcohol and amine, respectively. Enzymes that hydrolyze esters in which sugars play the role of acid are also considered as CEs (Biely, 2012). Of the 16 CE families known to date (September 2017) CEs deacetylating chitin and its derivatives are only found in CE families 4 and 14 [www.cazy.org (Lombard et al., 2014)]. The CE14 family only contains a few characterized chitin deacetylases, all being archaeal, deacetylating the non-reducing end of (GlcNAc)₂ as part of the chitinolytic pathway of the organism [see Section 1.4.3 (Tanaka et al., 2004, Mine et al., 2014)]. The CE4 family comprises several bacterial and eukaryotic esterases that deacetylate GlcNAc units in peptidoglycan, chitin, and CHOS. The CE4 family also contains enzymes capable of removing *O*-linked acetyl groups from acetyl xylan (Biely et al., 1996) and several family members are known to act both on xylan (*O*-deacetylation) and chitin [*N*-deacetylation (Caufrier et al., 2003, Puchart et al., 2006, Tang et al., 2011, Liu et al., 2017)]. The activity and structural features of CE4 deacetylases are discussed in more detail below.

1.3.1.4.1 Carbohydrate esterase family 4

The first deacetylase acting on chitin was found in extracts from the fungus *Mucor rouxii* (Araki and Ito, 1975). However, a few years earlier, Araki and co-workers described an enzyme from *Bacillus cereus* that deacetylates GlcNAc units in peptidoglycan (Araki et al., 1971). Both these enzymes are today classified into carbohydrate esterase family 4. CE4 deacetylases removing the *N*-acetyl group from GlcNAc units share five conserved sequence motifs: motif 1, T(F/x)DD; motif 2, H(S/T)xxH; motif 3, R(P/x)PY; motif 4, (Dxx)D(W/Y); motif 5, LxH (Blair et al., 2005). Blair et al. (2005) first proposed a catalytic mechanism for family CE4 deacetylases, which, notably, depends on a bound metal ion, preferably zinc or cobalt (Blair et al., 2005, Taylor et al., 2006, Andrés et al., 2014). They suggested a general acid/base reaction mechanism based on an extensive structural analysis (Fig. 6). In this catalytic cycle, the catalytic base (the first Asp in motif 1) activates a metal-bound water molecule, which subsequently performs a nucleophilic attack on the carbon in the scissile C-N bond, creating a tetrahedral oxyanion intermediate. The metal ion and the backbone

nitrogen of the tyrosine in motif 3 stabilize the negative charge on the carbonyl oxygen. The catalytic acid (His in motif 5) protonates the nitrogen in the substrate, generating a free amine on the deacetylated product and leading to release of acetate (Blair et al., 2005, Andrés et al., 2014).

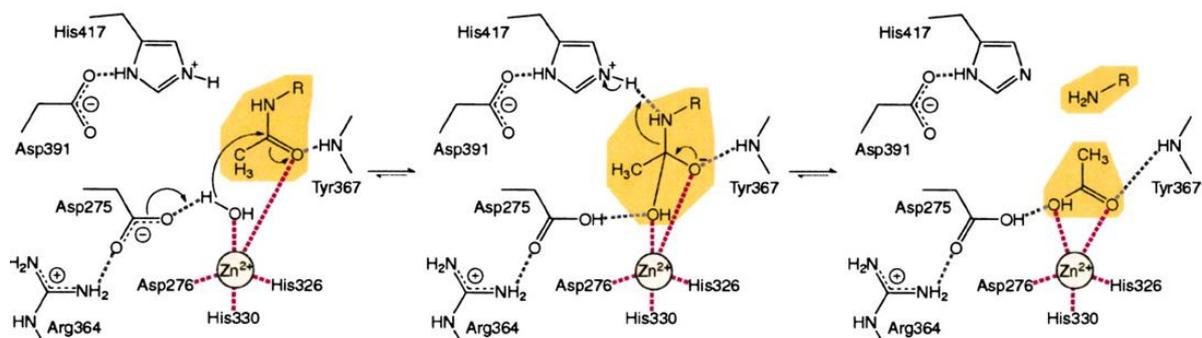


Figure 6. Reaction mechanism of CE4 deacetylases. The figure shows a proposed acid/base reaction mechanism for CE4 deacetylases that remove *N*-acetyl groups. The amino acid numbering is based on *SpPgdA* and the figure was taken from Blair et al. (2005). See the main text for detailed description of the catalytic mechanism.

Motif 1, 2, and 5 are highly conserved between different deacetylases, while motif 3 and 4 display more sequence variation (Fig. 7). The Asp-His-His metal binding triad is located in motif 1 and 2 (Fig. 7 and 8a). Motif 3 and 4 form one side of the active site groove each [Fig. 8a (Blair et al., 2005, Andrés et al., 2014)].

As shown in Fig. 7, some CE4 members have big insertions, representing loops that are located close to the active site (Fig. 8b). Andrés et al. (2014) proposed a “subsite capping model” involving six loops (indicated in Fig. 7) that cap the edges of the active site cleft. Such loops would contribute to substrate specificity and could endorse the deacetylase with sequence specificity, because they could define which substrates that can bind to the enzymes and which GlcNAc unit in the substrate that becomes deacetylated. The family CE4 representative, *VcCDA*, from *Vibrio cholera*, is special in that these loops are particularly long and form a buried active site (Fig. 8b). It was further suggested that these loops may rearrange depending on the length of the substrate (Andrés et al., 2014). Indeed, *VcCDA* is a highly specific enzyme that is restricted to deacetylate the GlcNAc next to the non-reducing end in CHOS. It is important to note that most other deacetylases in this family, including most other structurally characterized ones (Fig. 7), have shorter loops and, hence, more open

The GlcNAc residue to be deacetylated binds in subsite 0, and apart from the interactions between the enzyme and substrate in subsite 0, little experimental evidences for interactions between the enzyme and substrate in other subsites exist. The crystal structure of *VcCDA* in complex with (GlcNAc)₂ and (GlcNAc)₃ was determined by Andrés et al. (2014), representing the only structures of CE4 enzymes in complex with a relevant substrate. The dimer occupies subsite 0 and -1, while the trimer occupies subsite -1 to +1, with the non-reducing end in subsite -1 in both cases. *VcCDA* makes several interactions with the sugar in subsite -1, and binding of a sugar in subsite +1 requires rearrangement of several loops to allow its binding. As pointed out above and illustrated in Fig. 8, the structure of *VcCDA* is very different compared to other CE4 enzymes with known structure. Blair et al. (2006) performed a docking of (GlcNAc)₃ bound in subsite -1 to +1 in *CiCDA*, which indicated that there are no interactions between the enzyme and the sugar bound in subsite -1. This is different from what Andrés et al. (2014) found for *VcCDA*, underlining the need to obtain enzyme-substrate complexes with CE4 proteins with an open active site to get a deeper understanding of the substrate binding of these enzymes.

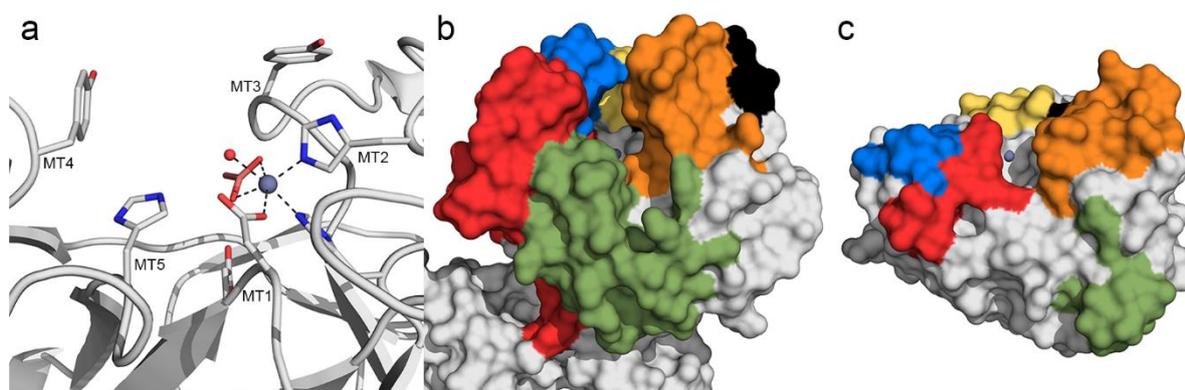


Figure 8. Structure of CE4 deacetylases. (a) *CiCDA* in cartoon representation zooming in on the active site. The side chains of the most important residues from the five catalytically important sequence motifs is shown as sticks. The metal ion is coordinated in an octahedral fashion (black dashed lines) by the metal binding triad, a water molecule (red sphere), and an acetate ion (sticks with pink carbons). (b) Surface representation of *VcCDA* (PDB id 4NY2) highlighting the loops that are proposed to be involved in the substrate-capping model. (c) Surface representation of *CiCDA* (PDB id 2IW0), showing the open active site. The loops are colored according to the color scheme used in Fig. 7: loop 1 in yellow, loop 2 in blue, loop 3 in red, loop 4 in orange, loop5 in green, and loop 6 in black. The metal ion is shown as a grey sphere. The figure was made using PyMol (Schrödinger, 2015).

From an applied point of view, utilization of deacetylases for production of chitosans and CHOS with defined P_A is of great interest since the P_A together with the F_A influences the biochemical properties of chitosan and CHOS, as previously described (Section 1.2). In this context, VcCDA and NodB from *Rhizobium* sp. GRH2 have been utilized to produce CHOS that are deacetylated specifically at the non-reducing end and the neighboring sugar unit (Hamer et al., 2015). Recently, a fungal deacetylase from *Puccinia graminis* was shown to deacetylate all GlcNAc units in different CHOS, except the two at the non-reducing end (Naqvi et al., 2016).

1.3.1.5 Chitinases

Chitinases are hydrolytic enzymes that hydrolyze the O-glycosidic bonds in chitosan and can cleave all types of glycosidic linkages found in chitosan, except, usually, GlcNAc-GlcNAc linkages. As mentioned in Section 1.3.1.1, GH18 chitinases can also hydrolyze chitosan, however, they require a GlcNAc unit in the -1 subsite, an important property that separates them from chitinases. Chitinases exist in GH families 5, 7, 8, 46, 75, and 80. GH families 5, 7, and 8 contain enzymes that hydrolyze a variety of polysaccharides, while GH families 46, 75, and 80 exclusively contain chitinases. Chitinases in GH families 8, 46, 75, and 80 hydrolyze chitosan through an inverting mechanism, contrary to chitinases in GH families 5 and 7, which use a retaining mechanism [www.cazy.org, (Hoell et al., 2010, Lombard et al., 2014)]. Based on substrate specificity, chitinases are sometimes divided into four subclasses; subclass I enzymes hydrolyze GlcNAc-GlcN and GlcN-GlcN linkages (Fukamizo et al., 1994), subclass II enzymes hydrolyze GlcN-GlcN linkages (Izume et al., 1992), subclass III enzymes cleave GlcN-GlcNAc and GlcN-GlcN linkages (Mitsutomi et al., 1996), while subclass IV can cleave all types of linkages (Hirano et al., 2012). Notably, some chitinases also cleave GlcNAc-GlcNAc linkages, albeit not very efficiently (Heggset et al., 2010).

1.3.1.6 Carbohydrate-binding modules

In recalcitrant polysaccharides such as chitin and cellulose, the substrate is often difficult to access for carbohydrate-active enzymes. Many CAZymes have solved this problem by including one or several non-catalytic CBMs that promote association of the enzyme to the polysaccharide and which may also contribute to correct (“productive”) positioning of the catalytic module (Boraston et al., 2004). A recent study showed that the beneficial effect of CBMs on enzyme efficiency is reduced at higher substrate concentrations, underpinning their

role in substrate binding (Várnai et al., 2013). Currently (September 2017), 81 different CBM families exist, and chitin-binding CBMs are found in families 1, 2, 3, 5, 12, 14, 18, 19, 37, 50, 54, 55, and 73 [www.cazy.org] (Lombard et al., 2014). CBM families 5 and 12 are distantly related and often referred to as CBM5/12. Of the different CBM families, the CBM50 family contains most entries. CBM50 proteins are also known as LysM domains, which bind to various GlcNAc containing carbohydrates such as peptidoglycan, chitin, and CHOS (Buist et al., 2008, Akcapinar et al., 2015).

In addition to the CAZy classification, CBMs have been divided into three types, based on structural and functional similarities: type A, surface-binding CBMs; type B, glycan-chain-binding CBMs; type C, small-sugar-binding CBMs (Boraston et al., 2004). The binding of CBMs to crystalline surfaces involves aromatic residues on the binding surface of the module, and several papers have demonstrated the importance of these aromates for the activity of the appended catalytic domains towards crystalline substrates [e.g. (Akagi et al., 2006, Viegas et al., 2008)]. CBMs are not usually considered to be important towards oligomeric substrates and, indeed, studies have shown that the aromatic residues in the binding surface of a CBM are not important for the activity of the appended catalytic domain towards such short (and soluble) substrates (Uchiyama et al., 2001, Katouno et al., 2004, Akagi et al., 2006).

1.3.2 Biological roles of chitin-active enzymes

Chitin-active enzymes are important for all organisms that contain or metabolize chitin. From a biotechnological and biorefining point of view, the conversion of chitin to soluble products that can be metabolized by different organisms is perhaps the most central. This subject of microbial chitin degradation is described in detail in Section 1.4. In addition to the metabolic function of chitin-active enzymes, a multitude of other functions have been described for this diverse group of enzymes. Chitin-active enzymes are widespread in Nature, found in all domains of life, and the biological role of these proteins varies.

Humans, although devoid of chitin, have two GH18 chitinases encoded in the genome: human chitotriosidase (Hollak et al., 1994) and acidic mammalian chitinase (Boot et al., 2001). The human chitotriosidase plays a role in the innate immune system against chitin-containing pathogens (van Eijk et al., 2005), while the acidic mammalian chitinase has gained attention due to its possible link to the pathophysiology of asthma (Zhu et al., 2004). In fungi, chitinases are postulated to have a wide variety of functions including degradation of exogenous chitin

for nutrition, remodeling of the (own) fungal cell wall (which contains chitin), and contributing to the defense against other fungi and arthropods (Seidl, 2008). Fungal chitinases have also been proposed to act as virulence factors in pathogen fungi infecting insects (Huang et al., 2016). The number of chitinases encoded by fungi varies from one to over 30, making it an extensive task to determine the exact role of each chitinase in a fungus encoding several chitinases (Gruber and Seidl-Seiboth, 2012, Langner and Göhre, 2016). This is discussed further in Section 1.4.4. In plants, chitinases are important in the defense against fungal attacks (Broglie and Chet, 1991, Collinge et al., 1993). Bacteria generally use chitinases to degrade chitin for utilizing it as a nutrient source (Watanabe et al., 1997, Orikoshi et al., 2005). However, there are indications that bacterial chitinases have additional roles, based on putative activities on non-chitin substrates such as glycoproteins (Adrangi and Faramarzi, 2013, Frederiksen et al., 2013). For example, virulence of *Listeria monocytogenes* in mammals is dependent on a chitinase, suggested to suppress the expression and activity of a nitric oxide synthase, an important part of the innate immune system in mammals (Chaudhuri et al., 2013).

Chitobiases have several biological roles, dependent on the organism and even the cell type. In addition to participating in chitin catabolism, several additional biological functions of bacterial chitobiases is proposed. For example, a chitobiase in *Escherichia coli* is shown to be important in cell wall recycling by hydrolyzing the β -1,4-linkage between GlcNAc and anhydro-*N*-acetylmuramic acid (Cheng et al., 2000). In the biofilm-forming bacterium *Actinobacillus actinomycetemcomitans*, a chitobiase is important for detachment of cells from the biofilm in order to enable spreading of the biofilm to other surfaces (Slámová et al., 2010). In fungi, chitobiases have several biological roles, especially in controlling the composition of chitin in the cell wall. Chitobiase activity is also important for nutrient release during saprophytic and mycoparasitic growth phases in fungi and is proposed to be involved in insect pathogenesis (Slámová et al., 2010).

Chitosanases are found in bacteria, fungi, and plants, having different biological functions (Thadathil and Velappan, 2014). Some organisms secrete chitosanases to degrade chitosan and utilize this as a nutrition source (Viens et al., 2015). Saito et al. (2009) demonstrated that a GH46 chitosanase from *Amycolatopsis* sp.CsO-2 had antifungal activity against *Rhizopus oryzae*. Chitosanases may also have a role in protection against the antimicrobial activity of chitosan and CHOS (Ghinat et al., 2010).

In addition to participating in chitin catabolism, LPMOs are suggested to act as virulence factors in some organisms. *Listeria monocytogenes* possess two chitinases and an LPMO possibly involved in the pathogenesis of this bacterium (Chaudhuri et al., 2010, Paspaliari et al., 2015). In *V. cholerae* an LPMO (Loose et al., 2014) termed GbpA binds mucin, and thereby enhances bacterial colonization of the intestine (Bhowmick et al., 2008). Mucin consist of glycoproteins, which are glycosylated with different carbohydrates including GlcNAc (Barchi, 2013); it is not yet known whether the LPMO domain of GbpA acts on the mucins. Another example of an LPMO potentially involved in bacterial virulence is *SmCBP21* from *S. marcescens*. Kawada et al. (2008) showed that knocking out this protein significantly decreased adhesion of the bacterium to colonic epithelial cells, and suggested that *SmCBP21* and similar proteins are involved in bacterial adhesion to such cells.

Deacetylases in the CE4 family are believed to have a role in pathogenesis of Gram-positive bacteria, since they, by modifying the peptidoglycan layer, make the bacteria less susceptible to the host innate immune system (Boneca, 2005, Zhao et al., 2010). The CE4 peptidoglycan deacetylase of *S. pneumonia* (*SpPgdA*) acts as a virulence factor by deacetylating GlcNAc residues in peptidoglycan, thereby obstructing lysozyme activity of the host. Knocking out the *PgdA* gene made the bacterium lysozyme sensitive (Vollmer and Tomasz, 2000, Vollmer and Tomasz, 2002). In fungi, CE4 deacetylases are thought to have a similar role in pathogenesis by deacetylating the chitin in the fungi cell wall to chitosan, creating a poorer substrate for host chitinases (El Gueddari et al., 2002, Cord-Landwehr et al., 2016). It has also been shown that deacetylases are required for yeast spore wall formation (Christodoulidou et al., 1996, Christodoulidou et al., 1999), which indicates that remodeling of chitin plays a role in this process. The heterodimer GlcNAc-GlcN, produced by highly specific chitin deacetylases in Gram-negative *Vibrio* species, induces the production of chitinases by functioning as a signal molecule in the catabolic chitin cascade (Hirano et al., 2009). Deacetylases in *Vibrios* and *Photobacteria* produce CHOS with a deacetylated sugar next to the penultimate GlcNAc, and the resulting products, GlcNAc-GlcN-(GlcNAc)_n (Li et al., 2007), resemble those produced by NodB, another specific CE4 deacetylase producing GlcN-(GlcNAc)₃₋₄ (Zhao et al., 2010). The products from NodB are intermediates in the biosynthesis of Nod factors, which are important in the communication between symbiotic nitrogen fixing bacteria and plants (Zhao et al., 2010). This suggests that the products produced by deacetylases in *Vibrios* and *Photobacteria* could be important in cellular

signaling (Li et al., 2007, Zhao et al., 2010). Due to the abovementioned biological functions of CE4 deacetylases, they are interesting targets for biological control of pathogenesis.

In addition to the important role of CBMs in promoting substrate binding when associated with catalytic domains, many other biological roles of CBMs have been suggested, including roles in virulence related to the ability of CBMs to bind to carbohydrate structures in the host (Guillén et al., 2010). Although CBMs often are associated with catalytic domains hydrolyzing different polysaccharides, they can also exist as single domains or with catalytic domains targeting non-polysaccharide substrates. For example, the plant pathogenic fungus *Cladosporium fulvum* secretes Av4, a CBM14 protein, which binds specifically to chitin in fungal cell walls, protecting the fungi from host chitinases during infection (van den Burg et al., 2006, van Esse et al., 2007). In the same fungus, a LysM domain (CBM50) mediates virulence through perturbation of chitin-triggered host immunity (de Jonge et al., 2010).

1.4 Microbial degradation and utilization of chitin

Chitinolytic microbes are widespread in Nature, and are able to degrade chitin under both aerobic and anaerobic conditions. In general, chitin catabolism is achieved by a microorganism through secretion of a set chitinolytic enzymes that attack the insoluble chitin, generating shorter soluble CHOS that are imported into the cell and processed further by other enzymes for utilization as a nutrient source. The strategy used by the microbe seems to depend on the environmental conditions in which the microbe dwells. For example, in a comparative study of chitinolytic proteins encoded by bacterial genomes performed by Bai et al. (2016), substantial differences were found when comparing aquatic and terrestrial bacteria. The modular composition of chitinases differed substantially between the two habitats and the terrestrial bacteria seemed more adapted to various chitin sources, having a higher number of chitinase genes, higher diversity of associated CBMs, and a higher number of LPMOs (Bai et al., 2016). Different microorganisms have developed different systems for chitin modification and utilization, and a few of such systems are discussed in detail below.

1.4.1 Chitin degradation by *Serratia marcescens*

S. marcescens is a Gram-negative bacterium and one of the most studied chitinolytic bacteria. The *SmCBP21* protein from *S. marcescens* B JL200, originally classified as CBM33, was the first of its kind to be identified as an LPMO (Vaaje-Kolstad et al., 2010, Horn et al., 2012b). The chitinolytic machinery of *S. marcescens* (Fig. 9) consist of three chitinases (*SmChiA*,

SmChiB, and *SmChiC*), *SmCBP21*, and a chitobiase. *SmChiA* and *SmChiB* are predominantly processive exo-acting enzymes, degrading the chitin from the reducing and non-reducing end, respectively (Perrakis et al., 1994, Van Aalten et al., 2000, Hult et al., 2005, Igarashi et al., 2014). Contrary to *SmChiA* and *SmChiB*, *SmChiC* is a non-processive endo-acting enzyme (Horn et al., 2006b, Horn et al., 2006c). As early as in 2005 (Vaaje-Kolstad et al., 2005), it was shown that *SmCBP21* promotes chitinase activity on chitin, and since 2010 (Vaaje-Kolstad et al., 2010), we know that *SmCBP21* is an LPMO that makes oxidative cleavages in crystalline regions of the substrate, creating new ends for the chitinases to bind to (Vaaje-Kolstad et al., 2013). The main product produced by the chitinases is $(\text{GlcNAc})_2$, which is cleaved into monomers by chitobiase (Fig. 9).

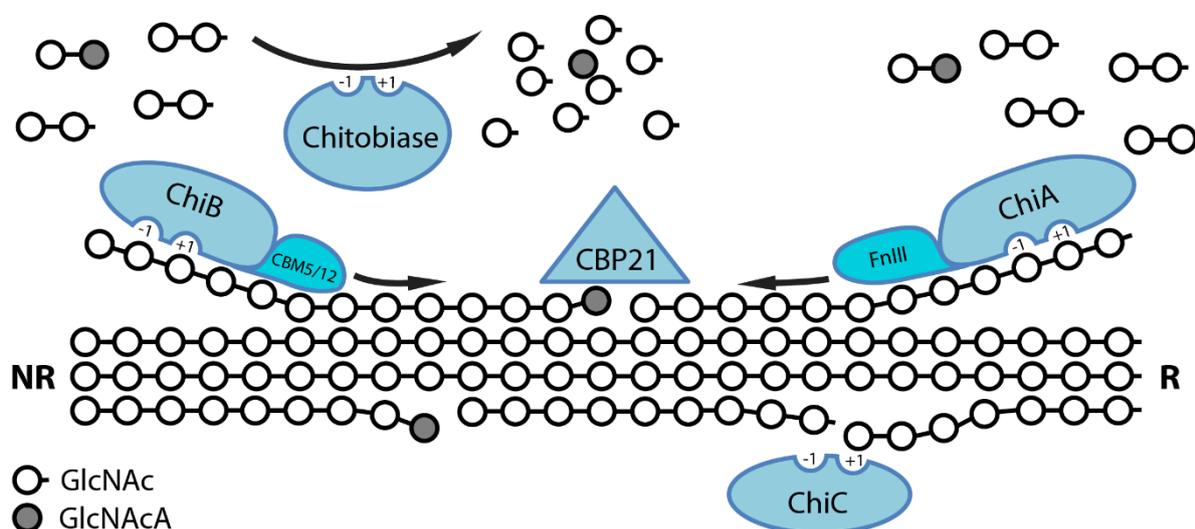


Figure 9. The chitinolytic machinery of *S. marcescens*. The chitinases *SmChiA*, *SmChiB*, and *SmChiC* hydrolyze glycosidic bonds in different fashions. *SmCBP21* is an LPMO, making oxidative cuts in the crystalline part of the chitin. The chitobiase degrades soluble products generated by the chitinases and the LPMO into monomers. Note that *SmChiC* has FnIII and CBM5/12 domains and that the chitobiase is a four-domain protein. Figure taken from Vaaje-Kolstad et al. (2013).

As mentioned in Section 1.3.1.6, catalytic domains degrading crystalline polysaccharides often have CBMs to aid in substrate binding. Accordingly, *SmChiA* has an N-terminal FnIII module contributing in substrate binding and improving catalytic activity (Perrakis et al., 1994, Uchiyama et al., 2001), while *SmChiB* has a C-terminal CBM5/12 module (Van Aalten et al., 2000). *SmChiC* has two C-terminal domains, one FnIII and one CBM5/12 module, however in culture supernatant two versions of *SmChiC* are commonly found; C1 and C2. The C1 version contain the catalytic domain and the two C-terminal domains, while C2 only

consist of the catalytic domain (Suzuki et al., 1999, Watanabe et al., 1997). The chitobiase consist of four domains designated I, II, III, and IV, with domain III being the catalytic domain. Domain I share structural similarities with a CBM2, while the functional roles of domain II and IV are unknown (Tews et al., 1996, Vaaje-Kolstad et al., 2013).

When grown on chitin, *S. marcescens* secretes all the three chitinases and *SmCBP21* (Brurberg et al., 1994, Brurberg et al., 1995, Suzuki et al., 1998, Hamilton et al., 2014). Interestingly, *SmChiB* and *SmChiC* lack a conventional signal peptide. Hamilton et al. (2014) showed that secretion of the chitinolytic machinery in *S. marcescens* Db10/11 is dependent on a holin-like protein and an endopeptidase, however the mechanism for this and the exact role of the holin/endopeptidase system is not clear (see Section 1.5 for details on protein secretion). Interestingly, the genes encoding *SmCBP21* and *SmChiB* are localized in a gene locus dedicated to chitin metabolism. This gene locus also contains the holin and endopeptidase genes, a LysR transcriptional regulator involved in regulation of chitinase expression, and two spanin genes (Suzuki et al., 2001, Hamilton et al., 2014). The chitobiase, whose expression is induced by the presence of $(\text{GlcNAc})_2$, is located in the periplasm, although small amounts of the enzyme have been detected in culture supernatants (Toratani et al., 2008). Fig. 10 shows a proposed model for the catabolic pathway for chitin utilization in *S. marcescens*. Generally, the catabolic pathway in *S. marcescens* is similar to the chitinolytic machinery described in other Gram-negative (Hunt et al., 2008) and Gram-positive bacteria (Nazari et al., 2013).

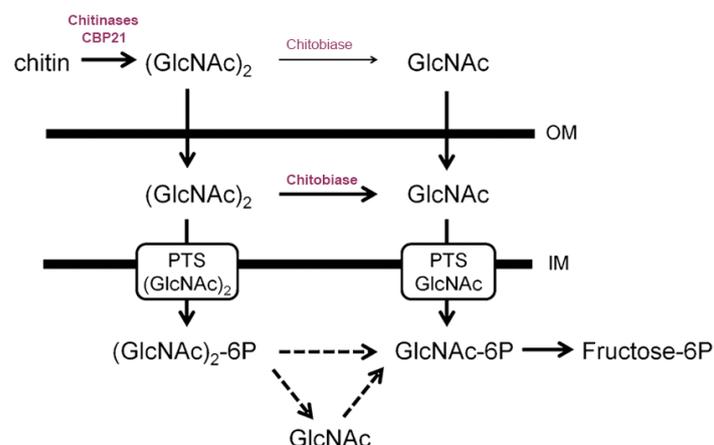


Figure 10. Catabolic pathway for chitin utilization in *S. marcescens*. Pathways supported by experimental data are shown with solid arrows, and enzymes of the chitinolytic machinery are indicated. Pathways with dashed arrows are not supported by experimental data. OM; outer membrane, IM; inner membrane. Figure adapted from Toratani et al. (2008).

1.4.2 Chitin degradation by bacteria in the Bacteroidetes phylum

The Bacteroidetes phylum contains Gram-negative bacteria that are generally recognized for effectively degrading carbohydrates. Although, most studied Bacteroidetes are anaerobes related to the gut microbiota of mammals, Bacteroidetes can be found in both anaerobic and aerobic environments (Thomas et al., 2011). The genes encoding the proteins needed for degrading a specific carbohydrate are often organized in gene clusters called polysaccharide utilization loci [PULs (Martens et al., 2009)]. The first described PUL is the starch utilization system of the anaerobic bacterium *Bacteroidetes thetaiotaimicron*, which, next to enzymes contains a so-called SusC/D pair, i.e. an outer membrane porin and a carbohydrate binding protein, respectively (D'Elia and Salyers, 1996, Reeves et al., 1997). SusC/D-like pairs are now used as identifiers for PULs in different organisms (Terrapon et al., 2015).

Recently, Larsbrink et al. (2016) described the first chitin utilization locus, which was found in the aerobic Bacteroidetes *Flavobacterium johnsoniae*, consisting of eleven genes encoding four enzymes, a predicted two-component sensor/regulator system, and two SusC/D-like pairs (Fig. 11). The four enzymes are a secreted multidomain GH18 chitinase (ChiA), an outer-membrane anchored GH18 chitinase (ChiB), a periplasmic chitobiase (GH20), and a

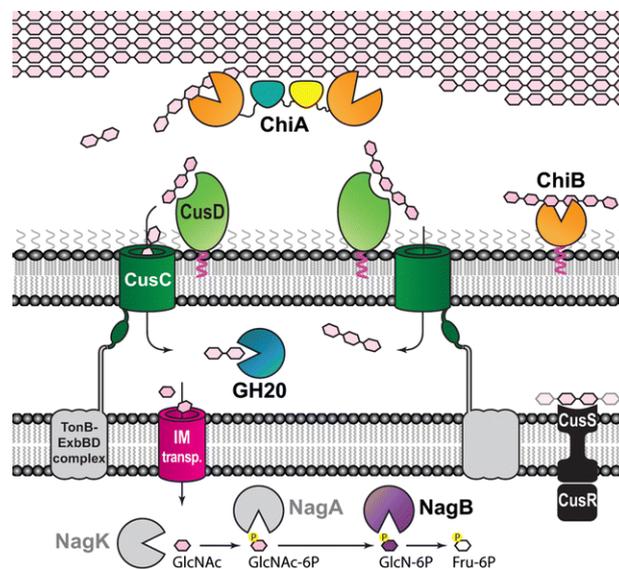


Figure 11. Proposed pathway for chitin utilization in *F. johnsoniae*. The secreted multi-domain ChiA converts chitin to CHOS. ChiB can also degrade CHOS and chitin, although not as efficiently as ChiA. The SusC/D-like pairs (here called CusC/D) capture and import CHOS to the periplasm. The chitobiase (GH20) converts CHOS to monomers, before translocation of the monomers to the cytoplasm where they are converted to fructose-6-phosphate by NagK, A and B. The proteins not encoded by the PUL are colored grey. Figure taken from Larsbrink et al. (2016).

glucosamine-6-phosphate deaminase (NagB). Interestingly, to date, no LPMOs have been discovered in Bacteroidetes (Hemsworth et al., 2016, Larsbrink et al., 2016). The absence of LPMOs may be related to the fact that most Bacteroidetes live under anaerobic conditions. It is worth noting, however, that, as mentioned in Section 1.3.1.3, recent findings suggest that molecular oxygen is not crucial for LPMO activity (Bissaro et al., 2017). Another explanation could be that the PULs in themselves are so powerful when it comes to polysaccharide degradation, that the need for LPMOs are relieved (Larsbrink et al., 2016).

1.4.3 Chitin degradation by *Thermococcus kodakaraensis*

Tanaka and co-workers have studied the rather peculiar chitinolytic pathway of the hyperthermophilic anaerobic archaeon *Thermococcus kodakaraensis* KOD1 (Tanaka et al., 1999, Tanaka et al., 2001, Tanaka et al., 2003, Tanaka et al., 2004). The chitin utilization system of this archaeon is organized in a gene cluster encoding a multimodular family GH18 chitinase (*Tk-ChiA*), a family CE14 deacetylase (*Tk-Dac*), a family GH35 glucosaminidase (*Tk-GlmA*), an ABC transporter system, and several hypothetical proteins. The chitinase has two catalytic GH18 domains and three CBMs. The catalytic GH18 domains show different activities, the first domain being an exo-type chitinase, while the second is an endo-type chitinase (Tanaka et al., 2001). Two of the CBMs belong to family 2 and can bind both chitin and cellulose, while the last CBM belongs to family 5 and binds to chitin (Tanaka et al., 1999, Lombard et al., 2014, Hanazono et al., 2016). *Tk-Dac* is a highly specific deacetylase that only deacetylates the non-reducing end of fully acetylated CHOS (Tanaka et al., 2004). *Tk-GlmA* is active against a range of fully deacetylated CHOS, although the dimer GlcN-GlcNAc is considered as the physiological substrate (Tanaka et al., 2003, Tanaka et al., 2004). In the proposed pathway for chitin conversion by *T. kodakaraensis* KOD1 (Fig. 12), *Tk-ChiA* produces (GlcNAc)₂ that is imported into the cells by the ABC transporter. The intracellular deacetylase, *Tk-Dac*, specifically deacetylates the non-reducing end of (GlcNAc)₂ after which the glucosaminidase, *Tk-GlmA*, hydrolyses the GlcN-GlcNAc dimer. Finally, *Tk-Dac* deacetylates the GlcNAc monomer to generate GlcN (Tanaka et al., 2004).

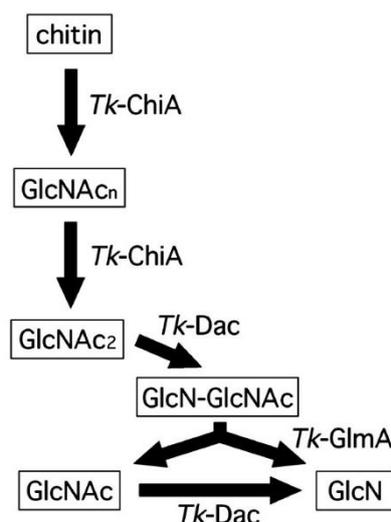


Figure 12. The chitinolytic pathway of *T. kodakaraensis*. *Tk-ChiA* converts chitin to $(\text{GlcNAc})_2$, which is deacetylated on the non-reducing end by *Tk-Dac*. *Tk-GlmA* then hydrolyses the GlcN-GlcNAc dimer, and *Tk-Dac* finally deacetylates GlcNAc, to produce GlcN. Figure taken from Tanaka et al. (2004).

1.4.4 Chitin degradation by fungi

Fungal genomes often encode several chitinases; it is not exceptional to find 10-25 different chitinases, which, all belong to family GH18 (Seidl, 2008, Gruber and Seidl-Seiboth, 2012). For example, the genome of the well-studied entomopathogen *Beauveria bassiana* encodes twenty GH18 chitinases possibly involved in pathogenesis by aiding the penetration of the chitin-containing exoskeleton of the host (St Leger et al., 1986, Xiao et al., 2012). *Trichoderma* species, which are known for biomass degradation as well as chitinase-mediated biocontrol functionalities, contain a various number of GH18 encoding genes, as many as 36 GH18s are encoded by *Trichoderma virens* (Seidl et al., 2005, Gruber and Seidl-Seiboth, 2012). Chitinases in fungi are sometimes divided into three subgroups, A, B, and C, based on their modular structures. Group A contains chitinases without extra domains, group B contains chitinases linked to CBM1s, while group C comprise chitinases with CBM18s and LysM (CBM 50) domains (Seidl et al., 2005, Seidl, 2008). Chitin is an important part of the fungal cell wall, which also contains β -1,3-glycan and a manno-protein layer [Fig. 13 (Bowman and Free, 2006)].

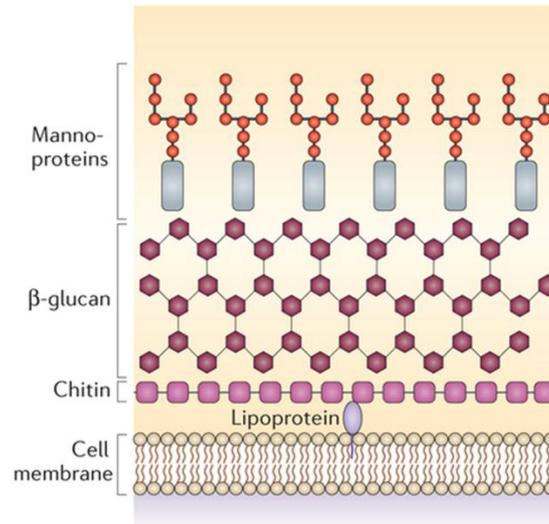


Figure 13. The fungal cell wall. Schematic representation of the fungal cell wall containing a cell membrane, a chitin layer, and β -glucan and manno-protein layers. Figure taken from Brown et al. (2015).

As mentioned in Section 1.3.2, fungal chitinases are involved in cell wall remodeling and in degradation of exogenous chitin, and an often raised issue is how the fungi discriminate between self and non-self chitin. It has been proposed that different chitinases are specialized to do different jobs. Gruber and Seidl-Seiboth (2012) hypothesize that differentiation between self and non-self is regulated by substrate accessibility (Seidl, 2008, van den Burg et al., 2006). The high number of chitinases often found in the genomes make genetic assessment of chitinase functionality a challenging task. For instance, knocking out one chitinase gene might not lead to any phenotypical change, and it is needed to make double or multiple knock-outs to get a phenotypic effect, indicating possible redundancy in fungi encoding many chitinases (Alcazar-Fuoli et al., 2011, Langner and Göhre, 2016).

1.5 Protein secretion in Gram-negative bacteria

While protein secretion in Gram-positive bacteria seems relatively simple and uniform, exploiting a limited set of secretion mechanisms (Schneewind and Missiakas, 2014), Gram-negative bacteria, such as the bacteria discussed in Section 3.2-3.4 of this thesis, have developed an array of secretion mechanisms to translocate different compounds (DNA, proteins, small molecules etc.) over their two membranes (Desvaux et al., 2009, Costa et al., 2015). The different mechanisms in Gram-negative bacteria, referred to as the type 1-9 secretion systems (T1-9SS; Fig. 14), can be divided into two groups based on the secretion mechanism, which may be a one-step or a two-step mechanism. T1SS, T3SS, T4SS, and T6SS

span the two cell membranes and secrete substrates directly from the cytoplasm to the target location outside the cell in one-step. The target location can be the extracellular space itself or a target cell. In the latter case, well known from pathogenic bacteria, the secretion system docks on the target cell and the secreted compound is injected into this cell (Fig. 14). The secretion systems with a two-step mechanism (T2SS, T5SS, T7SS, T8SS, and T9SS) depend on initial translocation of the substrate over the inner membrane by systems also found in Gram-positive bacteria, and subsequently the secretion system secretes the substrate from the periplasm over the outer membrane in a separate step. The target location of proteins secreted by these latter mechanisms can be the extracellular space or the compounds may end up anchored to the outer membrane (Costa et al., 2015).

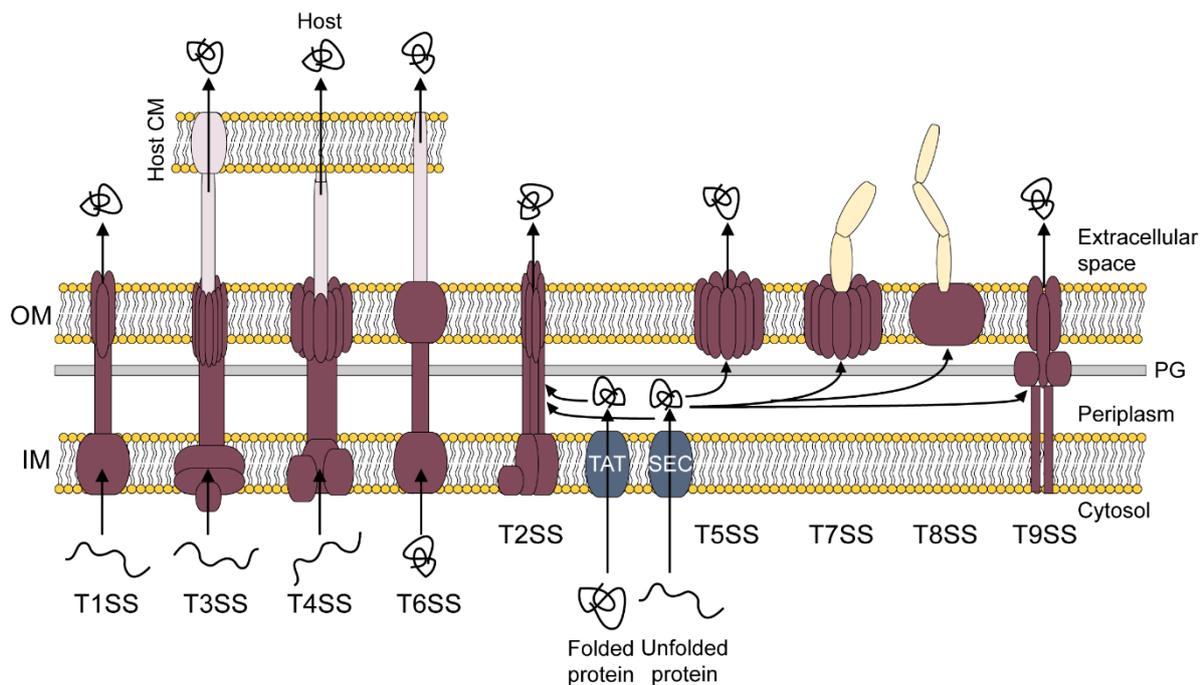


Figure 14. Secretion systems in Gram-negative bacteria. The figure shows a simplified schematic representations of the organization of the nine secretion systems and the Sec and Tat translocation machineries mentioned in the text. OM: outer membrane, IM: Inner membrane, CM: Cell membrane, PG: Peptidoglycan. See text for more details.

Proteins secreted in one-step with T1SS, T3SS, T4SS, or T6SS are guided to these systems by different signal motifs in the amino acid sequence. Proteins secreted through T1SS usually have a glycine rich motif at the C-terminal end (Costa et al., 2015), while T3SS secreted proteins have a secretion signal on the N-terminus of the amino acid sequence which is not cleaved upon secretion (Sory et al., 1995, Schesser et al., 1996). Proteins secreted via the T4SS share a conserved hydrophilic motif with a net positive charge on the C-terminal end

(Vergunst et al., 2005). Proteins secreted through the T6SS are proposed to contain an N-terminal motif named MIX (Salomon et al., 2014), however this does not seem to be universal as proteins without this motif have been shown to be secreted via T6SS (Liang et al., 2015).

Translocation over the inner membrane in the two-step mechanisms is commonly dependent on the SecYEG translocon (Sec) or the twin-arginine translocation pathway (Tat; Fig. 14). To be translocated via the Sec or Tat system, proteins have an N-terminal signal peptide that targets the protein to the correct processing system. The signal peptide is normally cleaved off during, or shortly after translocation by a signal peptidase. Lipoproteins anchored to the inner or outer membrane can also be translocated to the periplasm via the Sec system (see Section 1.5.1 for more details on lipoproteins). The major difference between the Sec- and Tat-pathway is that unfolded proteins are translocated via the Sec-pathway, while protein folded in the cytoplasm are translocated via the Tat-pathway (Natale et al., 2008). It is worth noting that folded proteins without a signal peptide may be translocated from the cytoplasm to the periplasm via holins. Holins are small membrane proteins originating from phages and are involved in translocation and activation of cell-wall hydrolyzing proteins (Desvaux and Hébraud, 2006). As mentioned in Section 1.4.1, secretion of the chitinolytic machinery in *S. marcescens* Db10/11 is dependent on a holin-endopeptidase system, similar to the lysis cassette found in phages (Hamilton et al., 2014). However, a precise description of the role of this system in the secretion of the chitinolytic machinery is not available. Section 3.4 discuss the secretion of the chitinolytic machinery in *S. marcescens* in more detail.

Compounds secreted through the T1SS are often associated with virulence, nutrient acquisition, and adhesion (Natale et al., 2008, Costa et al., 2015). T3SS, T4SS, and T6SS translocate different effector molecules from pathogenic bacteria into a target host cell. T4SS is somewhat special as translocation of DNA, in addition to proteins, is possible with this system (Costa et al., 2015). Proteins secreted via the T2SS are often hydrolytic enzymes and toxins (Costa et al., 2015). For example, the T2SS is needed for secretion of CAZymes involved in cellulose degradation by the Gram-negative bacteria *Cellvibrio japonicus* (Gardner and Keating, 2010) and secretion of cholera toxins produced by *V. cholerae* during infection (Hirst et al., 1984, Reichow et al., 2010). The T5SS, also referred to as the autotransporter system, meaning that the proteins contain a translocator in their C-terminus that directs secretion across the outer membrane or that a separate translocator protein facilitates the secretion. The T5SS is divided into five subfamilies, denoted a-e. Molecules

secreted through the T5SS are involved in adhesion, biofilm formation, or act as virulence factors (Chagnot et al., 2013, Costa et al., 2015). The T7SS represents the chaperone-usher pathway responsible for assembly and secretion of appendages called pili or fimbriae, which contribute to pathogenicity and biofilm formation, and play a role in initiation of host cell recognition and attachment (Chagnot et al., 2013, Costa et al., 2015). The T8SS refers to the extracellular nucleation-precipitation pathway involved in secreting and assembling curli, which are amyloid fibers associated with adhesion, biofilm formation and surface colonization (Desvaux et al., 2009, Costa et al., 2015). The most recently described secretion system in Gram-negative bacteria is the T9SS found exclusively in the phylum of Bacteroidetes. Proteins secreted through the T9SS contain, in addition to a classical N-terminal signal peptide, a C-terminal domain guiding them to the secretion system (Nguyen et al., 2007, Kharade and McBride, 2014). The T9SS is mainly thought to secrete proteins involved in motility and pathogenesis (Sato et al., 2010). However, ChiA of *F. johnsoniae*, described in Section 1.4.2, is shown to be secreted through the T9SS (Kharade and McBride, 2014).

It should be noted that not all secretion systems occur in all Gram-negative bacteria. Abby et al. (2016) predicted the presence of secretion systems by studying multiple Gram-negative genomes from various phyla, which revealed that T1SS and T5SS are the most widespread secretion systems.

1.5.1 *In silico* prediction of secreted proteins

N-terminal signal peptides targeting proteins to the Sec and Tat pathway share a similar tripartite architecture, with an n-region dominated by positively charged amino acids, a h-region dominated by hydrophobic amino acids, a c-region consisting of more polar amino acids, and with the cleavage site for a signal peptidase (von Heijne, 1990, Bendtsen et al., 2005b). Several algorithms have been developed for *in silico* prediction of different N-terminal signal peptides and, consequently the subcellular locations of their cognate proteins, in Gram-negative bacteria (Juncker et al., 2003, Bendtsen et al., 2005a, Bendtsen et al., 2005b, Bagos et al., 2010, Petersen et al., 2011). These prediction tools have since the 1980s developed from signal peptide prediction based on weight matrices to more sophisticated machine learning approaches (Caccia et al., 2013). Today, a combination of several bioinformatics tools are often used to obtain the most reliable predictions possible (Desvaux et al., 2009).

SignalP is one of the most popular algorithms used for prediction of signal peptides having a signal peptidase I (SpI) cleavage site (Bendtsen et al., 2004, Petersen et al., 2011). However, SignalP is not able to predict signal peptidase II (SpII) cleavage sites, which are the cleavage sites found in the signal peptides of lipoproteins, i.e. proteins that concomitant with transport over the membrane become post-translationally modified with N-terminal lipid(s) and covalently bound to the cell membrane. Both SpI and SpII signal peptides target the unfolded protein for translocation from the cytoplasm to the periplasm through the Sec-pathway. In the periplasm, peptidases cleave off the signal peptide, before the protein folds and is then potentially exported out of the cell via one of the secretion systems (Fig. 14) or anchored to one of the membranes (Juncker et al., 2003). Juncker et al. (2003) developed LipoP to predict lipoproteins that have a SpII signal peptide; LipoP is convenient, since it also predicts SpI, cytosolic, and trans-membrane proteins. SpII signal peptides have a region of four well-conserved amino acids (called the lipobox) located in the c-region around the cleavage site and the first amino acid of the mature protein is always a cysteine (Juncker et al., 2003). Studies exist showing that the amino acid following this cysteine determines the final localization of a lipoprotein, i.e. anchoring to the inner or outer membrane (Seydel et al., 1999).

Signal peptides targeting folded proteins to the twin-arginine-pathway contains, as the name implies, two arginines located in the n-region (Bagos et al., 2010). Tat signal peptides are in general longer than Sec signal peptides, with a lower degree of hydrophobicity in the h-region (Juncker et al., 2003). Several algorithms are available for prediction of Tat signal peptides (Rose et al., 2002, Bendtsen et al., 2005b, Bagos et al., 2010).

Although most secreted proteins have a signal peptide guiding them to the Sec or Tat pathway (or a signal sequence guiding the protein to any of the one-step secretion systems described in Section 1.5), there are secreted proteins that do not contain a classical signal peptide. Such proteins are subject to so-called non-classical secretion. As described in Section 1.4.1, and further discussed in Section 3.4 of this thesis, *SmChiB* and *SmChiC* are examples of such proteins. SecretomeP is an algorithm that predicts proteins secreted in a non-classical fashion (Bendtsen et al., 2005a). This algorithm combines several protein features, e.g. predicted protein disorder and amino acid composition, to predict whether a protein is secreted in a non-classical fashion or not. A SecretomeP score between 0 and 1 is reported, and a score above 0.5 indicates non-classical secretion in Gram-negative bacteria. Some secreted proteins that

lack a signal peptide may have a role both in the cytoplasm and extracellularly and are referred to as moonlighting proteins (Jeffery, 1999). Gram-negative bacteria can exploit vesicles to bring proteins out to the extracellular space [e.g. (Arntzen et al., 2017)]. These vesicles are often referred to as outer membrane vesicles, meaning that the vesicle contain proteins located in the periplasm, hence proteins must be translocated from the cytoplasm to the periplasm before they can be part of an outer membrane vesicle (Kim et al., 2015). However, cytosolic proteins, DNA and RNA have also been identified in outer membrane vesicles, indicating that the cargo of these vesicles is not restricted to periplasmic content (Lee et al., 2008). One such example is the cytotoxin ClyA in *E. coli* that does not have a predicted signal peptide, but is known to be secreted via vesicles. Interestingly, this protein receive a low SecretomeP score indicating that it is not secreted in a non-classical manner (Bendtsen et al., 2005a). This example emphasis that experimental data is also needed to complement the *in silico* prediction. When doing such experiments the experimental design is important as e.g. moonlighting proteins might only be secreted under certain conditions (Bendtsen et al., 2005a). This makes it a challenging task to reveal if a protein is secreted or not.

The presence of a signal peptide is commonly used as an indication of protein secretion (Caccia et al., 2013) and in the secretomics studies presented in this thesis (Section 3.2 and 3.4) proteins with a predicted SpI, SpII, or Tat signal peptide were included when calculating the fraction of secreted proteins. In addition, proteins predicted to be non-classically secreted were included. It should, however, be noted that the presence of a signal peptide is not a guarantee for secretion. In Gram-negative bacteria, a SpI or a Tat signal peptide is only guiding the proteins to the Sec or the Tat translocon systems, respectively, translocating the protein over the inner membrane. Whether proteins are actually exported out of the cell can be difficult to predict. A SpII signal peptide indicates a lipoprotein, however it is difficult to predict if an outer membrane anchored lipoprotein faces the periplasm or the extracellular space. Until recently, there was a general understanding that lipoproteins face the periplasm, anchored to the inner or outer membrane. However recent research suggests that several lipoproteins in fact are surface-exposed, i.e. facing the extracellular space (Wilson and Bernstein, 2016).

1.6 Proteomics as a tool for studying bacterial secretomes

The term “proteome” was coined by Marc Wilkins in 1994 and is defined as “the total set of proteins expressed in a given cell at a given time”. The study of proteomes is referred to as

proteomics (Dove, 1999). Since then, the field of proteomics has greatly evolved, from qualitative determination of a protein's presence, to proteome-wide protein quantification. The development of proteomics has gone hand in hand with the improvement of mass spectrometry (MS) methods suitable for proteins and peptides (Ong and Mann, 2005, Cox and Mann, 2008, Nahnsen et al., 2013). Another important factor for the development of the proteomics field, is the increased number of sequenced genomes which is essential in order to have databases to search against. The proteome of a bacterium is, in contrast to the genome, highly dynamic and shifts depending on environmental factors such as nutrient source, stress, interactions with a host and temperature. The workflow of a proteomics experiment (Fig. 15)

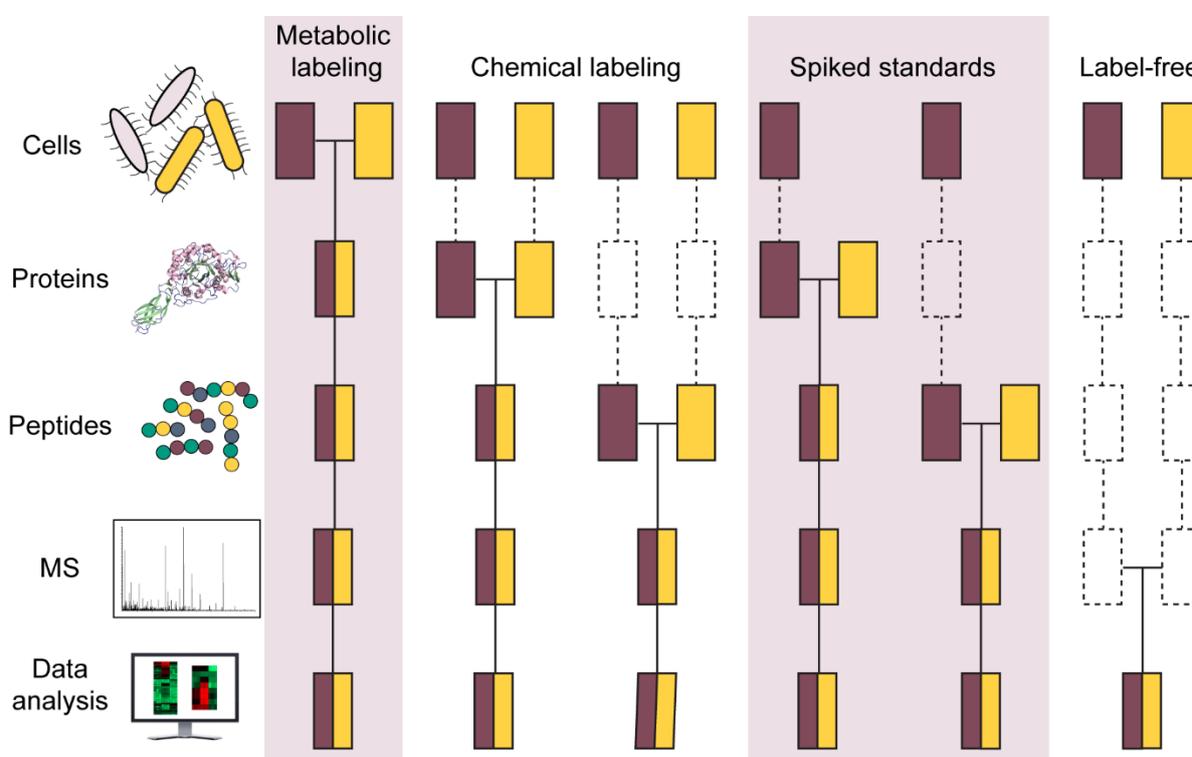


Figure 15. Common workflows in quantitative MS-based proteomics. The dark purple and yellow boxes represent two experimental conditions. Horizontal lines indicate when samples are combined, while dashed lines indicate points where experimental errors and quantification errors can occur. Figure adapted from Bantscheff et al. (2007).

typically starts by growing cells under specific conditions, before extracting, and, if desired, subsequent fractionation of the proteins. Denatured proteins are then subjected to an endoprotease with a specific cleavage pattern, generating peptides. Trypsin is commonly used since it specifically cleaves after arginine and lysine residues. Solid-phase extraction with Zip-Tips® or Stage-Tips (pipet tips packed with C-18 material) is then used to remove buffers

used in the sample preparation and to concentrate the sample making it ready for MS-analysis. Before reaching the MS, peptides are often separated by online reversed phase high-performance liquid chromatography to reduce sample complexity. The raw data from the mass spectrometer is then converted to a text-based format for protein identification using database search engines.

Protein quantification using MS is not straightforward due to a phenomenon called ionization efficiency. Peptides from a given protein have different amino acid composition and thus vary in their ability to pick up a charge during ionization (i.e. the transition from liquid to gas phase). This in turn cause different peptide intensities although they all originate from the same protein, a fact that may preclude protein quantification. However, a great effort has been done to circumvent this challenge (see below), and today, several strategies are available in quantitative proteomics (Fig. 15), each having different advantages and disadvantages.

Metabolic labelling using stable isotopes has long been considered as the gold standard in quantitative proteomics (Cox et al., 2014). Stable isotope labeling by amino acids in cell culture experiments is done by growing two cell populations separately, one with normal medium and one with medium containing isotope labeled amino acids (typically isotopically labeled arginine or lysine). A major advantage of metabolic labeling is that the two resulting cell samples (e.g. healthy versus sick cells, or cells with and without exposure to an external factor) are combined prior to further sample preparation and LC-MS, excluding potential data handling errors introduced during these steps. On the other hand, isotopes are expensive, and only a few labels are available, restricting the number of conditions to be compared to maximum five (Chen et al., 2015). Furthermore, this approach requires use of defined media, which limits the types of biological questions that can be addressed in this type of experiments.

Chemical labels can be introduced either at the protein or at the peptide level. Although universally applicable, the later introduction of the chemical labels leads to a loss in robustness, compared to metabolic labeling. (Cox et al., 2014). For absolute quantification of one or a few proteins of interest, samples can be spiked with isotopically labeled synthetic protein or peptide standards (Bantscheff et al., 2007). Like metabolic labeling, both chemical labeling and the use of spiked standards have restrictions when it comes to the number of samples that can be compared in the same experiment [referred to as multiplexing (Chen et

al., 2015)], but the amount of multiplexing is markedly higher with chemical labeling where up to 18 conditions in one run has been reported (McAlister et al., 2012).

In recent years, label-free quantification (LFQ) has gained massive popularity despite its relatively poor position in Fig. 15. This is mainly due to three important factors: 1) its simplicity, the samples do not need any kind of pre-treatment such as tedious labeling and successive clean-up steps, 2) free of charge, no need for expensive isotope reagents, and most importantly 3) a massive development in algorithms for data analysis. When using LFQ, all samples are treated separately prior to data analysis (Fig. 15) which would imply that high reproducibility in the sample preparation is required. However, thanks to the development of sophisticated software, such as the MaxLFQ algorithm implemented in the MaxQuant software package (Cox and Mann, 2008, Cox et al., 2014), these challenges have been largely overcome. In the MaxLFQ algorithm, a delayed normalization is introduced by using a nonlinear optimization model to minimize the overall peptide changes across all samples (Cox et al., 2014). Such a method is compatible with upfront pre-fractionations and makes it possible to quantitatively compare samples, even if there is some variation in the efficiency of sample preparation. It is worth noting that, when using LFQ, there is no limitation when it comes to the number of samples and conditions that can be compared in one experiment.

Proteomics is an excellent tool for studying proteins present in a cell-culture grown under specific conditions. Investigation of proteins located in the extracellular milieu is often referred to as secretomics, a term implying that only secreted proteins are studied. However, there is some debate on how to correctly define studies of proteins located in the extracellular milieu. Exoproteomics have been suggested as a more suitable term, since proteins that are thought not to be actively secreted often are found in the extracellular milieu (Desvaux et al., 2009). Proteins that are not predicted to be secreted provide a major challenge in secretomics/exoproteomics, since it is difficult to know if the occurrence of such a protein in the extracellular milieu is biologically relevant or an artefact due to e.g. cell lysis. When it comes to degradation of chitin-rich material and other biomasses, extracellular enzymes are thought to play a major role (see Sections 1.4 and 1.5). Secretomics will give insight into which enzymes an organism use to accomplish biomass conversion and may also lead to identification of “novel” proteins involved in conversion, i.e. proteins having similar expression profiles as enzymes known to be important in biomass conversion, but whose function in this process is not yet known. As an example, Takasuka et al. (2013) used

proteomics to study how the CAZyme composition in the secretome of a *Streptomyces* strain changed when growing on different biomasses. They found clear differences between cellulose- and chitin-grown cells, and could thus show how the bacteria adapt to the biomass that they are feeding on. There are several similar examples in the current literature and more examples appear in Sections 3.2 and 3.4 of this thesis.

Combining proteomics with other omics-techniques, in particular transcriptomics, allows an even deeper understanding of cellular responses to external effects. The transcriptome changes faster than the proteome and gives a snapshot of the proteins transcribed at the time of sampling. Due to variation in mRNA and protein lifetime, and due to the several (regulated) processes in between mRNA synthesis and the presence of the gene product (protein) at its proper location, proteomics and transcriptomics data may give rather different views on a system. These are complementary views, however, both provide important information.

2 OUTLINE AND PURPOSE OF THE RESEARCH PRESENTED IN THIS THESIS

Chitin, being the second most abundant biomass on Earth, is a valuable resource in a green economy. Potential downstream applications of chitin and its key derivatives, chitosan and CHOS, are numerous. However, today much of the available chitin-rich biomass, mainly crustaceans, is treated as waste and not utilized. In addition, the extraction of chitin from these sources involves hazardous chemicals, and in the shift towards a greener economy, alternative extraction processes must be considered. There is also a need for well-defined chitosans and CHOS, which have numerous applications and which can be obtained by using an array of enzymes with appropriate substrate specificities.

The research described in this thesis was part of a Norwegian national project (MARPOL) focusing on the valorization of marine polysaccharides, with an overall goal to develop innovative biomaterials and chemicals by enzyme technology. The research was focused on enzyme discovery and characterization of enzymes involved in chitin degradation and modification, and this enzyme work was conducted by using bioinformatics, proteomics, and biochemical experiments. The work included studies of the chitin utilization systems of bacteria, since such systems could be used to make the extraction process of chitin more environmentally friendly, which is of industrial interest, as part of the change towards a greener economy.

Deacetylases targeting GlcNAc units in chitin and CHOS are of special interest when it comes to production of chitosan and CHOS with defined F_A and P_A , properties which are known to be important for the functionality, and, thus, application of these compounds. Almost all known CE4 deacetylases acting on GlcNAc show broad substrate specificities and do not seem suitable for specific tailoring of chitosan or CHOS. The use of expanding sequence databases, including metagenomics data, opens a possibility to discover CE4s with different properties than already known proteins. **Paper I** describes the use of a (meta-)genomics approach to search for novel CE4 deacetylases. This approach yielded several candidates, one of which was selected for in-depth characterization. A crystal structure of an enzyme-substrate complex and biochemical data provide new insights in the substrate binding and specificity of CE4 deacetylases.

Many microorganisms degrade chitin of different origins and, if an enzymatic process is to be used for chitin extraction from e.g. shrimp shells, it is of interest to investigate the proteins secreted by these organisms when growing on chitin. The ability of *C. japonicus* to degrade plant cell wall polysaccharides has been studied in depth, however its chitinolytic potential has not been investigated. **Papers II** and **III** describe studies of the chitinolytic machinery of *C. japonicus*. **Paper II** describes the use of a novel plate method to investigate proteins found in the extracellular space during growth on chitin. The genome of *C. japonicus* encodes four putative chitinases belonging to glycoside hydrolase family 18, and proteomics data confirmed their importance in chitin utilization, in addition to revealing other proteins of possible importance in chitin utilization. The biochemical properties and biological importance of the GH18 chitinases expressed by the bacterium were examined in detail, as described in **Paper III**. The functional studies described in **Paper III**, which includes studies of knock-out mutants, showed that some of the chitinases were more important than others, thus pointing towards specific candidate enzymes for potential industrial use.

S. marcescens is one of the best studied chitinolytic bacteria, however, at the start of the research described in this thesis, the genome sequence of one of the most commonly used strains was not known. A proper investigation of the proteins secreted by this strain upon growth on chitin had neither been done. **Paper IV** describes both the genome sequence and a proteomics study of *S. marcescens* BJL200. The genome sequencing showed that *S. marcescens* encodes four GH18 chitinases, including one, *SmChiD*, whose role in chitin degradation had so far not been properly addressed. The proteomics data also revealed interesting target proteins for future investigation, due to their potential involvement in chitin degradation. The biochemical and biological properties of *SmChiD* were investigated and its interplay with other well-known *S. marcescens* chitinases in chitin degradation was assessed.

3 MAIN RESULTS AND DISCUSSION

3.1 Paper I – Structure and function of a CE4 deacetylase isolated from a marine environment

CE4 deacetylases acting on chitin and CHOS are of interest, especially due to the possible use of these enzymes to produce chitosan and CHOS with well-defined properties such as F_A and P_A . The use of (meta-)genomic data from relevant ecological niches opens up the possibility to discover proteins with different substrate specificities compared to already characterized members of this enzyme family.

A database containing annotated bacterial genomes and metagenomes of different origins was searched for members of the CE4 family using EC number 3.5.14.1. This search yielded 48 candidate CE4 proteins which were further evaluated using a set of criteria to sort out the best candidates. The protein sequences should: 1) Contain the five motifs conserved in GlcNAc active CE4s (see Section 1.3.1.4.1 for details on these motifs). 2) Have a predicted signal peptide since activity against chitin is desired, and for bacteria (and not necessarily for fungi), the chitin substrate is expected to be outside the cell. 3) Be single domains, to increase the chance of successful expression. This filtration resulted in one candidate protein consisting of 246 amino acids including a predicted signal peptide from amino acid 1 to 31. The gene originates from *Arthrobacter* sp. AW19M34-1, a Gram-positive bacterium isolated from a tunicate at 77 meters depth in Vestfjorden, Norway. The protein, named *ArCE4A*, was successfully expressed in *E.coli*.

Two crystal structures were obtained for the protein, one with and one without a $(\text{GlcNAc})_2$ ligand (**Table 1, Paper I**). The protein shows a somewhat deformed $(\beta/\alpha)_8$ topology (**Fig. 1a, Paper 1**), characteristic for CE4 proteins (Blair et al., 2005, Blair et al., 2006, Taylor et al., 2006, Andrés et al., 2014), with a relatively open active site. The structure without $(\text{GlcNAc})_2$, contains a Ni^{2+} ion, which is coordinated in an octahedral fashion by the metal binding triad and three water molecules (**Fig.1b, Paper I**). One of these water molecules is thought to act as a nucleophile during catalysis, whereas the other two likely are displaced by substrate binding (Fig. 16). CE4 deacetylases are known to require a metal ion for catalysis, and in order to get an enzyme-substrate complex, EDTA was mixed with the protein to remove metals prior to adding the $(\text{GlcNAc})_4$ substrate. Although a tetramer was used as substrate in the crystallization, only a dimer could be refined from the electron density. Apparently, the two other sugars were not stabilized by any protein-substrate interactions, which leaves them

flexible and hence no electron density was observed for them. The refined GlcNAc dimer occupied subsite 0 and +1, with the non-reducing end bound in subsite 0. The binding of the sugar in subsite +1 seems to be dominated by stacking interactions with Trp171 in motif 4, while the GlcNAc in subsite 0 makes several interactions with the protein, mainly involving fully conserved residues in GlcNAc active deacetylases. The hydroxyl-group at C3 interacts with the metal ion and makes a hydrogen bond with Asp56 in motif 1, while the hydroxyl-group at C4 seems to have an indirect interaction with the backbone carboxyl of Trp171 through a water molecule (Fig. 16).

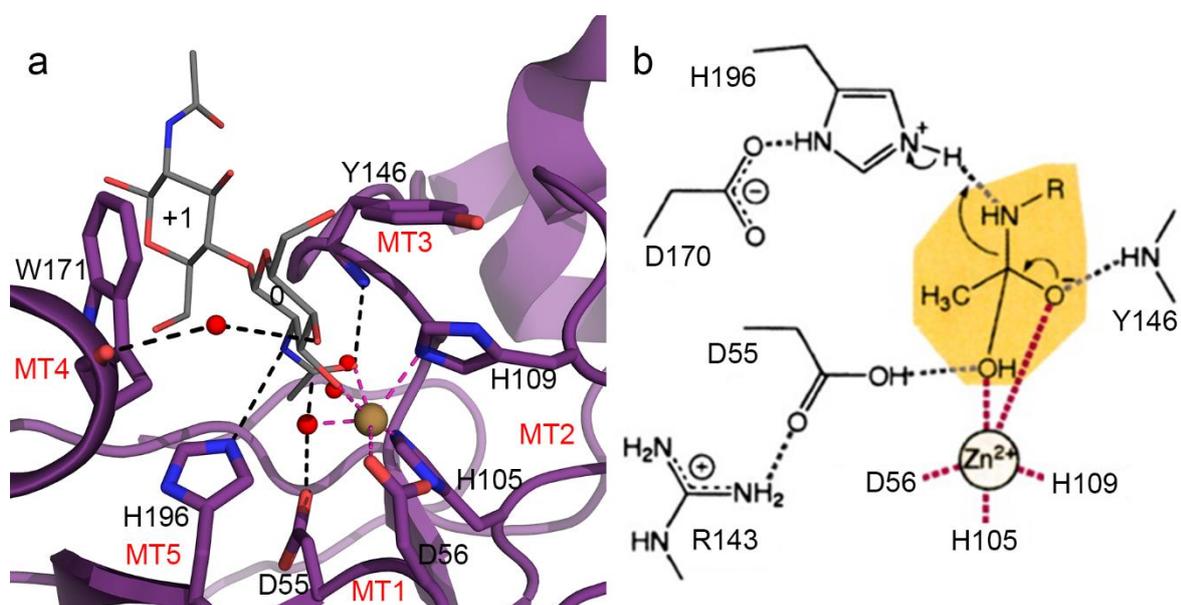


Figure 16. Enzyme-substrate interactions in the crystal structure and proposed catalytic intermediate state of *ArCE4A*. (a) *ArCE4A* in cartoon representation (purple), with residues interacting with the substrate shown as sticks (purple carbons). Interactions coordinating the metal ion (brown sphere) is shown as pink dashed lines, while interactions between the (GlcNAc)₂ substrate (sticks with grey carbons) and the enzyme is shown with black dashed lines. Water molecules are shown as red spheres. The sequence motifs which the residues belong to are shown in red text. (b) The proposed oxyanion tetrahedral intermediate state of the acid/base catalysis of CE4 deacetylases. The figure is adapted from Blair et al. (2005), showing amino acid numbering corresponding to *ArCE4A*. Since the sugar ring is absent in the figure, the interaction between the C3 hydroxyl group and the metal ion is absent. See Section 1.3.1.4.1 and Fig. 6 for complete description of the catalytic mechanism.

The backbone amine of Tyr146 is thought to stabilize the oxyanion intermediate by interacting with the oxygen in the acetyl group. In the structure of the complex, the distance

between these two atoms is 3.3 Å (**Fig. 1d, Paper I**), which is normally considered to be a bit too long for this sort of interaction, but minor positional changes affecting this distance during catalysis are conceivable. The conserved catalytic acid in motif 5, His196, interacts with the nitrogen atom in the acetamido group through its Ne nitrogen. The distance between these two atoms is 3.5 Å (**Fig. 1d, Paper I**), similar to the distance of 3.7 Å proposed in docking studies with other CE4 deacetylases (Blair et al., 2005, Blair et al., 2006). The conserved catalytic base, Asp55 in motif 1, does not make any direct interactions with the substrate. The role of Asp55 during catalysis is to activate a water molecule which subsequently performs a nucleophilic attack on the carbon in the scissile C-N bond. No water molecules were refined in the active site cleft of the structure with the (GlcNAc)₂ ligand, probably due to the lack of a metal ion. However, a superposition of the two *ArCE4A* crystal structures (Figs. 16 and **1e, Paper I**) shows that the position of one of the three water molecules visible in the structure without (GlcNAc)₂ could represent the nucleophilic water activated by Asp55. In addition, the oxygen of the acetyl group and the hydroxyl on carbon 3 in the subsite 0-bound sugar occupy the same positions as the other two water molecules observed in the metal containing substrate-free structure. This indicates that the oxygen of the acetyl group and the hydroxyl on carbon 3 of the sugar bound in subsite 0 coordinate the metal ion during catalysis.

The biochemical experiments with *ArCE4A* show that this deacetylase has a broad substrate specificity, being able to deacetylate CHOS, chitosan, acetyl xylan, and, to a smaller extent, chitin (**Table 2, Paper I**). The low activity against chitin is likely due to the crystalline nature of chitin, making most of the acetyl-groups unavailable for *ArCE4A*. Liu et al. (2017) showed that adding an LPMO to the reaction increased the activity of a deacetylase from *A. nidulans*, probably by making more acetyl-groups available for the deacetylase, either by disturbing crystal packing or by releasing soluble products. The use of a reducing end labelling method allowed mapping of how, i.e. at which positions, *ArCE4A* deacetylates soluble chitin fragments. Studies with (GlcNAc)₅ showed that deacetylation of internal sugars is preferred over deacetylation of the ends, and that the reducing end sugar is rarely deacetylated under the conditions used in this study (**Fig. 3, Paper I**). Accordingly, *CICDA*, an enzyme with a similar active site architecture as *ArCE4A*, was shown to deacetylate the reducing end sugar of (GlcNAc)₄ much slower than the other positions (Hekmat et al., 2003). Similarly, analysis of the P_A of CHOS generated by *PgtCDA* from (GlcNAc)₄₋₆ showed that deacetylation of the reducing end is less favored (Naqvi et al., 2016). Collectively, these results indicate that binding of the reducing end in subsite 0 is not favored, which is in agreement with the

crystallographically observed binding mode of the (GlcNAc)₂ ligand in *Ar*CE4A. With the methods used here, it was not possible to determine if *Ar*CE4A has preferences when it comes to deacetylation of one of the three middle sugars in chitopentaose. However, based on the signal intensities from the MS2 spectra (**Fig. 3, Paper I**), although giving only qualitative data, one might speculate that the sugar next to the reducing end is slightly favored over the other positions. Using a reacylation and ¹⁸O-labeling method before MS analysis, Naqvi et al. (2016) found that *Pgt*CDA prefers the sugar next to the reducing end during deacetylation of a pentamer.

The crystal structure of *Ar*CE4A with the (GlcNAc)₂ ligand is the first of its kind and seems to indicate that only two sugars in a tetramer made stable interactions with the protein. However, the activity of *Ar*CE4A towards (GlcNAc)₂ was low compared to longer CHOS, suggesting that other interactions between the protein and the ligand are beneficial for activity. This is also consistent with the labeling experiments showing that *Ar*CE4A preferred to deacetylate the middle sugars in a pentamer. As already mentioned in the Section 1.3.1.4.1, a docking study led Blair et al. (2006) to suggested that a (GlcNAc)₃ bound in subsite -1 to +1 in *C/C*CDA does not make any interactions with the enzyme in subsite -1. This is in consensus with the observed binding mode of the (GlcNAc)₂ ligand in *Ar*CE4A, occupying subsite 0 and +1, and the fact that only two of four sugars were refined in the structure. Unlike *Ar*CE4A, *C/C*CDA has a short loop between motif 1 and 2 containing a Trp (**Fig. 2, Paper I**), which could create a -2 subsite [**Fig. 1f, Paper I**, (Blair et al., 2006)]. This loop is one of the six loops defined in the “subsite capping model” introduced by Andrés et al. (2014). The “subsite capping model” is based on studies of *Vc*CDA, with its much more closed substrate-binding site (Fig. 8), and suggests that substrate binding could lead to conformational changes of certain loops in CE4 deacetylases. If in *Ar*CE4A conformational changes occur upon substrate binding, this could imply that the (static) crystal structures presented cannot be used to speculate about substrate-protein interactions beyond subsites 0 and +1. Notably, as discussed in Section 1.3.1.4.1, *Vc*CDA has large insertions forming flexible loops and hence a very different active site conformation compared to *Ar*CE4A, which has short loops and a much more open active site.

Of the substrates tested, *Ar*CE4 showed the highest activity towards acetyl xylan (**Table 2, Paper I**), suggesting that it might be an acetyl xylan deacetylase. Broad substrate specificity is common among CE4 enzymes, including enzymes that are generally considered as chitin

deacetylases (Caufrier et al., 2003, Liu et al., 2017). Recent data on a fungal CE4 showed a similar activity profile as *Ar*CE4A, with highest activity towards acetyl xylan. Importantly, this deacetylase, from *A. nidulans*, is structurally similar to *Ar*CE4A (Liu et al., 2017). All the CE4 enzymes having a broad specificity have open active sites with seemingly few interactions with the substrate beyond subsite 0, which could explain their broad specificity. The structures of chitin, acetylated xylan and peptidoglycan (another substrate of CE4s; see below) are different and one may argue that the main similarity between these substrates is the acetyl group (Figs. 1 and 17). On the other hand, an acetylated xylose and a GlcNAc sugar are stereochemically very similar, with all hydroxyl groups in equatorial orientation. The main difference is that xylose has a C2 *O*-acetylation, while GlcNAc has a C2 *N*-acetylation. The enzyme-substrate complex presented in **Paper I** shows that the enzyme only had direct interactions with the sugar bound in subsite 0 (the interactions in subsite +1 are stacking interactions), hence the substrate, regardless of whether it is an acetylated xylose or GlcNAc, may provide all necessary interactions for an deacetylation to occur.

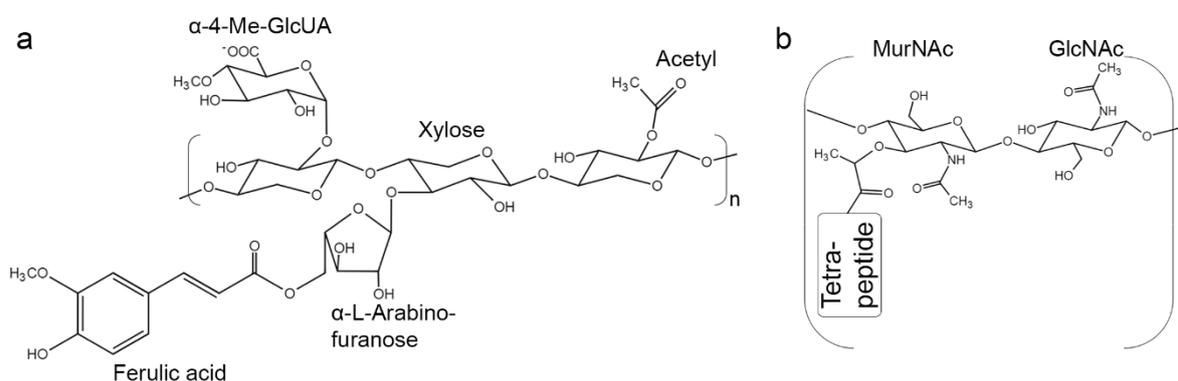


Figure 17. Structures of non-chitin polysaccharides that are deacetylated by CE4s. (a) Structure of xylan, comprising a backbone of β -1,4-linked xylose, with acetyl, 4-O-methyl- α -D-glucuronic acid (α -4-Me-GlcUA), α -L-arabinose, and ferulic acid substitutions. (b) Structure of peptidoglycan containing alternating units of β -1,4-linked GlcNAc and MurNAc (*N*-acetylmuramic acid).

Natural occurring xylan can have many substitutions, and among them are *O*-acetylations mainly on the C2 and/or C3 positions (Adesioye et al., 2016). For CE4 deacetylases to access the C2 acetyl groups in xylan, there is need to overcome the structural bulkiness caused by the other substitutions (Fig. 17a). This could explain why the active site of xylan-active CE4s is open and makes few interactions with the substrate beyond subsites 0 and +1. Structurally, the substitutions in xylan are larger than an *N*-acetyl group, and it is thus possible that an

enzyme such as *Ar*CE4A interacts with the substrate in subsites beyond 0 and +1, if the substrate is a substituted xylan. So far, there is no information on such interactions and additional crystallographic experiments with xylan fragments would thus be of interest.

Peptidoglycan consists of alternating GlcNAc and MurNAc (*N*-acetylmuramic acid). In addition to being *N*-acetylated on C2, MurNAc is substituted with a polypeptide on C3 (Fig. 17b). As the sugars are rotated 180° relative to each other, the C3 substitution is not likely to create problems for CE4 deacetylases with *Ar*CE4A-like active sites that act on the GlcNAc residue in peptidoglycan. Indeed, the active site of *Sp*PgdA, a well-known peptidoglycan deacetylase acting on GlcNAc, resembles the active site of *Ar*CE4A. Conversely, for CE4 enzymes deacetylating the MurNAc residue, the C3 substitution add some challenges when it comes to adapting the polypeptide in the active site. A peptidoglycan deacetylase from *B. subtilis* (*Bs*PdaA) is active against MurNAc, but has no activity towards CHOS (Fukushima et al., 2005). The sequence and the structure show that this enzyme has three additional basic amino acids lining the catalytic groove, not seen in GlcNAc-active CE4s, which could be important in the recognition and/or binding of the C3 substitution in MurNAc (Blair and van Aalten, 2004). In addition, the structure shows that the active site cleft is not as narrow as seen in GlcNAc-active CE4s, and, moreover *Bs*PdaA has a deep pocket next to the active site, possibly for fitting the tetrapeptide substitution on MurNAc residues. Another CE4 deacetylase from *B. subtilis* (*Bs*PdaC) can deacetylate both MurNAc residues in peptidoglycan and (GlcNAc)₁₋₄ (Kobayashi et al., 2012). The amino acid sequence of *Bs*PdaC is more similar to that *Ar*CE4A, compared to *Bs*PdaA. These observations on Pdas show that, while CE4 deacetylases of the *Ar*CE4A-type generally seem to have broad substrate specificities and primarily seem to interact with the acetylgroup of the substrate, subtle changes in substrate specificity do occur, which may provide leads for engineering of the substrate specificity of these enzymes.

The broad substrate specificity seen in *Ar*CE4A and similar CE4s could be due to these enzymes having a dual biological role, at least in fungi and Gram-positive bacteria. In these organisms CE4s can function as a defense mechanism against exposure to external chitinases and lysozyme by deacetylating the chitin and peptidoglycan layer in the cell wall. In addition, the CE4s could be important in virulence and/or for nutrient utilization by being part of a plant cell wall or chitin degradation system. In Gram-negative bacteria, the peptidoglycan layer is protected by the outer cell membrane, making deacetylation of the peptidoglycan layer

as a defense mechanism redundant. It is worth noting that genome sequencing of *S. marcescens* BJL200 (**Paper IV**) showed that this efficient chitin-degrading Gram-negative bacterium does not encode any CE4s. *S. marcescens* does neither contain a plant cell wall degrading machinery, adding to the notion that this bacterium does not need CE4s. On the other hand, Gram-negative *C. japonicus*, having an extensive plant cell wall-degrading enzyme machinery next to its chitinolytic system has four putative CE4 proteins (**Paper II**).

In conclusion, the data presented in **Paper I** give new insight into the active site cleft and substrate specificity of CE4 deacetylases with an open active site. However, as the discussion above implies, a protein structure in complex with a longer substrate is needed to get an even better understanding of these enzymes. While scientifically interesting, the data presented in **Paper I** do not provide sufficient leads for achieving one of the long term biotechnological goals of this study, namely to engineer CE4s that generate specific deacetylation patterns on their substrates. To achieve this goal, enzyme-substrate interactions beyond subsites 0 and +1 need to be mapped.

3.2 Paper II – Proteomic investigation of the secretome of *Cellvibrio japonicus* during growth on chitin

The ability of the Gram-negative bacterium *C. japonicus* to degrade plant cell wall polysaccharides is documented well (Ueda et al., 1952, DeBoy et al., 2008, Gardner and Keating, 2010, Gardner et al., 2014), and the genome encodes a large number of CAZymes. These CAZymes include several putatively chitinolytic enzymes: two AA10s, four GH18s, four CE4s, two GH20s, one GH19, and one GH46. The two AA10s have been characterized and it is known that *CjLPMO10A* is chitin-active (Forsberg et al., 2016), while *CjLPMO10B* is cellulose-active (Gardner et al., 2014). The aim of the work described in **Paper II** was to do a quantitative proteomics study using LFQ to investigate the secretome from *C. japonicus* during growth with chitin as the sole carbon source. By doing so, one could identify which of the recognized CAZymes are important in chitin conversion and one could identify proteins of unknown function that could play additional roles in this process. Experiments were done using α -chitin and β -chitin as the relevant substrates and glucose as a control substrate.

The first approach in this study was to grow *C. japonicus* in shaking flask cultures. Samples were taken at different time points, and after removing cells and chitin particles, proteins present in the extracellular milieu were analyzed. After MS analysis and LFQ quantification using the MaxLFQ algorithm incorporated into the MaxQuant software (Cox and Mann, 2008, Cox et al., 2014), we found a high content of proteins predicted to be cytosolic (around 50 %, **Fig. 1, Paper II**). It is not unusual to have cytosolic proteins in secretome samples, for example, around 80 % of the proteins found in the secretome of *Staphylococcus epidermidis* were predicted to be cytosolic (Siljamäki et al., 2014). In a secretome analysis of *Thermobifida fusca* after growth on cellulose, 62 % of the identified proteins were predicted to not have a signal peptide (Adav et al., 2012). Although the presence of cytosolic proteins is common and usually not considered as a major problem, it is desirable to reduce the cytosolic fraction as much as possible for several reasons: 1) *In silico* prediction of cellular localization is not 100 % accurate and it is thus desirable to be as certain as possible that the detected “secreted” protein indeed is secreted. 2) During MS analysis, the presence of cytosolic proteins in high amounts may mask other (secreted) proteins present in low amounts.

One reason for the high content of cytosolic proteins in the shaking flask culture could be cell lysis caused by shearing forces created during the shaking with large chitin particles. Bengtsson et al. (2016) developed a plate method to look at secretomes from fungi growing

on insoluble substrates. They showed that the secretomes prepared contain a high fraction of proteins predicted to be secreted and we therefore adapted this method to bacteria. The results showed a reduction in the cytosolic fraction from around 50% using the shaking flask method, to around 30% using the plate method (**Fig. 1, Paper II**).

An additional advantage of the plate method concerns the sample preparation. During growth on chitin, it is expected that CAZymes are secreted, and many of them will bind to the insoluble substrate, especially those containing CBMs, which promote enzyme-substrate interactions. Enzymes acting on chitin may still be attached to chitin particles during sampling for proteomics analysis. In the shaking flask method, the chitin particles are removed before sample preparation, hence proteins bound to the chitin will not be identified. On the other hand, when using the plate method, the chitin particles are present during the sample preparation, which includes a boiling step, meaning that there is a bigger chance that proteins bound to the chitin will be solubilized and subsequently identified. Another advantage of the plate method is that it possibly resembles the natural growth conditions of *C. japonicus* better than the shaking flask cultures, as *C. japonicus* has rather stationary (i.e. non-liquid) growth conditions in its natural habitat in soil.

After establishment of the plate method and quantification of identified proteins using the MaxLFQ algorithm in MaxQuant, a thorough investigation of the identified proteins was performed, with initial focus on CAZymes, including enzymes containing a CBM but without any further CAZy annotation. A wide array of CAZymes was identified (**Fig. 2, Paper II**) and there were clear differences in abundance levels, between the enzymes and between the carbon sources. The 30 proteins belonging to Cluster 1 and 2 (**Fig. 2A, Paper II**) contain CAZymes that are clearly more abundant in the chitin samples, compared to the glucose sample, suggesting their potential importance in chitin degradation. Among these proteins, cluster 1 contains *CjLPMO10A*, *CjChi18C*, and *CjChi18D*, while cluster 2 contains *CjChi18A*, *CjChi18B*, and *CjChi19A*. So, of the 30 proteins in these two clusters, six have clear bioinformatics-based predicted functions that link them to chitin degradation. The heat map in **Fig. 2A of paper II** is based on the second time point taken during sampling. In total, three time points were analyzed in order to get an impression of the protein profile at an early, middle and late stage of the growth. At the early time point, chitin-modifying enzymes were only detected in the chitin samples, indicating that *C. japonicus* indeed utilizes its predicted putative chitinolytic machinery to degrade chitin. Additionally, of the putative chitinases only

CjChi18A, *CjChi18C*, and *CjChi18D*, were identified at the early time point in the α -chitin sample, suggesting that these are the most important chitinases in the beginning of the chitin degradation process (see Section 3.3 and **Paper III** for further discussion of the roles of the individual chitinases).

The presence of *CjChi18A* in the secretome is somewhat unexpected since it is predicted to be a lipoprotein (having a SpII signal peptide), that is attached to one of the cell membranes. Cleavage of the SpII signal peptide leaves a cysteine as the first amino acid in the mature protein, and it has been proposed that the amino acid following this cysteine determines if the protein should be attached to the inner or the outer membrane. A serine following the cysteine indicates anchoring to the periplasmic side of the inner membrane, while any other amino acid indicates anchoring to the outer membrane (Yamaguchi et al., 1988, Seydel et al., 1999). *CjChi18A* has a glutamine following this cysteine, suggesting anchoring to the outer membrane. Whether *CjChi18A* faces the periplasm or the extracellular space cannot be predicted and its functional characteristics (**Paper III**) do not suggest a particular location. Cell lysis is one explanation for finding *CjChi18A* in the secretome. However since *CjChi18A* is present in relatively high amounts one might speculate that it faces the extracellular space, and that it somehow has loosened from the cell surface. The orientation of *CjChi18A* is discussed further in Section 3.3 and **Paper III**. It should be noted that lipoproteins are commonly observed in secretomics studies [e.g. (Tjalsma et al., 2004, Fazzini et al., 2011, Tschumi et al., 2012)].

The genome of *C. japonicus* encodes two putative chitobiases belonging to GH family 20; *CjHex20A* and *CjHex20B*. *CjHex20A* has a predicted SpI signal peptide, while LipoP (Juncker et al., 2003) predicts *CjHex20B* to have a transmembrane helix. Only *CjHex20B* was observed in the secretome in low amounts and only in the β -chitin sample for the second time point, which is somewhat unexpected as *CjHex20B* is predicted to have a TMH. However, in some cases it is difficult to differentiate between signal peptides and TMHs. The TMHMM 2.0 server (Krogh et al., 2001) predicts *CjHex20B* to have one TMH, but the details of the prediction show that this (N-terminal) TMH could actually be a signal peptide. As described in the Introduction, the main task of chitobiases in a chitinolytic machinery is to degrade products formed by the chitinases, and chitobiases are often located in the periplasm (Toratani et al., 2008), which may be the reason to why only small amounts of *CjHex20B* and no *CjHex20A* are observed in the secretome samples. It is not certain whether *C. japonicus*

need or use both as part of the chitin utilization machinery. However, the transcriptomic analysis presented in **Paper III** showed that both *CjHex20A* and *CjHex20B* are up-regulated during growth on chitin, indicating a potential role of these proteins in chitin utilization. This is discussed further in Section 3.3 concerning **Paper III**.

Two of the four putative CE4 proteins encoded by *C. japonicus* were detected in low amounts. One of these is part of a multimodular protein (*CjXyn11A*), which has a GH11, a CE4, a CBM60, and a CBM10 domain and which was only found in the β -chitin samples. This protein is known to hydrolyze β -1,4 linkages and remove acetyl groups in xylan (Emami et al., 2002). Thus, while the CE4 domain of *CjXyn11A* may be active against chitin and CHOS, a purposeful role of this CE4 domain in chitin metabolism in *C. japonicus* is not very likely. The other CE4 protein was found in low amounts in all samples (cluster 6, **Fig. 2A, Paper II**), and a specific role in chitin utilization cannot be assigned for this protein. In addition, this putative CE4 protein lacks three of the five conserved motifs commonly found in chitin-active CE4 deacetylases (see Section 1.3.1.4.1). Notably, the other two putative CE4 proteins, not identified in the secretome and not predicted to be secreted, also lacks some of the conserved motifs characteristic for GlcNAc-active deacetylases.

Several other CAZymes without an obvious role in chitin degradation were enriched in the chitin samples, often occurring in high amounts (cluster 1 and 2, **Fig. 2a, Paper II**). Some of these proteins have putative roles in cellulose degradation, while some have putative activity against other plant cell wall polysaccharides. Additionally, some of these proteins have GH and CBM domains belonging to families for which chitin related activity has been described, in particular GH5, GH9, GH16, and CBM2 (Lombard et al., 2014). *CjCel5A* and *CjCel6A* (both having a CBM2 and CBM10 domain) were, together with *CjLPMO10B*, among the most abundant proteins in the chitin samples, and it is possible that these enzymes can have a role in the chitin degradation. Another possibility is that the presence of chitin leads to production of cellulose degrading enzymes as well, since the structure of chitin and cellulose is quite similar. This is discussed further in **Paper III** (Section 3.3).

Two other proteins found in high amounts in the chitin samples, without any obvious role in chitin degradation, are *CjCbp2D* and *CjCbp2E* (**Fig. 2A, Paper II**). Interestingly, these proteins are known to be important for *C. japonicus* when growing on cellulose (Gardner et al., 2014), and the proteomics data in **Paper II**, indicate that they also are important when the

bacterium grows on chitin. *CjCpb2E* only occurred in the chitin samples, while *CjCbp2D* occurred in 1000-fold higher amounts in the chitin samples compared to the glucose samples. In addition to a CBM2 domain, known to bind chitin and cellulose, these two proteins have a so-called YceI-like domain. Studies of a *CjCbp2E* homologue from *Saccharophagus degradans* have shown that its YceI-like domain binds polyisoprenoids, which are electron rich molecules (Vincent et al., 2010). *CjCbp2D* contains in addition two cytochrome C-like domains with potential roles in electron transfer (Vincent et al., 2010). Based on these observations, it is conceivable that *CjCbp2E* and *CjCbp2D* play a role in redox processes that relate to chitin degradation.

Gardner et al. (2014) showed that *CjCbp2D* and *CjCbp2E* are upregulated during growth on cellulose and that knocking out *CjCbp2D* and *CjCbp2E* led to impaired growth of *C. japonicus* on cellulose. The authors speculate that these proteins could be important in the oxidative cleavage of recalcitrant substrates, possibly by donating electrons to the LPMOs, similar to the activation of LPMOs by cellobiose dehydrogenase in fungi (Phillips et al., 2011, Loose et al., 2016). If electron donation to the LPMOs is the role of these proteins, one can envisage that *CjCbp2D* and *CjCbp2E*, with help of the CBM2 domains, bind close to the binding site of the LPMO on the substrate. This might cause a favorable proximity effect, ensuring reduction of the LPMO close to the substrate. Reduced LPMOs free in solution are unstable, and may produce reactive-oxygen species that can be harmful for the LPMO itself (Bissaro et al., 2017) and the cells. Although the role of *CjCbp2D* and *CjCbp2E* in degradation of cellulose and chitin remains unclear, the proteomics data presented in **Paper II**, together with the study by Gardner et al. (2014) and the transcriptomics data described in **Paper III**, strongly indicate an important function of these proteins in the degradation of crystalline substrates.

The secretomes did not only contain CAZymes and investigation of other proteins found in high abundance in the chitin samples compared to the glucose sample was of interest, especially from an enzyme discovery point of view. In Nature, chitin usually occurs in a network with proteins and minerals (see Section 1.1 for details), meaning that proteases that are tailored for dealing with chitin-rich biomass are of particular interest. Of the proteins identified in the secretomes, 31 proteins were given a protease annotation when searching the MEROPS peptidase database (Rawlings et al., 2016). Six of these proteins could be of interest as they occur in higher amounts in chitin samples (4 proteins), compared to glucose samples,

or because they contain a CBM that may bind chitin (2 proteins). One of the predicted proteases has a CBM6, and although chitin binding has not been reported for this CBM family, the finding of a CBM6 containing protein exclusively in the chitin samples is an indication that the CBM6s possibly can bind chitin (cluster 2, **Fig. 2A, Paper II**). The other protease contain a CBM50 domain commonly found associated with chitin and peptidoglycan active glycosyl hydrolases, in addition to proteases targeting peptidoglycan (Lombard et al., 2014). However, the CBM50 containing protease was found in relative equal amounts in all samples (cluster 6, **Fig. 3A, Paper II**), and does not seem to be chitin specific.

Another interesting prospect that opens up when investigating non-CAZymes is the possibility to discover proteins with unknown functions. Many of the detected non-CAZymes were found in similar amounts in all three secretomes, however, some were found exclusively or in higher amounts in the chitin samples (**Table 1, Paper II**). Of these, several have unknown or putative functions, hence they are interesting targets for future investigation of their biochemical properties.

Paper II demonstrates that several putative chitin-degrading proteins indeed are expressed in high levels when *C. japonicus* grows on chitin. Since the novel plate method used to generate the secretome samples showed a significant reduction in contamination of cytosolic proteins, compared to a shaking flask method, it is likely that several of the other detected proteins also play a role in chitin conversion and are not just “contaminants”. The proteases and the proteins with yet unknown functions found in higher amounts in the chitin samples make up an interesting target list for enzyme discovery and further investigation of their potential role in biomass conversion is warranted.

3.3 Paper III – Chitin degradation by *Cellvibrio japonicus* is dependent on the non-redundant *CjChi18D* chitinase

The four GH18 chitinases encoded by the *C. japonicus* genome were all detected in the secretome studies described in **Paper II** and their abundance in the various secretomes described in **Paper II** are summarized in Table 2. To further investigate the importance of these four GH18s in chitin degradation, transcriptomics, knock-out studies, and biochemical experiments were conducted as a follow-up to **Paper II**. The chitin-active LPMO, *CjLPMO10A* was also included in this study. This discussion focuses on the biochemical characterization of the chitinases, since the other parts of **Paper III** were mainly produced by collaborators.

Table 2: Key features of the GH18 chitinases and *CjLPMO10A*. The four GH18 chitinases encoded by *C. japonicus* and *CjLPMO10A* are listed with their LipoP prediction, log₁₀ of LFQ values (average of three biological parallels) from the proteomics study presented in **Paper II**, and their domain organization (see also **Fig. 2, Paper III**). ND = not detected.

Chitinase	LipoP	LFQ intensities proteomics (log ₁₀)			Domain organization
		α -chitin	β -chitin	Glucose	
<i>CjChi18A</i>	SpII	9.1	8.9	ND	GH18
<i>CjChi18B</i>	SpI	8.4	10.4	7.4	CBM73-GH18-CBM5
<i>CjChi18C</i>	SpI	9.1	10.8	8.2	CBM5-CBM73-CH18
<i>CjChi18D</i>	SpI	9.5	11.0	8.7	CBM73-CBM5-GH18
<i>CjLPMO10A</i>	SpI	9.4	10.9	8.4	AA10-CBM5-CBM73

By making and testing several deletion mutants of the GH18 chitinases, it became clear that *CjChi18D* is essential for the ability of *C. japonicus* to utilize α -chitin as a sole carbon source. The $\Delta chi18D$ single mutant was unable to grow on α -chitin and crab shells, while the other single chitinase mutants had similar growth characteristics as the wild type (**Figs. 3A and 3C, Paper III**). In a previous study it was shown that a $\Delta lpmo10A$ mutant led to growth defects when *C. japonicus* grew on β -chitin and crab shells (Forsberg et al., 2016). However, the growth defect was not as severe as observed for the $\Delta chi18D$ mutant. The double mutant $\Delta chi18B \Delta chi18C$ showed a 37 % reduction in growth rate on α -chitin compared to the wild type. The double mutants $\Delta chi18A \Delta chi18B$ and $\Delta chi18A \Delta chi18C$ showed similar growth as the wild type. The triple mutant $\Delta chi18A \Delta chi18B \Delta chi18C$ showed the same growth defect as the double mutant $\Delta chi18B \Delta chi18C$ (**Figs. 3B and 3D, Paper III**). These results suggest

that *CjChi18A* is not critical for growth on chitin and that *CjChi18B* and *CjChi18C* play partly overlapping roles (since single knock-outs had no effect, whereas the double knock-out had an effect). It is worth noting that, when using crab shells as carbon source, all double mutants not including $\Delta chi18D$, had wild type-like phenotypes. Apparently, with this substrate, the contribution of *CjChi18B* and *CjChi18C* is not rate-limiting for growth of *C. japonicus*.

The knock-out results indicated different roles of the four chitinases, and biochemical characterization of the chitinases was done to explore this further. Unfortunately, full-length proteins were difficult to obtain, and all other versions, except *CjChi18A_{cat}*, formed inclusion bodies during production. A denaturing and refolding method enabled production of soluble catalytic domains of *CjChi18B*, *CjChi18C*, and *CjChi18D*. In addition, *CjChi18B_{cat}+CBM5* and *CjChi18D_{cat}+CBM5* were produced in soluble form. Testing the ability to degrade α -chitin revealed differences in enzyme properties (**Fig. 5, Paper III**), with *CjChi18D_{cat}+CBM5* having highest activity. *CjChi18A_{cat}* and *CjChi18B_{cat}* are the poorest α -chitin degraders, while *CjChi18B_{cat}+CBM5* and *CjChi18C_{cat}* perform equally well, but still less well than *CjChi18D_{cat}+CBM5* and *CjChi18D_{cat}*. The activity data for α -chitin, showing that *CjChi18D* is particularly powerful, align well with the studies of the knock-out mutants which showed that only *CjChi18D* is essential for chitin degradation.

The characterization of the chitinases revealed additional functional differences that are interesting. After degradation of α -chitin, the product profile for *CjChi18A_{cat}* revealed GlcNAc as the main product, while for the other enzymes (GlcNAc)₂ was the main product (**Fig. S3, Paper III**). Synergy experiments (**Fig. 6, Paper III**) done by mixing the catalytic domains of the four chitinases and *CjLPMO10A* in various combinations, showed that the presence of *CjChi18A_{cat}* increased the amount of monomeric products, indicating that *CjChi18A_{cat}* degrades the dimers produced by the other chitinases into monomers. Synergistic effects were observed for various combinations of enzymes indicating that the enzymes have different roles during chitin conversion. Addition of *CjLPMO10A* (catalytic domain) to the reactions had a positive effect on the activity of *CjChi18A_{cat}*, *CjChi18B_{cat}*, and *CjChi18C_{cat}*, but had limited effect on the activity of *CjChi18D_{cat}*. Generally, the effects of the LPMO were rather limited, which is probably due to the nature of the substrate, as discussed in **Paper III**. The synergy experiments showed that *CjChi18B_{cat} + CjChi18D_{cat}* resulted in higher product formation than *CjChi18C_{cat} + CjChi18D_{cat}*, which may seem surprising, as *CjChi18C_{cat}* is a more efficient chitin degrader than *CjChi18B_{cat}* (**Fig. 5, Paper III**). On the other hand, it may

seem that both *CjChi18C* and *CjChi18D* have a considerable endo-character, whereas *CjChi18B* likely is exo-processive (discussed further below) and is thus more likely to act synergistically with *CjChi18D* than *CjChi18C*. The highest product yields were obtained with enzyme mixtures containing *CjChi18B_{cat}*, *CjChi18C_{cat}*, and *CjChi18D_{cat}* and the product levels could not be improved further by adding *CjChi18A_{cat}*. Thus, all chitinases, with the possible exception of *CjChi18A*, seem needed for optimal chitin degradation *in vitro*. Degradation reactions with an enzyme cocktail composed of the four chitinases and *CjLPMO10A* according to the quantitative proteomics data presented in **Paper II** and summarized in Table 2 did not give higher product yields (**Fig. 5, Paper III**).

To further investigate substrate specificity and product formation by the chitinases, degradation of (GlcNAc)₆ was tested, using the catalytic domains. The results showed that *CjChi18C_{cat}* was the most efficient degrader of this substrate (**Fig. 7A, Paper III**) with an initial rate of $300.2 \pm 7.7 \text{ min}^{-1}$. *CjChi18A_{cat}*, *CjChi18B_{cat}*, and *CjChi18D_{cat}* had initial rates of 118.1 ± 12 , 27.0 ± 0.8 , and $44.5 \pm 2.6 \text{ min}^{-1}$, respectively. It is interesting to note that the four enzymes rank differently when it comes to activity on (soluble) chitohexaose, compared to α -chitin. Again, clear functional differences between the chitinases are observed.

The product profiles emerging during degradation of (GlcNAc)₆ can give hints concerning the mode of action of the chitinases. After two minutes reaction time, when only a minor fraction of the substrate had been converted, there were clear differences in the product profiles of different chitinases: *CjChi18A_{cat}* produced all product types [i.e. (GlcNAc)₁₋₅], *CjChi18B_{cat}* produced mainly dimers and trimers, *CjChi18C_{cat}* produced (GlcNAc)₂₋₄, while *CjChi18D_{cat}* produce mainly dimers and tetramers (**Fig. 7B, paper III**). For the *CjChi18D_{cat}* reaction, a peak eluting prior to the hexamer peak evolved over time (**Figs. 7B and S4, Paper III**), suggesting production of CHOS longer than (GlcNAc)₆, which again indicates that *CjChi18D_{cat}* has transglycosylating abilities.

The ratio between different degradation products during degradation of α -chitin and (GlcNAc)₆ can give information about a potential processive mode of action. Processive enzymes acting on chitin are expected to give a high (GlcNAc)₂/GlcNAc ratio since for processive enzymes only initial binding can lead to production of an odd-numbered product (which will usually lead to eventual production of a monomer), whereas all further catalytic events resulting from the same initial productive enzyme-substrate association will yield

dimers. Enzymes acting processively on (GlcNAc)₆ are expected to show a high (i.e. non-stoichiometric) (GlcNAc)₂/(GlcNAc)₄ ratio, as processive activity on a hexamer would result in production of three (GlcNAc)₂. A non-processive mode of action towards (GlcNAc)₆ would result in (GlcNAc)₄ + (GlcNAc)₂ or (GlcNAc)₃ + (GlcNAc)₃. Using these criteria, only *CjChi18B_{cat}* showed clear signs of processivity (**Table S4, Paper III**). Interestingly, *CjChi18B_{cat}* is slowest of the chitinases towards (GlcNAc)₆, which coincides well with previous observations that processive activity comes at the cost of a reduced rate against soluble substrates (Horn et al., 2006a). On the other hand, one might expect a processive enzyme like *CjChi18B* to show high activity towards chitin, which was not observed for *CjChi18B_{cat}*. While *CjChi18B_{cat}+CBM5* shows higher activity than the catalytic domain alone, this enzyme is still not as good as *CjChi18D_{cat}* and *CjChi18D_{cat}+CBM5* in degrading α -chitin. Notably, both proteomics (**Paper II**) and transcriptomics data (discussed below, **Paper III**) suggest that *CjChi18B* is not the most highly expressed chitinase.

It is difficult to address processivity experimentally (Horn et al., 2012a) and the results discussed above give indications but are not conclusive. There is a need for more experimental data to confirm whether the chitinases are processive or not, preferable with the full-length enzymes. As described in Section 1.3.1.1, the amino acid sequences and three-dimensional structures of known processive chitinases show some characteristic features. The most important features are the presence of aromatic residues in certain subsites (Horn et al., 2006a, Zakariassen et al., 2009) and a “deep” substrate-binding cleft where depth is provided in by the presence of the so-called $\alpha+\beta$ domain as well as loops protruding from this domain or from other parts of the catalytic domain (Perrakis et al., 1994, Van Aalten et al., 2000, Vaaje-Kolstad et al., 2013). The crystal structures of the GH18 chitinases of *C. japonicus* are not available; however, structural modelling of the proteins and investigation of the amino acid sequence was feasible (**Figs. S6 and S7, Paper III**).

The structural models of the catalytic domains of the chitinases show that *CjChi18B* has a deep, tunnel-shaped substrate-binding cleft formed by loops, including the $\alpha+\beta$ domain which is characteristic for processive chitinases. In addition, *CjChi18B* has aromatic residues that seem to be conserved in processive chitinases, affecting substrate binding in subsites -3 to +2. Notably, the structural model of *CjChi18B* indicates that it may have a more closed active site compared to the well-studied processive enzymes *SmChiA* and *SmChiB*. Thus, both the predicted structural features and the biochemical data indicate that *CjChi18B* is a processive

enzyme. *CjChi18A* has a relatively deep binding pocket due to the presence of the $\alpha+\beta$ domain, but there is no tunnel and several of the aromatic residues found in processive chitinases are absent. Together with the biochemical data, these structural features suggest that *CjChi18A* is a non-processive enzyme. Its deep active site cleft may allow relatively strong binding to shorter substrates, and this coincides with the enzyme's high rate towards chitohexaose and its putative function in processing of soluble substrates. The structural model of *CjChi18C* shows an open active site typical for non-processive endo-acting enzymes, whereas aromatic residues typical for processive enzymes are lacking. So, the predicted structural features indicate that *CjChi18C* is non-processive. The substrate-binding cleft of *CjChi18D* is more open compared to *CjChi18A* as it lacks loops on the opposite side of the substrate-binding cleft relative to the $\alpha+\beta$ domain. In contrast to *CjChi18A*, *CjChi18D* does have the aromatic residues that are characteristic for processive chitinases (in subsites -3 to +2). It may seem as if *CjChi18D* is a "hybrid" enzyme, with an open active site, enabling endo-type binding to the substrate (see below), and having aromatic residues to allow strong binding to crystalline material and perhaps giving some degree of processivity (that has remained undetected in the biochemical assays).

Experimental determination of endo- versus exo-activity is not easy neither, but also here the structural models give hints. It seems rather certain that *CjChi18B* is exo-processive, whereas *CjChi18C* is endo-non-processive. Whether *CjChi18A* and *CjChi18D* are endo- or exo-enzymes is difficult to predict, but by analogy to studies on *S. marcescens* chitinases (Sikorski et al., 2006) their relative open active sites indicate that both enzymes have considerable endo-activity. Judged by the structural model, *CjChi18D* seems endo-processive and it is somewhat surprising that this is not supported by observed product ratios. The combination of high activity on crystalline substrates and lower activity on soluble substrates, as is the case for *CjChi18D* has been observed previously for processive enzymes (Horn et al., 2006a). The product profile for *CjChi18D* upon degradation of (GlcNAc)₆, showing a peak of CHOS with DP longer than 6 (**Figs. 7B and S4, Paper III**), indicates that this chitinase has transglycosylating properties. If *CjChi18D* uses e.g. (GlcNAc)₂ as substrate during transglycosylation, this will of course affect the product profile, potentially leading to a wrong impression when it comes to whether this enzyme is processive or not.

It should be noted that the conclusions drawn from the biochemical data presented in **Paper III** must be considered critically, since we were not able to produce the full-length proteins.

The lack of one or more CBMs could affect the enzyme characteristics that were studied, especially when it comes to chitin degradation, as CBMs are known to aid in enzyme-substrate interactions (Boraston et al., 2004). *CjChi18B*, *CjChi18C*, and *CjChi18D* all contain a CBM5 and a CBM73 organized differently relative to the catalytic domains (Table 2 and **Fig. 2, Paper III**). The organization of the CBMs might be intentional and affect the enzyme specificities. For example, the lack of CBMs could explain the somewhat poor activity of *CjChi18B* towards chitin, compared to *CjChi18D*. If *CjChi18B* is a processive chitinase, it might be more dependent on the CBMs in order to stay attached to the substrate. Regardless of the presence of the CBMs or not, the biochemical data for the catalytic domains show clear functional differences between the enzymes, which are accompanied by clear differences in their predicted structures.

The genome of *C. japonicus* and the proteomics data described in **Paper II** suggest that other proteins could be involved in chitin utilization. To investigate this further and to complement the proteomics study (**Paper II**), RNAseq was used to investigate the change in gene expression during growth on α -chitin. In the exponential growth phase, 73 CAZymes were significantly up-regulated when comparing the transcriptome of α -chitin- versus glucose-grown cells (**Table S3A, Paper III**). Thirteen of these genes encode proteins with a predicted chitin-related activity and the four chitinases and *CjLPMO10A* were among the proteins with highest log₂ fold change (**Fig. 8 and Table S3A, Paper III**). A similar comparison was also done in the early stationary growth phase, revealing that 47 CAZyme-encoding genes were up-regulated during growth on chitin (**Fig. S5A and Table S3B, Paper III**), and seven of these have a putative role in chitin utilization. In both growth phases, the *lpmo10A* and *chi18D* genes showed the highest fold change, supporting their crucial role in chitin degradation, as shown by Forsberg et al. (2016) and in **Paper III**, respectively.

Other chitin utilization related genes that were up-regulated during growth on chitin were *hex20B*, *nag9A*, *hex20A*, *csn46F*, and *chi19A*. The *hex20* genes encode GH20 chitobiases likely to cleave GlcNAc dimers into monomers, a process likely to take place in the periplasm (Toratani et al., 2008). The *nag9A* gene encodes a cytoplasmic *N*-acetylglucosamine-6-phosphate deacetylase (CE family 9) with a putative role in GlcNAc utilization. Neither *CjHex20A* nor *CjNag9A* were identified in the proteomics study presented in **Paper II**, and *CjHex20B* was found in very low amounts only in the β -chitin samples, supporting the notion that the main location of these proteins is inside the cell. If cell lysis was a problem in **Paper**

II, the transcriptomics data suggest that *CjHex20A*, *CjNag9A* and *CjHex20B* would have been found in relatively high amounts in the proteomics data as well. The fact that this is not the case may be taken to indicate that cell lysis was not really a problem, adding confidence to the results presented in **Paper II**. One consequence of this is that we may assume that the detection of lipo-anchored *CjChi18A* in the proteomic data in **Paper II** indeed indicates an extracellular orientation of this lipoprotein. In addition, an *N*-acetylhexosaminidase-like role of *CjChi18A* in the periplasm seems redundant as *CjHex20A* and/or *CjHex20B* are expected to convert (GlcNAc)₂ into GlcNAc in the periplasm. If *CjChi18A* is anchored to the outer membrane facing the extracellular space the production of GlcNAc monomers could for instance be important in some cell signaling processes or control of the (GlcNAc)₂ concentrations can be important to avoid product inhibition of the other chitinases.

In accordance with the findings described in **Paper II**, many genes encoding CAZymes related to degradation of other polysaccharides were up-regulated suggesting a general response to chitin, although chitin-degrading enzymes in general showed higher log₂ fold changes (**Table S3, Paper III**). The same general response was found when *C. japonicus* grows on cellulose (Gardner et al., 2014), while the response when growing on xylan is more specific (J.G. Gardner et al., unpublished observations). It is perhaps not surprising that chitin and cellulose elicit similar general responses, considering the similarities of these polysaccharides, including their recalcitrance.

Paper II lists 24 non-CAZyme proteins (**Table 1, Paper II**) that occur in higher abundance in the chitin samples compared to glucose. The RNAseq data produced for **Paper III** show that the expression of nine of these proteins is significantly up-regulated during growth on chitin in the exponential growth phase. Four of these proteins were exclusively in the secretomes of chitin-grown cells (Table 3), making these proteins of special interest for future investigation of their potential role in chitin utilization by *C. japonicus*. Another gene of particular interest, listed among the CAZymes (**Table S3A, Paper III**), encodes a CBM6 attached to a predicted protease domain, and also this protein was exclusively found in the chitin samples in the proteomics study (**Paper II**). This gene is also significantly up-regulated during growth on chitin (Table 3), suggesting a potential role of this protease in utilization of chitin-rich biomasses.

Table 3. Proteins potentially involved in degradation of chitin-rich biomasses. The table lists five proteins that seem of particular interest for future investigations of natural chitin converting enzyme systems, as explained in the text. The LFQ- and fold change-values listed are averages of three biological replicates (except the LFQ-value for B3PL02, which is an average of two biological replicates). For comparison, the \log_{10} LFQ- and \log_2 fold change-values for *CjChi18D*, the most expressed GH18 chitinase, and *CjChi18B*, the least expressed GH18 chitinase upon growth on α -chitin, were 9.5 and 6.66, and 8.4 and 2.45, respectively.

Uniprot ID	Locus ID	Predicted function	LFQ intensities proteomics (\log_{10})	Fold change transcriptomics (\log_2), α -chitin vs glucose
			α -chitin	Exponential growth
B3PL02	CJA_2515	Uncharacterized protein	7.9	1.67
B3PB26	CJA_0998	Putative lipoprotein	8.5	1.21
B3PFC4	CJA_0040	Pectin metylesterase, putative	8.2	1.4
B3PKZ2	CJA_2505	Uncharacterized protein	8.2	1.2
B3PH79	CJA_0276	CBM6 + protease	8.3	2.67

Collectively the data presented in **Paper III** show that *CjChi18D* is the most important chitinase for the ability of *C. japonicus* to utilize α -chitin. The biochemical characterization of the chitinases confirmed that *CjChi18D* is the most efficient chitin degrader and a structural model suggests that this enzyme indeed is well suited to interact with crystalline material, possibly acting endo-processively. The data also suggest that *CjChi18B* is a processive exo-enzyme, whose contribution to chitin degradation possibly overlaps with the role of the putatively non-processive endo-acting *CjChi18C*. The transcriptomics data, together with the proteomics data in **Paper II**, suggest that *C. japonicus* responds to chitin in a general way, leading to transcription and secretion of a variety of CAZymes.

3.4 Paper IV – Genomic, proteomic and biochemical analysis of the chitinolytic machinery of *Serratia marcescens* BJL200

The chitinolytic machinery of *S. marcescens* has been studied in detail [reviewed by Vaaje-Kolstad et al. (2013)], and includes three GH18 chitinases, one AA10 LPMO, and a GH20 chitobiase (see Section 1.4.1). However, a thorough proteomic study on this bacterium upon growth on chitin had not been conducted at the start of this work. In addition, the genome of one of the most studied strains (BJL200) was not available. **Paper IV** describes the genome sequence of *S. marcescens* BJL200 as well as a proteomics study performed using the plate method developed for *C. japonicus* and described in **Paper II**. The genome sequencing resulted in more than 5000 predicted coding sequences, of which 160 were predicted to encode CAZymes. Interestingly, these CAZymes includes a fourth putative GH18 protein, designated *SmChiD*. The genome does not contain any genes encoding CE4 proteins, indicating that chitin deacetylases are not important for chitin utilization by this organism. Detailed investigation of the genome revealed that the genes encoding the chitinase *SmChiB* and the LPMO *SmCBP21* are organized in a similar manner as in *S. marcescens* Db10/11 (see Section 1.4.1), together with a LysR family regulator, a holin, an endolysin and two spanins (Hamilton et al., 2014). Comparable to *S. marcescens* Db10/11, *S. marcescens* BJL200 lacks a T2SS, which in e.g. *C. japonicus* is shown to be essential for secretion of CAZymes (Gardner and Keating, 2010).

Unexpectedly, having **Paper II** in mind, the proteomics data showed a rather high content of cytosolic proteins in the secretome samples (45 – 60 %; **Fig. 1, Paper IV**). Next to unintended cell lysis, there are several explanations for this observation: 1) These proteins could be so-called “moonlighting proteins”, without a classical signal peptide, that have both a cytosolic and an extracellular function (Bendtsen et al., 2005a). 2) Outer membrane vesicles are produced by Gram-negative bacteria including *S. marcescens* (McMahon et al., 2012); these vesicles often contain cytosolic proteins which could end up in the secretome. 3) Intentional cell lysis that is initiated in some cells to benefit the rest of the community (Allocati et al., 2015).

Another possible explanation for the high cytosolic content relates to the not yet fully understood system *S. marcescens* uses to secrete the chitinolytic machinery. This machinery, which was briefly introduced in Section 1.4.1 and 1.5, may also be used to secrete other proteins, which likely would be annotated as cytosolic. Studies of *S. marcescens* Db10/11

have shown that secretion of the chitinolytic machinery depends on a holin-endolysin system (Hamilton et al., 2014). While *SmChiB* and *SmChiC* lack a classical signal peptide Hamilton et al. (2014) showed that secretion of *SmChiA* and *SmCBP21*, both having a signal peptide, also is dependent on this system. Holins are thought to translocate folded proteins without a signal peptide across the inner membrane (Desvaux and Hébraud, 2006), but the holin of *S. marcescens* Db10/11 is not essential for translocation of *SmChiC* to the periplasm (Hamilton et al., 2014). Hence, the translocation process of *SmChiC* and *SmChiB* to the periplasm remains unclear, whereas, according to Hamilton et al. (2014), *SmChiA* and *SmCBP21* are translocated to the periplasm through the Sec-system. Holin-endopeptidase systems are known to be involved in phage-induced cell lysis in Gram-negative bacteria. In this process the holin translocates the endopeptidase from the cytoplasm to the periplasm. In the periplasm, the endopeptidase makes pores in the peptidoglycan layer, and the spanin proteins are thought to then bring the inner and outer membranes together, which leads to cell lysis and release of proteins to the extracellular space (Rajaure et al., 2015). However, Hamilton et al. (2014) show evidence against cell lysis being the process for releasing the chitinolytic machinery into the extracellular space. Our proteomics data support the hypothesis that the detected extracellular proteins were not released into the extracellular space due to cell lysis. If cell lysis had occurred, we would expect to find the periplasmic GH20 chitobiase in the secretome samples, and this is not the case (**Fig. 2, Paper IV**). A control experiment with lysed cells confirmed that chitobiase is indeed produced in large amounts when *S. marcescens* grows on chitin. Since unintended lysis is an unlikely explanation for our observations, other, yet unknown, mechanisms must be involved in the secretion of the chitinolytic machinery and other proteins from the periplasm to the extracellular space.

Genome analysis shows that *S. marcescens* BJL200 possesses the T5SS, which could be an alternative route to translocate proteins from the periplasm to the extracellular space. As mentioned in Section 1.5, T5SS mainly secretes virulence factors and proteins involved in cell-cell adhesion and biofilm formation. Interestingly, both chitinases and LPMOs are proposed to have a role in virulence in pathogenic bacteria. In addition, Kawada et al. (2008) showed that knocking out *SmCBP21* leads to decreased ability of *S. marcescens* to adhere to colonic epithelial cells. Taken together, translocation of the chitinolytic machinery over the outer membrane through the T5SS is perhaps a reasonable hypothesis. However, this hypothesis does not explain the role of the holin-endopeptidase system nor explains the relatively high content of cytosolic proteins in the secretome samples.

The secretome samples showed that *SmChiA*, *SmChiB*, *SmChiC* and *SmCBP21* all were more abundant in the samples from cells grown on chitin samples compared to the glucose-grown cells (**Fig. 2, Paper IV**), confirming their important role in chitin degradation by *S. marcescens*. *SmChiA* and *SmCBP21* were more abundant than *SmChiB* and *SmChiC*, suggesting that the two former proteins are more important in chitin degradation. Interestingly, functional data indicate that *SmChiA* has higher activity towards α -chitin, compared to *SmChiB* and *SmChiC* (Horn et al., 2006b), and it is now well known that *SmCBP21*, a chitin-active LPMO, boosts the hydrolytic activity of the chitinases (Vaaje-Kolstad et al., 2010).

Interestingly, few additional CAZymes were found in the secretome, indicating that there is a specific response to chitin when *S. marcescens* utilizes chitin as a carbon source. One of the few CAZymes found in the secretomes is the putative fourth chitinase, *SmChiD*, which is not present in the glucose samples, and only in low amounts in the chitin samples, suggesting that this enzyme is chitin-specific, however, not important for chitin degradation.

In an attempt to elucidate the possible biological role of *SmChiD*, we carried out a biochemical characterization of the enzyme. *SmChiD* was able to degrade chitin, although with very low efficiency compared to the other chitinases (**Fig. 5, Paper IV**). An interesting observation was that (GlcNAc)₂ products decreased over time, while GlcNAc continued to increase (**Figure 4, Paper IV**), suggesting that *SmChiD* was able to hydrolyze (GlcNAc)₂ into monomers. The main degradation product produced by the other chitinases was (GlcNAc)₂, therefore the effect of *SmChiD* on the product profile when combined with the other chitinases were tested. Indeed, addition of *SmChiD* led to increased levels of GlcNAc (**Fig. 5, Paper IV**), but the enzyme had no positive effect on the overall efficiency of chitin degradation. These results suggest that the role of *SmChiD* could be to convert (GlcNAc)₂ into GlcNAc, which would mean that *SmChiD* has a similar role as the GH20 chitobiase. Comparing the Michaelis-Menten kinetics of chitobiase and *SmChiD* clearly shows that the chitobiase is much more powerful when it comes to hydrolyzing (GlcNAc)₂, making 1000-fold more cuts per second (**Fig. 6, paper IV**). Hence, it seems unlikely that conversion of (GlcNAc)₂ into monomers is the biological role of *SmChiD*.

Noteworthy, (GlcNAc)₂ type of structures are found in other biological systems, e.g. in glycoproteins, and chitinases have been proposed to target glycoproteins during pathogenesis

(Frederiksen et al., 2013). The activity of *SmChiD* against the glycoprotein RNaseB, containing a (GlcNAc)₂ at the core of its glycosyl moieties, was tested, however no activity could be detected against this glycoprotein (**Fig. 7, Paper IV**). Nevertheless, the possibility that *SmChiD* targets (other) glycoproteins cannot be ruled out.

A homologue of *SmChiD* found in *S. proteamaculans* (*SpChiD*) is well-characterized although its role in chitin conversion has not been addressed (Madhuprakash et al., 2012, Purushotham and Podile, 2012, Madhuprakash et al., 2014, Vaikuntapu et al., 2016). The two proteins share 85.7 % sequence identity, and, interestingly, the structure of *SpChiD* shows a loop that occludes part of the substrate binding cleft (Madhuprakash et al., 2013). This loop is also present in *SmChiD* but absent in the other *S. marcescens* chitinases, and is likely affecting the catalytic properties of *SpChiD*. In addition to being capable of hydrolyzing different CHOS, *SpChiD* shows considerable transglycosylating activity. The high sequence identity between *SmChiD* and *SpChiD* suggests that the two enzymes have similar properties. Although not important in chitin degradation, *SmChiD* could be useful for biotechnological production of CHOS due to its probable transglycosylating properties.

As discussed in connection to **Paper II**, proteomics studies open up the possibility to discover new proteins possibly involved in chitin degradation. Proteases are of special interest as chitin in Nature tends to exist in complex co-polymeric networks that are rich in proteins (Nikolov et al., 2011). Cluster 1 in **Fig. 2, Paper IV**, comprising proteins that were found in all samples, also including the glucose samples, contains three proteins annotated as secreted alkaline metalloproteases and a Blast against the MEROPS database [www.ebi.ac.uk/merops/ (Rawlings et al., 2016)] showed that they belong to metallo-peptidase family M10B. This protease family contains serralysins, which are considered as virulence factors in some pathogenic *Serratia* strains. Interestingly members of the M10B family are synthesized without a signal peptide and secreted by an unknown mechanism. Additionally, it has been shown that a serralysin in *Serratia* sp. KCK enhances chitinolytic activity towards crude chitin (Kim et al., 2007). Together, these observations and considerations are suggestive of a role in chitin degradation, but the proteomics data indicate that their expression is not regulated by the presence of chitin. Clusters 4, 5 and 6 (**Fig. 2, Paper IV**) comprise proteins that were found in higher amounts during growth on chitin compared to glucose, and include five hypothetical proteins that occur in higher amounts in the chitin samples, and that thus could be of interest for future studies.

The data presented in **Paper IV** confirmed that the well-characterized chitinolytic machinery indeed is secreted in high amounts when *S. marcescens* grows on chitin. The genome sequencing revealed a fourth putative chitinase encoded by the bacterium, and although this chitinase (*SmChiD*), is able to degrade chitin, the data indicate that its biological role is not related to chitin degradation. The proteomics data reveal a few interesting targets for further investigation of their possible role in conversion of natural chitin, which is of interest in order to make the chitin extraction process more environmentally friendly, as discussed in Section 1. As for *SmChiD*, further functional studies, e.g. involving the studies of the physiology of knock-out mutants of *S. marcescens*, could be of interest, for example to assess a potential role in virulence, a process that could also involve some of the proteases detected in this study.

4 CONCLUDING REMARKS AND PERSPECTIVES

Exploration of enzymes that may be used in the valorization of chitin-rich biomass is important to facilitate the shift towards a green economy. The water-soluble chitin derivatives chitosan and CHOS, commonly produced using harsh chemicals, have several applications in e.g. biomedicine, agriculture, and the food industry. The bioactivity of chitosans and CHOS is dependent on several physicochemical properties like F_A , DP, and P_A . Increased use of enzymes will not only lead to more environmentally friendly production processes but may also allow production of well-defined chitosans and CHOS. In light of this, investigation of enzymes used by microorganisms growing on chitin is of interest. Chitinases and LPMOs from various chitinolytic microorganism have been characterized. However, as chitin in chitin-rich biomasses exist in complexes with proteins and minerals, other enzymes used by the microorganisms could also be of industrial interest e.g. by making the chitin extraction process greener and more efficient. Recent data (Mekasha et al., 2017) indicate that chitin-rich biomass is remarkably recalcitrant, suggesting that there also is a need for more powerful chitinases.

CE4 deacetylases acting on GlcNAc can in principle be used to produce chitosan and CHOS with well-defined F_A and P_A . For optimal utilization of these enzymes, understanding their substrate specificities is important. Only a few CE4 deacetylases with very specific deacetylation pattern have been described (Andrés et al., 2014, Hamer et al., 2015, Cord-Landwehr et al., 2016, Naqvi et al., 2016). **Paper I** describes the structural and functional characterization of a CE4 deacetylase and include the first crystal structure of an enzyme-substrate complex of a CE4 deacetylase with an open active site. The enzyme showed broad substrate specificity, and the enzyme-substrate complex provided valuable insight into how the enzyme may interact with CHOS. Considering the broad substrate specificity of CE4 deacetylases with an open active site, a structural complex with another substrate, e.g. an acetyl xylan, would be of great interest (and necessary) to get an even better understanding of how these enzymes interact with their substrates. Available data on CE4 deacetylases with an open active site (Blair et al., 2005, Blair et al., 2006, Liu et al., 2017), including the one described in **Paper I**, suggest that utilization of these enzymes for tailoring the F_A and P_A of chitosans and CHOS can be challenging, since they have broad specificities and since little is known about enzyme-substrate interactions, if any, beyond subsites 0 and +1. Randomization of (putative) substrate binding sites by mutagenesis followed by screening for mutants that

generate changed acetylation patterns in model substrates could be one way to go, both to obtain more insight and to develop better biocatalysts.

Papers II and IV describe proteomic investigations of *C. japonicus* and *S. marcescens*, respectively, upon growth on chitin. By adapting a novel plate-based method originally developed for fungi, we were able to quantify proteins secreted by the bacteria and compare the abundance levels during growth on different substrates. In addition to confirming that assumed chitinolytic proteins indeed are produced in high amounts during growth on chitin, the proteomics studies revealed other (non-CAZy) proteins whose expression seemed to be chitin specific. These proteins are interesting targets for future investigation, since they may play a role in degradation of chitin-rich biomass.

A chitin extraction process and production of chitosan and CHOS from chitin-rich biomass exclusively using enzymes is theoretically possible. However, such a process must be optimized to be as efficient as the chemical processes used today, meaning that one needs more than only chitinases, e.g. to degrade the protein fraction. The experiments of Paper II and Paper IV were carried out with model substrates (extracted chitin) and the bacteria probably produce other proteins than the ones targeted in **Paper II** and **Paper IV** when growing on true natural substrates. Thus, proteomics studies with non-processed chitin-rich biomasses, would be of interest. Lab-scale enzymatic chitin extraction from chitin-rich biomasses has been tested, however, so far, enzymatic processes are not able to compete with the chemical processes when it comes to the purity of the chitin (Younes and Rinaudo, 2015, Younes et al., 2016).

Paper II describes studies of the chitinolytic machinery of *C. japonicus*, which showed that the four putative GH18 chitinases encoded by the bacterium were produced in high amounts. The study presented in **Paper III** was conducted to investigate the importance of the individual GH18 chitinases. This latter study showed that *CjChi18D* is crucial for the bacterium's ability to utilize chitin as a carbon source. The biochemical characterization of the chitinases revealed different properties when hydrolyzing chitin and CHOS, suggesting that the enzymes have varying and complementary roles in the chitin utilization. The finding that *CjChi18D* is essential for *C. japonicus* to survive with chitin as the sole carbon source suggests that this is a powerful chitinase, and further exploration of this enzyme in industrial valorization of chitin is thus of interest. Notably, Mekasha et al. (2017) recently showed that

the well-known and supposedly powerful chitinase cocktail produced by *S. marcescens* is capable of degrading industrially extracted chitin, but only at rather high enzyme dosages. It would be interesting to see if the *C. japonicus* cocktail or perhaps *CjChi18D* only could improve the degradation efficiency.

In addition to describing a proteomics investigation of *S. marcescens* growing on chitin, **Paper IV** reports the genome sequence of *S. marcescens* BJL200. The genome sequence showed that a fourth chitinase is encoded by this Gram-negative bacterium and the proteomics data suggested that this chitinase is not important in chitin utilization. Biochemical characterization of this chitinase, *SmChiD* showed that this GH18 chitinase is a poor chitin degrader compared to the three other chitinases encoded by *S. marcescens*. Thus, another, hereto unknown, biological role is proposed for this chitinase. However, what this potential biological role could be, remains unclear. Considering the seemingly minimal role of *SmChiD* in chitin degradation it seems that the lack of this enzyme in earlier work on the potential of the other *S.marcescens* chitinases, *SmChiA*, *SmChiB* and *SmChiC*, in the degradation of chitin (Horn et al., 2006b, Horn et al., 2006c, Sikorski et al., 2006, Synstad et al., 2008, Vaaje-Kolstad et al., 2013, Mekasha et al., 2017) does not reduce the validity of that work.

Collectively, the data presented in this thesis provide new knowledge about enzymes involved in chitin utilization and modification. Natural chitinolytic machineries have been mapped and novel enzymes have been characterized. More work is needed to understand chitin degradation in Nature and to identify more natural tools with industrial potential. For example, the novel target proteins coming out of the omics work described in **Papers II-IV** still need to be assessed for functionality and potential industrial applicability. Further proteomics studies using natural chitin-rich biomasses like crab- or shrimp-shells may yield additional targets for further work. While the results presented in this thesis have been discussed primarily in an applied perspective, it is important to note that answering unresolved questions considering the chitinolytic machinery of microorganisms eventually may yield insights into important biological processes where chitinolytic enzymes perhaps play a role (Kawada et al., 2008, Chaudhuri et al., 2013, Frederiksen et al., 2013), including e.g. virulence and biofilm formation.

5 REFERENCES

- AAM, B. B., HEGGSET, E. B., NORBERG, A. L., SØRLIE, M., VÅRUM, K. M. & EIJSINK, V. G. H. 2010. Production of chitooligosaccharides and their potential applications in medicine. *Marine drugs*, 8, 1482-1517.
- ABBY, S. S., CURY, J., GUGLIELMINI, J., NÉRON, B., TOUCHON, M. & ROCHA, E. P. C. 2016. Identification of protein secretion systems in bacterial genomes. *Scientific Reports*, 6, 23080.
- ADAV, S. S., CHEOW, E. S. H., RAVINDRAN, A., DUTTA, B. & SZE, S. K. 2012. Label free quantitative proteomic analysis of secretome by *Thermobifida fusca* on different lignocellulosic biomass. *Journal of Proteomics*, 75, 3694-3706.
- ADESIOYE, F. A., MAKHALANYANE, T. P., BIELY, P. & COWAN, D. A. 2016. Phylogeny, classification and metagenomic bioprospecting of microbial acetyl xylan esterases. *Enzyme and Microbial Technology*, 93, 79-91.
- ADRANGI, S. & FARAMARZI, M. A. 2013. From bacteria to human: A journey into the world of chitinases. *Biotechnology Advances*, 31, 1786-1795.
- AGNIHOTRI, S. A., MALLIKARJUNA, N. N. & AMINABHAVI, T. M. 2004. Recent advances on chitosan-based micro- and nanoparticles in drug delivery. *Journal of Controlled Release*, 100, 5-28.
- AKAGI, K.-I., WATANABE, J., HARA, M., KEZUKA, Y., CHIKAISHI, E., YAMAGUCHI, T., AKUTSU, H., NONAKA, T., WATANABE, T. & IKEGAMI, T. 2006. Identification of the substrate interaction region of the chitin-binding domain of *Streptomyces griseus* chitinase C. *Journal of Biochemistry*, 139, 483-493.
- AKCAPINAR, G. B., KAPPEL, L., SEZERMAN, O. U. & SEIDL-SEIBOTH, V. 2015. Molecular diversity of LysM carbohydrate-binding motifs in fungi. *Current Genetics*, 61, 103-113.
- ALCAZAR-FUOLI, L., CLAVAUD, C., LAMARRE, C., AIMANIANDA, V., SEIDL-SEIBOTH, V., MELLADO, E. & LATGÉ, J.-P. 2011. Functional analysis of the fungal/plant class chitinase family in *Aspergillus fumigatus*. *Fungal Genetics and Biology*, 48, 418-429.
- ALLOCATI, N., MASULLI, M., DI ILIO, C. & DE LAURENZI, V. 2015. Die for the community: an overview of programmed cell death in bacteria. *Cell Death & Disease*, 6, e1609.
- ANDRÉS, E., ALBESA-JOVÉ, D., BIARNÉS, X., MOERSCHBACHER, B. M., GUERIN, M. E. & PLANAS, A. 2014. Structural basis of chitin oligosaccharide deacetylation. *Angewandte Chemie International Edition*, 53, 6882-6887.
- ANITHA, A., SOWMYA, S., KUMAR, P. T. S., DEEPTHI, S., CHENNAZHI, K. P., EHRlich, H., TSURKAN, M. & JAYAKUMAR, R. 2014. Chitin and chitosan in selected biomedical applications. *Progress in Polymer Science*, 39, 1644-1667.
- ARAKI, Y., FUKUOKA, S., OBA, S. & ITO, E. 1971. Enzymatic deacetylation of *N*-acetylglucosamine residues in peptidoglycan from *Bacillus cereus* cell walls. *Biochemical and Biophysical Research Communications*, 45, 751-758.
- ARAKI, Y. & ITO, E. 1975. A Pathway of Chitosan Formation in *Mucor rouxii*. *European Journal of Biochemistry*, 55, 71-78.
- ARANDA-MARTINEZ, A., NARANJO ORTIZ, M. Á., ABIHSSIRA GARCÍA, I. S., ZAVALA-GONZALEZ, E. A. & LOPEZ-LLORCA, L. V. 2017. Ethanol production from chitosan by the nematophagous fungus *Pochonia chlamydosporia* and the entomopathogenic fungi *Metarhizium anisopliae* and *Beauveria bassiana*. *Microbiological Research*, 204, 30-39.
- ARNTZEN, M. Ø., VÁRNAI, A., MACKIE, R. I., EIJSINK, V. G. H. & POPE, P. B. 2017. Outer membrane vesicles from *Fibrobacter succinogenes* S85 contain an array of carbohydrate-active enzymes with versatile polysaccharide-degrading capacity. *Environmental Microbiology*, 19, 2701-2714.
- BAGOS, P. G., NIKOLAOU, E. P., LIAKOPOULOS, T. D. & TSIRIGOS, K. D. 2010. Combined prediction of Tat and Sec signal peptides with hidden Markov models. *Bioinformatics*, 26, 2811-2817.

- BAI, Y., EIJSINK, V. G. H., KIELAK, A. M., VAN VEEN, J. A. & DE BOER, W. 2016. Genomic comparison of chitinolytic enzyme systems from terrestrial and aquatic bacteria. *Environmental Microbiology*, 18, 38-49.
- BAJAJ, M., FREIBERG, A., WINTER, J., XU, Y. & GALLERT, C. 2015. Pilot-scale chitin extraction from shrimp shell waste by deproteination and decalcification with bacterial enrichment cultures. *Applied Microbiology and Biotechnology*, 99, 9835-9846.
- BANTSCHIEFF, M., SCHIRLE, M., SWEETMAN, G., RICK, J. & KUSTER, B. 2007. Quantitative mass spectrometry in proteomics: a critical review. *Analytical and Bioanalytical Chemistry*, 389, 1017-1031.
- BARCHI, J. J. 2013. Mucin-type glycopeptide structure in solution: Past, present, and future. *Biopolymers*, 99, 713-723.
- BEESON, W. T., VU, V. V., SPAN, E. A., PHILLIPS, C. M. & MARLETTA, M. A. 2015. Cellulose degradation by polysaccharide monoxygenases. *Annual review of biochemistry*, 84, 923-946.
- BENDTSEN, J., KIEMER, L., FAUSBOLL, A. & BRUNAK, S. 2005a. Non-classical protein secretion in bacteria. *BMC Microbiology*, 5, 58.
- BENDTSEN, J., NIELSEN, H., VON HEIJNE, G. & BRUNAK, S. 2004. Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology*, 340, 783 - 795.
- BENDTSEN, J., NIELSEN, H., WIDDICK, D., PALMER, T. & BRUNAK, S. 2005b. Prediction of twin-arginine signal peptides. *BMC Bioinformatics*, 6, 167.
- BENGTSSON, O., ARNTZEN, M. Ø., MATHIESEN, G., SKAUGEN, M. & EIJSINK, V. G. H. 2016. A novel proteomics sample preparation method for secretome analysis of *Hypocrea jecorina* growing on insoluble substrates. *Journal of Proteomics*, 131, 104-112.
- BHOWMICK, R., GHOSAL, A., DAS, B., KOLEY, H., SAHA, D. R., GANGULY, S., NANDY, R. K., BHADRA, R. K. & CHATTERJEE, N. S. 2008. Intestinal adherence of *Vibrio cholerae* involves a coordinated interaction between colonization factor GbpA and mucin. *Infection and Immunity*, 76, 4968-4977.
- BIBI, R., AHMAD, Z., IMRAN, M., HUSSAIN, S., DITTA, A., MAHMOOD, S. & KHALID, A. 2016. Algal bioethanol production technology: A trend towards sustainable development. *Renewable and Sustainable Energy Reviews*, 71, 976-985.
- BIELY, P. 2012. Microbial carbohydrate esterases deacetylating plant polysaccharides. *Biotechnology Advances*, 30, 1575-1588.
- BIELY, P., CÔTÉ, G. L., KREMNIČKÝ, L., GREENE, R. V., DUPONT, C. & KLUEPFEL, D. 1996. Substrate specificity and mode of action of acetylxyylan esterase from *Streptomyces lividans*. *FEBS Letters*, 396, 257-260.
- BISSARO, B., RØHR, A. K., MÜLLER, G., CHYLENSKI, P., SKAUGEN, M., FORSBERG, Z., HORN, S. J., VAAJE-KOLSTAD, G. & EIJSINK, V. G. H. 2017. Oxidative cleavage of polysaccharides by monocopper enzymes depends on H₂O₂. *Nature Chemical Biology*, Epub ahead of print.
- BLACKWELL, J. 1969. Structure of β -chitin or parallel chain systems of poly- β -(1 \rightarrow 4)-*N*-acetyl-D-glucosamine. *Biopolymers*, 7, 281-298.
- BLAIR, D. E., HEKMAT, O., SCHÜTTELKOPF, A. W., SHRESTHA, B., TOKUYASU, K., WITHERS, S. G. & VAN AALTEN, D. M. F. 2006. Structure and mechanism of chitin deacetylase from the fungal pathogen *Colletotrichum lindemuthianum*. *Biochemistry*, 45, 9416-9426.
- BLAIR, D. E., SCHUTTELKOPF, A. W., MACRAE, J. I. & VAN AALTEN, D. M. F. 2005. Structure and metal-dependent mechanism of peptidoglycan deacetylase, a streptococcal virulence factor. *Proceedings of the National Academy of Sciences, USA*, 102, 15429-34.
- BLAIR, D. E. & VAN AALTEN, D. M. F. 2004. Structures of *Bacillus subtilis* PdaA, a family 4 carbohydrate esterase, and a complex with *N*-acetyl-glucosamine. *FEBS Letters*, 570, 13-9.
- BONECA, I. G. 2005. The role of peptidoglycan in pathogenesis. *Current Opinion in Microbiology*, 8, 46-53.

REFERENCES

- BOOT, R. G., BLOMMAART, E. F. C., SWART, E., GHAUHARALI-VAN DER VLUGT, K., BIJL, N., MOE, C., PLACE, A. & AERTS, J. M. F. G. 2001. Identification of a novel acidic mammalian chitinase distinct from chitotriosidase. *Journal of Biological Chemistry*, 276, 6770-6778.
- BORASTON, A. B., BOLAM, D. N., GILBERT, H. J. & DAVIES, G. J. 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochemical Journal*, 382, 769-781.
- BOWMAN, S. M. & FREE, S. J. 2006. The structure and synthesis of the fungal cell wall. *Bioessays*, 28, 799-808.
- BRAMUCCI, E., PAIARDINI, A., BOSSA, F. & PASCARELLA, S. 2012. PyMod: sequence similarity searches, multiple sequence-structure alignments, and homology modeling within PyMOL. *BMC Bioinformatics*, 13, S2.
- BROGLIE, K. & CHET, I. 1991. Transgenic plants with enhanced resistance to the fungal pathogen *Rhizoctonia solani*. *Science*, 254, 1194-1197.
- BROWN, L., WOLF, J. M., PRADOS-ROSALES, R. & CASADEVALL, A. 2015. Through the wall: extracellular vesicles in Gram-positive bacteria, mycobacteria and fungi. *Nature Reviews: Microbiology*, 13, 620-630.
- BRURBERG, M. B., EIJSINK, V. G. H., HAANDRIKMAN, A. J., VENEMA, G. & NES, I. F. 1995. Chitinase B from *Serratia marcescens* BJL200 is exported to the periplasm without processing. *Microbiology*, 141, 123-131.
- BRURBERG, M. B., EIJSINK, V. G. H. & NES, I. F. 1994. Characterization of a chitinase gene (*chiA*) from *Serratia marcescens* BJL200 and one-step purification of the gene product. *FEMS Microbiology Letters*, 124, 399-404.
- BUIST, G., STEEN, A., KOK, J. & KUIPERS, O. P. 2008. LysM, a widely distributed protein motif for binding to (peptido)glycans. *Molecular Microbiology*, 68, 838-847.
- CACCIA, D., DUGO, M., CALLARI, M. & BONGARZONE, I. 2013. Bioinformatics tools for secretome analysis. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1834, 2442-2453.
- CANTAREL, B. L., COUTINHO, P. M., RANCUREL, C., BERNARD, T., LOMBARD, V. & HENRISSAT, B. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research*, 37, D233-D238.
- CARLSTRÖM, D. 1957. The crystal structure of α -chitin (poly-*N*-acetyl-D-glucosamine). *The Journal of biophysical and biochemical cytology*, 3, 669-683.
- CAUFRIER, F., MARTINOU, A., DUPONT, C. & BOURIOTIS, V. 2003. Carbohydrate esterase family 4 enzymes: substrate specificity. *Carbohydrate Research*, 338, 687-692.
- CHAGNOT, C., ZORGANI, M. A., ASTRUC, T. & DESVAUX, M. 2013. Proteinaceous determinants of surface colonization in bacteria: bacterial adhesion and biofilm formation from a protein secretion perspective. *Frontiers in microbiology*, 4, 303.
- CHAUDHURI, S., BRUNO, J. C., ALONZO, F., XAYARATH, B., CIANCOTTO, N. P. & FREITAG, N. E. 2010. Contribution of chitinases to *Listeria monocytogenes* pathogenesis. *Applied and Environmental Microbiology*, 76, 7302-7305.
- CHAUDHURI, S., GANTNER, B. N., YE, R. D., CIANCOTTO, N. P. & FREITAG, N. E. 2013. The *Listeria monocytogenes* ChiA chitinase enhances virulence through suppression of host innate immunity. *mBio*, 4, e00617-12.
- CHEN, X., WEI, S., JI, Y., GUO, X. & YANG, F. 2015. Quantitative proteomics using SILAC: Principles, applications, and developments. *Proteomics*, 15, 3175-3192.
- CHENG, Q., LI, H., MERDEK, K. & PARK, J. T. 2000. Molecular characterization of the β -*N*-Acetylglucosaminidase of *Escherichia coli* and its role in cell wall recycling. *Journal of bacteriology*, 182, 4836-4840.
- CHO, Y.-W., CHO, Y.-N., CHUNG, S.-H., YOO, G. & KO, S.-W. 1999. Water-soluble chitin as a wound healing accelerator. *Biomaterials*, 20, 2139-2145.
- CHRISTODOULIDOU, A., BOURIOTIS, V. & THIREOS, G. 1996. Two sporulation-specific chitin deacetylase-encoding genes are required for the ascospore wall rigidity of *Saccharomyces cerevisiae*. *Journal of Biological Chemistry*, 271, 31420-31425.

- CHRISTODOULIDOU, A., BRIZA, P., ELLINGER, A. & BOURIOTIS, V. 1999. Yeast ascospore wall assembly requires two chitin deacetylase isozymes. *FEBS Letters*, 460, 275-279.
- COLLINGE, D. B., KRAGH, K. M., MIKKELSEN, J. D., NIELSEN, K. K., RASMUSSEN, U. & VAD, K. 1993. Plant chitinases. *The Plant Journal*, 3, 31-40.
- CORD-LANDWEHR, S., MELCHER, R. L. J., KOLKENBROCK, S. & MOERSCHBACHER, B. M. 2016. A chitin deacetylase from the endophytic fungus *Pestalotiopsis* sp. efficiently inactivates the elicitor activity of chitin oligomers in rice cells. *Scientific Reports*, 6, 38018.
- COSTA, T. R., FELISBERTO-RODRIGUES, C., MEIR, A., PREVOST, M. S., REDZEJ, A., TROKTER, M. & WAKSMAN, G. 2015. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nature Reviews Microbiology*, 13, 343-359.
- COX, J., HEIN, M. Y., LUBER, C. A., PARON, I., NAGARAJ, N. & MANN, M. 2014. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular and Cellular Proteomics*, 13, 2513-2526.
- COX, J. & MANN, M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, 26, 1367-1372.
- D'ELIA, J. N. & SALYERS, A. A. 1996. Effect of regulatory protein levels on utilization of starch by *Bacteroides thetaiotaomicron*. *Journal of bacteriology*, 178, 7180-7186.
- DAVIES, G. & HENRISSAT, B. 1995. Structures and mechanisms of glycosyl hydrolases. *Structure*, 3, 853-859.
- DE JONGE, R., PETER VAN ESSE, H., KOMBRINK, A., SHINYA, T., DESAKI, Y., BOURS, R., VAN DER KROL, S., SHIBUYA, N., JOOSTEN, M. H. A. J. & THOMMA, B. P. H. J. 2010. Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. *Science*, 329, 953-955.
- DEBOY, R. T., MONGODIN, E. F., FOUTS, D. E., TAILFORD, L. E., KHOURI, H., EMERSON, J. B., MOHAMOUD, Y., WATKINS, K., HENRISSAT, B., GILBERT, H. J. & NELSON, K. E. 2008. Insights into plant cell wall degradation from the genome sequence of the soil bacterium *Cellvibrio japonicus*. *Journal of Bacteriology*, 190, 5455-63.
- DESVAUX, M. & HÉBRAUD, M. 2006. The protein secretion systems in *Listeria*: inside out bacterial virulence. *FEMS Microbiology Reviews*, 30, 774-805.
- DESVAUX, M., HEBRAUD, M., TALON, R. & HENDERSON, I. R. 2009. Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue. *Trends in Microbiology*, 17, 139-145.
- DICKINSON, E. 2017. Biopolymer-based particles as stabilizing agents for emulsions and foams. *Food Hydrocolloids*, 68, 219-231.
- DOVE, A. 1999. Proteomics: translating genomics into products? *Nature Biotechnology*, 17, 233-236.
- DROUILLARD, S., ARMAND, S., DAVIES, J. G., VORGIAS, E. C. & HENRISSAT, B. 1997. *Serratia marcescens* chitobiase is a retaining glycosidase utilizing substrate acetamido group participation. *Biochemical Journal*, 328, 945-949.
- EL GUEDDARI, N. E., RAUCHHAUS, U., MOERSCHBACHER, B. M. & DEISING, H. B. 2002. Developmentally regulated conversion of surface-exposed chitin to chitosan in cell walls of plant pathogenic fungi. *New Phytologist*, 156, 103-112.
- EL HADRAMI, A., ADAM, L. R., EL HADRAMI, I. & DAAYF, F. 2010. Chitosan in plant protection. *Marine Drugs*, 8, 968-987.
- EMAMI, K., NAGY, T., FONTES, C. M. G. A., FERREIRA, L. M. A. & GILBERT, H. J. 2002. Evidence for Temporal Regulation of the Two *Pseudomonas cellulosa* Xylanases Belonging to Glycoside Hydrolase Family 11. *Journal of Bacteriology*, 184, 4124-4133.
- FAZZINI, R. A. B., LEVICAN, G. & PARADA, P. 2011. *Acidithiobacillus thiooxidans* secretome containing a newly described lipoprotein Licanantase enhances chalcocopyrite bioleaching rate. *Applied Microbiology and Biotechnology*, 89, 771-780.
- FORSBERG, Z., NELSON, C. E., DALHUS, B., MEKASHA, S., LOOSE, J. S. M., CROUCH, L. I., RØHR, Å. K., GARDNER, J. G., EIJSINK, V. G. H. & VAAJE-KOLSTAD, G. 2016. Structural and functional

REFERENCES

- analysis of a lytic polysaccharide monoxygenase important for efficient utilization of chitin in *Cellvibrio japonicus*. *Journal of Biological Chemistry*, 291, 7300-7312.
- FORSBERG, Z., VAAJE-KOLSTAD, G., WESTERENG, B., BUNÆS, A. C., STENSTRØM, Y., MACKENZIE, A., SØRLIE, M., HORN, S. J. & EIJSINK, V. G. H. 2011. Cleavage of cellulose by a CBM33 protein. *Protein Science*, 20, 1479-1483.
- FREDERIKSEN, R. F., PASPALIARI, D. K., LARSEN, T., STORGAARD, B. G., LARSEN, M. H., INGMER, H., PALCIC, M. M. & LEISNER, J. J. 2013. Bacterial chitinases and chitin-binding proteins as virulence factors. *Microbiology*, 159, 833-847.
- FREIER, T., MONTENEGRO, R., KOH, H. S. & SHOICHET, M. S. 2005. Chitin-based tubes for tissue engineering in the nervous system. *Biomaterials*, 26, 4624-4632.
- FUKAMIZO, T., OHKAWA, T., IKEDA, Y. & GOTO, S. 1994. Specificity of chitosanase from *Bacillus pumilus*. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 1205, 183-188.
- FUKUSHIMA, T., KITAJIMA, T. & SEKIGUCHI, J. 2005. A polysaccharide deacetylase homologue, PdaA, in *Bacillus subtilis* acts as an *N*-acetylmuramic acid deacetylase in vitro. *Journal of Bacteriology*, 187, 1287-1292.
- GARDNER, J. G., CROUCH, L., LABOUREL, A., FORSBERG, Z., BUKHMAN, Y. V., VAAJE-KOLSTAD, G., GILBERT, H. J. & KEATING, D. H. 2014. Systems biology defines the biological significance of redox-active proteins during cellulose degradation in an aerobic bacterium. *Molecular Microbiology*, 94, 1121-1133.
- GARDNER, J. G. & KEATING, D. H. 2010. Requirement of the type II secretion system for utilization of cellulosic substrates by *Cellvibrio japonicus*. *Applied and Environmental Microbiology*, 76, 5079-5087.
- GAUTIER, S., XHAUFLAIRE-UHODA, E., GONRY, P. & PIÉRARD, G. E. 2008. Chitin-glucan, a natural cell scaffold for skin moisturization and rejuvenation. *International Journal of Cosmetic Science*, 30, 459-469.
- GHINET, M. G., ROY, S., POULIN-LAPRADE, D., LACOMBE-HARVEY, M.-È., MOROSOLI, R. & BRZEZINSKI, R. 2010. Chitosanase from *Streptomyces coelicolor* A3 (2): biochemical properties and role in protection against antibacterial effect of chitosan. *Biochemistry and Cell Biology*, 88, 907-916.
- GLOSTER, T. M. & DAVIES, G. J. 2010. Glycosidase inhibition: assessing mimicry of the transition state. *Organic & biomolecular chemistry*, 8, 305-320.
- GRUBER, S. & SEIDL-SEIBOTH, V. 2012. Self versus non-self: fungal cell wall degradation in *Trichoderma*. *Microbiology*, 158, 26-34.
- GUILLÉN, D., SÁNCHEZ, S. & RODRÍGUEZ-SANOJA, R. 2010. Carbohydrate-binding domains: multiplicity of biological roles. *Applied Microbiology and Biotechnology*, 85, 1241-1249.
- HAMED, I., ÖZOGUL, F. & REGENSTEIN, J. M. 2016. Industrial applications of crustacean by-products (chitin, chitosan, and chitooligosaccharides): A review. *Trends in Food Science & Technology*, 48, 40-50.
- HAMER, S. N., CORD-LANDWEHR, S., BIARNES, X., PLANAS, A., WAEGEMAN, H., MOERSCHBACHER, B. M. & KOLKENBROCK, S. 2015. Enzymatic production of defined chitosan oligomers with a specific pattern of acetylation using a combination of chitin oligosaccharide deacetylases. *Scientific Reports*, 5, 8716.
- HAMILTON, J. J., MARLOW, V. L., OWEN, R. A., COSTA MDE, A., GUO, M., BUCHANAN, G., CHANDRA, G., TROST, M., COULTHURST, S. J., PALMER, T., STANLEY-WALL, N. R. & SARGENT, F. 2014. A holin and an endopeptidase are essential for chitinolytic protein secretion in *Serratia marcescens*. *Journal of Cell Biology*, 207, 615-26.
- HANAZONO, Y., TAKEDA, K., NIWA, S., HIBI, M., TAKAHASHI, N., KANAI, T., ATOMI, H. & MIKI, K. 2016. Crystal structures of chitin binding domains of chitinase from *Thermococcus kodakarensis* KOD1. *FEBS Letters*, 590, 298-304.

- HART, P. J., PFLUGER, H. D., MONZINGO, A. F., HOLLIS, T. & ROBERTUS, J. D. 1995. The refined crystal structure of an endochitinase from *Hordeum vulgare* L. seeds at 1.8 Å resolution. *Journal of molecular biology*, 248, 402-413.
- HEGGSET, E. B., DYBVIK, A. I., HOELL, I. A., NORBERG, A. L., SØRLIE, M., EIJSINK, V. G. H. & VÅRUM, K. M. 2010. Degradation of chitosans with a Family 46 chitosanase from *Streptomyces coelicolor* A3(2). *Biomacromolecules*, 11, 2487-2497.
- HEKMAT, O., TOKUYASU, K. & WITHERS, S. G. 2003. Subsite structure of the endo-type chitin deacetylase from a Deuteromycete, *Colletotrichum lindemuthianum*: an investigation using steady-state kinetic analysis and MS. *Biochemical Journal*, 374, 369-380.
- HEMSWORTH, G. R., DÉJEAN, G., DAVIES, G. J. & BRUMER, H. 2016. Learning from microbial strategies for polysaccharide degradation. *Biochemical Society Transactions*, 44, 94-108.
- HENRISSAT, B. & BAIROCH, A. 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal*, 293, 781-788.
- HERNÁNDEZ-LAUZARDO, A. N., BAUTISTA-BAÑOS, S., VELÁZQUEZ-DEL VALLE, M. G., MÉNDEZ-MONTEALVO, M. G., SÁNCHEZ-RIVERA, M. M. & BELLO-PÉREZ, L. A. 2008. Antifungal effects of chitosan with different molecular weights on in vitro development of *Rhizopus stolonifer* (Ehrenb.:Fr.) Vuill. *Carbohydrate Polymers*, 73, 541-547.
- HIRANO, K., WATANABE, M., SEKI, K., ANDO, A., SAITO, A. & MITSUTOMI, M. 2012. Classification of chitosanases by hydrolytic specificity toward N^1, N^4 -diacetylchitohexaose. *Bioscience, Biotechnology, and Biochemistry*, 76, 1932-1937.
- HIRANO, T., KADOKURA, K., IKEGAMI, T., SHIGETA, Y., KUMAKI, Y., HAKAMATA, W., OKU, T. & NISHIO, T. 2009. Heterodisaccharide 4-O-(*N*-acetyl- β -D-glucosaminyl)-D-glucosamine is a specific inducer of chitinolytic enzyme production in *Vibrios* harboring chitin oligosaccharide deacetylase genes. *Glycobiology*, 19, 1046-1053.
- HIRST, T. R., SANCHEZ, J., KAPER, J. B., HARDY, S. J. & HOLMGREN, J. 1984. Mechanism of toxin secretion by *Vibrio cholerae* investigated in strains harboring plasmids that encode heat-labile enterotoxins of *Escherichia coli*. *Proceedings of the National Academy of Sciences, USA*, 81, 7752-7756.
- HOELL, I. A., VAAJE-KOLSTAD, G. & EIJSINK, V. G. H. 2010. Structure and function of enzymes acting on chitin and chitosan. *Biotechnology and Genetic Engineering Reviews*, 27, 331-366.
- HOLLAK, C. E., VAN WEELY, S., VAN OERS, M. H. & AERTS, J. M. 1994. Marked elevation of plasma chitotriosidase activity. A novel hallmark of Gaucher disease. *Journal of Clinical Investigation*, 93, 1288-1292.
- HORN, S. J., SIKORSKI, P., CEDERKVIST, J. B., VAAJE-KOLSTAD, G., SØRLIE, M., SYNSTAD, B., VRIEND, G., VÅRUM, K. M. & EIJSINK, V. G. H. 2006a. Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides. *Proceedings of the National Academy of Sciences, USA*, 103, 18089-18094.
- HORN, S. J., SØRBOTTEN, A., SYNSTAD, B., SIKORSKI, P., SØRLIE, M., VÅRUM, K. M. & EIJSINK, V. G. H. 2006b. Endo/exo mechanism and processivity of family 18 chitinases produced by *Serratia marcescens*. *FEBS Journal*, 273, 491-503.
- HORN, S. J., SØRLIE, M., VAAJE-KOLSTAD, G., NORBERG, A. L., SYNSTAD, B., VÅRUM, K. M. & EIJSINK, V. G. H. 2006c. Comparative studies of chitinases A, B and C from *Serratia marcescens*. *Biocatalysis and Biotransformation*, 24, 39-53.
- HORN, S. J., SØRLIE, M., VÅRUM, K. M., VÄLJAMÄE, P. & EIJSINK, V. G. H. 2012a. Measuring Processivity. *Methods in Enzymology*, 510, 69-95.
- HORN, S. J., VAAJE-KOLSTAD, G., WESTERENG, B. & EIJSINK, V. G. H. 2012b. Novel enzymes for the degradation of cellulose. *Biotechnology for Biofuels*, 5, 45.
- HUANG, Z., HAO, Y., GAO, T., HUANG, Y., REN, S. & KEYHANI, N. O. 2016. The Ifchit1 chitinase gene acts as a critical virulence factor in the insect pathogenic fungus *Isaria fumosorosea*. *Applied Microbiology and Biotechnology*, 100, 5491-5503.

REFERENCES

- HULT, E.-L., KATOONO, F., UCHIYAMA, T., WATANABE, T. & SUGIYAMA, J. 2005. Molecular directionality in crystalline β -chitin: hydrolysis by chitinases A and B from *Serratia marcescens* 2170. *Biochemical Journal*, 388, 851-856.
- HUNT, D. E., GEVERS, D., VAHORA, N. M. & POLZ, M. F. 2008. Conservation of the chitin utilization pathway in the *Vibrionaceae*. *Applied and Environmental Microbiology*, 74, 44-51.
- IGARASHI, K., UCHIHASHI, T., UCHIYAMA, T., SUGIMOTO, H., WADA, M., SUZUKI, K., SAKUDA, S., ANDO, T., WATANABE, T. & SAMEJIMA, M. 2014. Two-way traffic of glycoside hydrolase family 18 processive chitinases on crystalline chitin. *Nature Communications*, 5, 3975.
- IZUME, M., NAGAE, S. I., KAWAGISHI, H., MITSUTOMI, M. & OHTAKARA, A. 1992. Action pattern of *Bacillus* sp. no. 7-M chitosanase on partially *N*-acetylated chitosan. *Bioscience, Biotechnology, and Biochemistry*, 56, 448-453.
- JANG, M. K., KONG, B. G., JEONG, Y. I., LEE, C. H. & NAH, J. W. 2004. Physicochemical characterization of α -chitin, β -chitin, and γ -chitin separated from natural resources. *Journal of Polymer Science Part A: Polymer Chemistry*, 42, 3423-3432.
- JEFFERY, C. J. 1999. Moonlighting proteins. *Trends in Biochemical Sciences*, 24, 8-11.
- JUNCKER, A. S., WILLENBROCK, H., VON HEIJNE, G., BRUNAK, S., NIELSEN, H. & KROGH, A. 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Science*, 12, 1652-62.
- KANDRA, P., CHALLA, M. M. & KALANGI PADMA JYOTHI, H. 2012. Efficient use of shrimp waste: present and future trends. *Applied Microbiology and Biotechnology*, 93, 17-29.
- KATOONO, F., TAGUCHI, M., SAKURAI, K., UCHIYAMA, T., NIKAIDOU, N., NONAKA, T., SUGIYAMA, J. & WATANABE, T. 2004. Importance of exposed aromatic residues in chitinase B from *Serratia marcescens* 2170 for crystalline chitin hydrolysis. *Journal of Biochemistry*, 136, 163-168.
- KAUR, S. & DHILLON, G. S. 2015. Recent trends in biological extraction of chitin from marine shell wastes: a review. *Critical Reviews in Biotechnology*, 35, 44-61.
- KAWADA, M., CHEN, C. C., ARIHIRO, A., NAGATANI, K., WATANABE, T. & MIZOGUCHI, E. 2008. Chitinase 3-like-1 enhances bacterial adhesion to colonic epithelial cells through the interaction with bacterial chitin-binding protein. *Lab Invest*, 88, 883-895.
- KHARADE, S. S. & MCBRIDE, M. J. 2014. *Flavobacterium johnsoniae* chitinase ChiA is required for chitin utilization and is secreted by the type IX secretion system. *Journal of Bacteriology*, 196, 961-970.
- KIM, H.-S., GOLYSHIN, P. N. & TIMMIS, K. N. 2007. Characterization and role of a metalloprotease induced by chitin in *Serratia* sp. KCK. *Journal of Industrial Microbiology & Biotechnology*, 34, 715-721.
- KIM, J. H., LEE, J., PARK, J. & GHO, Y. S. 2015. Gram-negative and Gram-positive bacterial extracellular vesicles. *Seminars in Cell & Developmental Biology*, 40, 97-104.
- KIM, S.-K. 2010. *Chitin, chitosan, oligosaccharides and their derivatives: biological activities and applications*, CRC Press.
- KOBAYASHI, K., SUDIARTA, I. P., KODAMA, T., FUKUSHIMA, T., ARA, K., OZAKI, K. & SEKIGUCHI, J. 2012. Identification and characterization of a novel polysaccharide deacetylase C (PdaC) from *Bacillus subtilis*. *Journal of Biological Chemistry*, 287, 9765-9776.
- KONG, M., CHEN, X. G., XING, K. & PARK, H. J. 2010. Antimicrobial properties of chitosan and mode of action: A state of the art review. *International Journal of Food Microbiology*, 144, 51-63.
- KROGH, A., LARSSON, B., VON HEIJNE, G. & SONNHAMMER, E. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology*, 305, 567 - 580.
- KURITA, K. 2006. Chitin and chitosan: Functional biopolymers from marine crustaceans. *Marine Biotechnology*, 8, 203-226.
- LANGNER, T. & GÖHRE, V. 2016. Fungal chitinases: function, regulation, and potential roles in plant/pathogen interactions. *Current Genetics*, 62, 243-254.

- LARSRINK, J., ZHU, Y., KHARADE, S. S., KWIATKOWSKI, K. J., EIJSINK, V. G. H., KOROPATKIN, N. M., MCBRIDE, M. J. & POPE, P. B. 2016. A polysaccharide utilization locus from *Flavobacterium johnsoniae* enables conversion of recalcitrant chitin. *Biotechnology for Biofuels*, 9, 260.
- LEE, E.-Y., CHOI, D.-S., KIM, K.-P. & GHO, Y. S. 2008. Proteomics in gram-negative bacterial outer membrane vesicles. *Mass Spectrometry Reviews*, 27, 535-555.
- LEVASSEUR, A., DRULA, E., LOMBARD, V., COUTINHO, P. M. & HENRISSAT, B. 2013. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnology for Biofuels*, 6, 41.
- LI, X., WANG, L.-X., WANG, X. & ROSEMAN, S. 2007. The chitin catabolic cascade in the marine bacterium *Vibrio cholerae*: characterization of a unique chitin oligosaccharide deacetylase. *Glycobiology*, 17, 1377-1387.
- LIANG, X., MOORE, R., WILTON, M., WONG, M. J. Q., LAM, L. & DONG, T. G. 2015. Identification of divergent type VI secretion effectors using a conserved chaperone domain. *Proceedings of the National Academy of Sciences, USA*, 112, 9106-9111.
- LIU, Z., GAY, L. M., TUVENG, T. R., AGGER, J. W., WESTERENG, B., MATHIESEN, G., HORN, S. J., VAAJE-KOLSTAD, G., VAN AALLEN, D. M. F. & EIJSINK, V. G. H. 2017. Structure and function of a broad-specificity chitin deacetylase from *Aspergillus nidulans* FGSC A4. *Scientific Reports*, 7, 1746.
- LOMBARD, V., GOLACONDA RAMULU, H., DRULA, E., COUTINHO, P. M. & HENRISSAT, B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research*, 42, D490-D495.
- LOOSE, J. S., FORSBERG, Z., FRAAIJE, M. W., EIJSINK, V. G. H. & VAAJE-KOLSTAD, G. 2014. A rapid quantitative activity assay shows that the *Vibrio cholerae* colonization factor GbpA is an active lytic polysaccharide monoxygenase. *FEBS Letters*, 588, 3435-3440.
- LOOSE, J. S. M., FORSBERG, Z., KRACHER, D., SCHEIBLBRANDNER, S., LUDWIG, R., EIJSINK, V. G. H. & VAAJE-KOLSTAD, G. 2016. Activation of bacterial lytic polysaccharide monoxygenases with cellobiose dehydrogenase. *Protein Science*, 25, 2175-2186.
- MADHUPRAKASH, J., SINGH, A., KUMAR, S., SINHA, M., KAUR, P., SHARMA, S., PODILE, A. R. & SINGH, T. P. 2013. Structure of chitinase D from *Serratia proteamaculans* reveals the structural basis of its dual action of hydrolysis and transglycosylation. *International Journal of Biochemistry and Molecular Biology*, 4, 166-78.
- MADHUPRAKASH, J., TANNEERU, K., KARLAPUDI, B., GURUPRASAD, L. & PODILE, A. R. 2014. Mutagenesis and molecular dynamics simulations revealed the chito oligosaccharide entry and exit points for chitinase D from *Serratia proteamaculans*. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1840, 2685-2694.
- MADHUPRAKASH, J., TANNEERU, K., PURUSHOTHAM, P., GURUPRASAD, L. & PODILE, A. R. 2012. Transglycosylation by chitinase D from *Serratia proteamaculans* improved through altered substrate interactions. *Journal of Biological Chemistry*, 287, 44619-44627.
- MARSH, K. & BUGUSU, B. 2007. Food packaging—Roles, materials, and environmental issues. *Journal of Food Science*, 72, R39-R55.
- MARTENS, E. C., KOROPATKIN, N. M., SMITH, T. J. & GORDON, J. I. 2009. Complex glycan catabolism by the Human gut microbiota: the Bacteroidetes Sus-like paradigm. *Journal of Biological Chemistry*, 284, 24673-24677.
- MCALISTER, G. C., HUTTLIN, E. L., HAAS, W., TING, L., JEDRYCHOWSKI, M. P., ROGERS, J. C., KUHN, K., PIKE, I., GROTHE, R. A. & BLETHROW, J. D. 2012. Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Analytical chemistry*, 84, 7469-7478.
- MCMAHON, K. J., CASTELLI, M. E., VESCOVI, E. G. & FELDMAN, M. F. 2012. Biogenesis of outer membrane vesicles in *Serratia marcescens* is thermoregulated and can be induced by activation of the Rcs phosphorelay system. *Journal of Bacteriology*, 194, 3241-3249.
- MEKASHA, S., BYMAN, I. R., LYNCH, C., TOUPALOVÁ, H., ANDĚRA, L., NÆS, T., VAAJE-KOLSTAD, G. & EIJSINK, V. G. H. 2017. Development of enzyme cocktails for complete saccharification of

- chitin using mono-component enzymes from *Serratia marcescens*. *Process Biochemistry*, 56, 132-138.
- MINAGAWA, T., OKAMURA, Y., SHIGEMASA, Y., MINAMI, S. & OKAMOTO, Y. 2007. Effects of molecular weight and deacetylation degree of chitin/chitosan on wound healing. *Carbohydrate Polymers*, 67, 640-644.
- MINE, S., NIYAMA, M., HASHIMOTO, W., IKEGAMI, T., KOMA, D., OHMOTO, T., FUKUDA, Y., INOUE, T., ABE, Y., UEDA, T., MORITA, J., UEGAKI, K. & NAKAMURA, T. 2014. Expression from engineered *Escherichia coli* chromosome and crystallographic study of archaeal *N,N'*-diacetylchitobiose deacetylase. *FEBS Journal*, 281, 2584-96.
- MITSUTOMI, M., UEDA, M., ARAI, M., ANDO, A. & WATANABE, T. 1996. Action patterns of microbial chitinases and chitosanases on partially *N*-acetylated chitosan. *Chitin Enzymology*, 2, 273-284.
- NAHNSEN, S., BIELOW, C., REINERT, K. & KOHLBACHER, O. 2013. Tools for label-free peptide quantification. *Molecular & Cellular Proteomics*, 12, 549-556.
- NAKAGAWA, Y. S., EIJSINK, V. G. H., TOTANI, K. & VAAJE-KOLSTAD, G. 2013. Conversion of α -chitin substrates with varying particle size and crystallinity reveals substrate preferences of the chitinases and lytic polysaccharide monoxygenase of *Serratia marcescens*. *Journal of Agricultural and Food Chemistry*, 61, 11061-11066.
- NAKAGAWA, Y. S., KUDO, M., LOOSE, J. S., ISHIKAWA, T., TOTANI, K., EIJSINK, V. G. H. & VAAJE-KOLSTAD, G. 2015. A small lytic polysaccharide monoxygenase from *Streptomyces griseus* targeting α - and β -chitin. *FEBS Journal*, 282, 1065-1079.
- NAM, K.-S., KIM, M.-K. & SHON, Y.-H. 2007. Inhibition of proinflammatory cytokine-induced invasiveness of HT-29 cells by chitosan oligosaccharide. *Journal of Microbiology and Biotechnology*, 17, 2042-2045.
- NAQVI, S., CORD-LANDWEHR, S., SINGH, R., BERNARD, F., KOLKENBROCK, S., EL GUEDDARI, N. E. & MOERSCHBACHER, B. M. 2016. A recombinant fungal chitin deacetylase produces fully defined chitosan oligomers with novel patterns of acetylation. *Applied and Environmental Microbiology*, 82, 6645-6655.
- NATALE, P., BRÜSER, T. & DRIESSEN, A. J. 2008. Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane—distinct translocases and mechanisms. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1778, 1735-1756.
- NAZARI, B., KOBAYASHI, M., SAITO, A., HASSANINASAB, A., MIYASHITA, K. & FUJII, T. 2013. Chitin-induced gene expression in secondary metabolic pathways of *Streptomyces coelicolor* A3(2) grown in soil. *Applied and Environmental Microbiology*, 79, 707-713.
- NGUYEN, K.-A., TRAVIS, J. & POTEMPA, J. 2007. Does the importance of the C-terminal residues in the maturation of RgpB from *Porphyromonas gingivalis* reveal a novel mechanism for protein export in a subgroup of Gram-negative bacteria? *Journal of Bacteriology*, 189, 833-843.
- NIKOLOV, S., FABRITIUS, H., PETROV, M., FRIÁK, M., LYMPERAKIS, L., SACHS, C., RAABE, D. & NEUGEBAUER, J. 2011. Robustness and optimal use of design principles of arthropod exoskeletons studied by ab initio-based multiscale simulations. *Journal of the Mechanical Behavior of Biomedical Materials*, 4, 129-145.
- ONG, S.-E. & MANN, M. 2005. Mass spectrometry-based proteomics turns quantitative. *Nature Chemical Biology*, 1, 252-262.
- ORIKOSHI, H., NAKAYAMA, S., MIYAMOTO, K., HANATO, C., YASUDA, M., INAMORI, Y. & TSUJIBO, H. 2005. Roles of four chitinases (ChiA, ChiB, ChiC, and ChiD) in the chitin degradation system of marine bacterium *Alteromonas* sp. strain O-7. *Applied and Environmental Microbiology*, 71, 1811-1815.
- PASPALIARI, D. K., LOOSE, J. S. M., LARSEN, M. H. & VAAJE-KOLSTAD, G. 2015. *Listeria monocytogenes* has a functional chitinolytic system and an active lytic polysaccharide monoxygenase. *FEBS Journal*, 282, 921-936.

- PAYNE, C. M., BABAN, J., HORN, S. J., BACKE, P. H., ARVAI, A. S., DALHUS, B., BJØRÅS, M., EIJSINK, V. G. H., SØRLIE, M., BECKHAM, G. T. & VAAJE-KOLSTAD, G. 2012. Hallmarks of processivity in glycoside hydrolases from crystallographic and computational studies of the *Serratia marcescens* chitinases. *The Journal of Biological Chemistry*, 287, 36322-36330.
- PERRAKIS, A., TEWS, I., DAUTER, Z., OPPENHEIM, A. B., CHET, I., WILSON, K. S. & VORGIAS, C. E. 1994. Crystal structure of a bacterial chitinase at 2.3 Å resolution. *Structure*, 2, 1169-1180.
- PETERSEN, T. N., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, 8, 785-786.
- PHILLIPS, C. M., BEESON IV, W. T., CATE, J. H. & MARLETTA, M. A. 2011. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chemical Biology*, 6, 1399-1406.
- POCHANAVANICH, P. & SUNTORNSUK, W. 2002. Fungal chitosan production and its characterization. *Letters in Applied Microbiology*, 35, 17-21.
- PUCHART, V., GARIÉPY, M.-C., SHARECK, F. & DUPONT, C. 2006. Identification of catalytically important amino acid residues of *Streptomyces lividans* acetylxyylan esterase A from carbohydrate esterase family 4. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1764, 263-274.
- PURUSHOTHAM, P. & PODILE, A. R. 2012. Synthesis of long-chain chitooligosaccharides by a hypertransglycosylating processive endochitinase of *Serratia proteamaculans* 568. *Journal of Bacteriology*, 194, 4260-4271.
- QUINLAN, R. J., SWEENEY, M. D., LEGGIO, L. L., OTTEN, H., POULSEN, J.-C. N., JOHANSEN, K. S., KROGH, K. B., JØRGENSEN, C. I., TOVBORG, M. & ANTHONSEN, A. 2011. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proceedings of the National Academy of Sciences, USA*, 108, 15079-15084.
- RAJAURE, M., BERRY, J., KONGARI, R., CAHILL, J. & YOUNG, R. 2015. Membrane fusion during phage lysis. *Proceedings of the National Academy of Sciences, USA*, 112, 5497-5502.
- RAVAL, N. P., SHAH, P. U., LADHA, D. G., WADHWANI, P. M. & SHAH, N. K. 2016. Comparative study of chitin and chitosan beads for the adsorption of hazardous anionic azo dye Congo Red from wastewater. *Desalination and Water Treatment*, 57, 9247-9262.
- RAWLINGS, N. D., BARRETT, A. J. & FINN, R. 2016. Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research*, 44, D343-D350.
- REEVES, A. R., WANG, G. & SALYERS, A. A. 1997. Characterization of four outer membrane proteins that play a role in utilization of starch by *Bacteroides thetaiotaomicron*. *Journal of bacteriology*, 179, 643-649.
- REICHOW, S. L., KOROTKOV, K. V., HOL, W. G. J. & GONEN, T. 2010. Structure of the cholera toxin secretion channel in its closed state. *Nature Structural & Molecular Biology*, 17, 1226-1232.
- RINAUDO, M. 2006. Chitin and chitosan: Properties and applications. *Progress in Polymer Science*, 31, 603-632.
- ROSE, R., BRÜSER, T., KISSINGER, J. C. & POHLSCHRÖDER, M. 2002. Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. *Molecular Microbiology*, 45, 943-950.
- RUDALL, K. M. 1963. The chitin/protein complexes of insect cuticles. *Advances in Insect Physiology*, Volume 1, 257-313.
- RYE, C. S. & WITHERS, S. G. 2000. Glycosidase mechanisms. *Current Opinion in Chemical Biology*, 4, 573-580.
- SAITO, A., OOYA, T., MIYATSUCHI, D., FUCHIGAMI, H., TERAKADO, K., NAKAYAMA, S.-Y., WATANABE, T., NAGATA, Y. & ANDO, A. 2009. Molecular characterization and antifungal activity of a family 46 chitosanase from *Amycolatopsis* sp. CsO-2. *FEMS Microbiology Letters*, 293, 79-84.
- SAITO, Y., KUMAGAI, H., WADA, M. & KUGA, S. 2002. Thermally reversible hydration of β -chitin. *Biomacromolecules*, 3, 407-410.

REFERENCES

- SAITO, Y., OKANO, T., GAILL, F., CHANZY, H. & PUTAUX, J.-L. 2000. Structural data on the intracrystalline swelling of β -chitin. *International Journal of Biological Macromolecules*, 28, 81-88.
- SALOMON, D., KINCH, L. N., TRUDGIAN, D. C., GUO, X., KLIMKO, J. A., GRISHIN, N. V., MIRZAEI, H. & ORTH, K. 2014. Marker for type VI secretion system effectors. *Proceedings of the National Academy of Sciences, USA*, 111, 9271-9276.
- SATO, K., NAITO, M., YUKITAKE, H., HIRAKAWA, H., SHOJI, M., MCBRIDE, M. J., RHODES, R. G. & NAKAYAMA, K. 2010. A protein secretion system linked to bacteroidete gliding motility and pathogenesis. *Proceedings of the National Academy of Sciences, USA*, 107, 276-281.
- SCHESSER, K., FRITZH-LINDSTEN, E. & WOLF-WATZ, H. 1996. Delineation and mutational analysis of the *Yersinia pseudotuberculosis* YopE domains which mediate translocation across bacterial and eukaryotic cellular membranes. *Journal of Bacteriology*, 178, 7227-7233.
- SCHNEEWIND, O. & MISSIAKAS, D. 2014. Sec-secretion and sortase-mediated anchoring of proteins in Gram-positive bacteria. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1843, 1687-1697.
- SCHRÖDINGER, L. L. C. 2015. The PyMOL Molecular Graphics System Version 1.8.
- SEIDL, V. 2008. Chitinases of filamentous fungi: a large group of diverse proteins with multiple physiological functions. *Fungal Biology Reviews*, 22, 36-42.
- SEIDL, V., HUEMER, B., SEIBOTH, B. & KUBICEK, C. P. 2005. A complete survey of *Trichoderma chitinases* reveals three distinct subgroups of family 18 chitinases. *FEBS Journal*, 272, 5923-5939.
- SEYDEL, A., GOUNON, P. & PUGSLEY, A. P. 1999. Testing the '+2 rule' for lipoprotein sorting in the *Escherichia coli* cell envelope with a new genetic selection. *Molecular Microbiology*, 34, 810-821.
- SIKORSKI, P., SØRBOTTEN, A., HORN, S. J., EIJNSINK, V. G. H. & VÅRUM, K. M. 2006. *Serratia marcescens* chitinases with tunnel-shaped substrate-binding grooves show endo activity and different degrees of processivity during enzymatic hydrolysis of chitosan. *Biochemistry*, 45, 9566-9574.
- SILJAMÄKI, P., VARMANEN, P., KANKAINEN, M., SUKURA, A., SAVIJOKI, K. & NYMAN, T. A. 2014. Comparative exoprotein profiling of different *Staphylococcus epidermidis* strains reveals potential link between nonclassical protein export and virulence. *Journal of Proteome Research*, 13, 3249-3261.
- SLÁMOVÁ, K., BOJAROVÁ, P., PETRÁSKOVÁ, L. & KŘEN, V. 2010. β -N-Acetylhexosaminidase: What's in a name...? *Biotechnology Advances*, 28, 682-693.
- SORY, M.-P., BOLAND, A., LAMBERMONT, I. & CORNELIS, G. R. 1995. Identification of the YopE and YopH domains required for secretion and internalization into the cytosol of macrophages, using the *cyaA* gene fusion approach. *Proceedings of the National Academy of Sciences, USA*, 92, 11998-12002.
- ST LEGER, R. J., COOPER, R. M. & CHARNLEY, A. K. 1986. Cuticle-degrading enzymes of entomopathogenic fungi: regulation of production of chitinolytic enzymes. *Microbiology*, 132, 1509-1517.
- SUZUKI, K., SUZUKI, M., TAIYOJI, M., NIKAIKIDOU, N. & WATANABE, T. 1998. Chitin binding protein (CBP21) in the culture supernatant of *Serratia marcescens* 2170. *Bioscience, Biotechnology, and Biochemistry*, 62, 128-135.
- SUZUKI, K., TAIYOJI, M., SUGAWARA, N., NIKAIKIDOU, N., HENRISSAT, B. & WATANABE, T. 1999. The third chitinase gene (*chiC*) of *Serratia marcescens* 2170 and the relationship of its product to other bacterial chitinases. *Biochemical Journal*, 343, 587.
- SUZUKI, K., UCHIYAMA, T., SUZUKI, M., NIKAIKIDOU, N., REGUE, M. & WATANABE, T. 2001. LysR-type transcriptional regulator ChiR is essential for production of all chitinases and a chitin-binding protein, CBP21, in *Serratia marcescens* 2170. *Bioscience, Biotechnology, and Biochemistry*, 65, 338-347.

- SYNSTAD, B., VAAJE-KOLSTAD, G., CEDERKVIST, F. H., SAUA, S. F., HORN, S. J., EIJSINK, V. G. H. & SØRLIE, M. 2008. Expression and Characterization of Endochitinase C from *Serratia marcescens* BJL200 and Its Purification by a One-Step General Chitinase Purification Method. *Bioscience, Biotechnology, and Biochemistry*, 72, 715-723.
- SØRBOTTEN, A., HORN, S. J., EIJSINK, V. G. H. & VÅRUM, K. M. 2005. Degradation of chitosans with chitinase B from *Serratia marcescens*. *FEBS Journal*, 272, 538-549.
- TAKASUKA, T. E., BOOK, A. J., LEWIN, G. R., CURRIE, C. R. & FOX, B. G. 2013. Aerobic deconstruction of cellulosic biomass by an insect-associated *Streptomyces*. *Scientific Reports*, 3, 1030.
- TANAKA, T., FUJIWARA, S., NISHIKORI, S., FUKUI, T., TAKAGI, M. & IMANAKA, T. 1999. A unique chitinase with dual active sites and triple substrate binding sites from the hyperthermophilic archaeon *Pyrococcus kodakaraensis* KOD1. *Applied and Environmental Microbiology*, 65, 5338-5344.
- TANAKA, T., FUKUI, T., ATOMI, H. & IMANAKA, T. 2003. Characterization of an exo- β -D-glucosaminidase involved in a novel chitinolytic pathway from the hyperthermophilic Archaeon *Thermococcus kodakaraensis* KOD1. *Journal of Bacteriology*, 185, 5175-5181.
- TANAKA, T., FUKUI, T., FUJIWARA, S., ATOMI, H. & IMANAKA, T. 2004. Concerted action of diacetylchitobiose deacetylase and exo-beta-D-glucosaminidase in a novel chitinolytic pathway in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *Journal of Biological Chemistry*, 279, 30021-30027.
- TANAKA, T., FUKUI, T. & IMANAKA, T. 2001. Different cleavage specificities of the dual catalytic domains in chitinase from the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *Journal of Biological Chemistry*, 276, 35629-35635.
- TANG, M.-C., NISOLE, A., DUPONT, C., PELLETIER, J. N. & WALDRON, K. C. 2011. Chemical profiling of the deacetylase activity of acetyl xylan esterase A (AxeA) variants on chitooligosaccharides using hydrophilic interaction chromatography–mass spectrometry. *Journal of Biotechnology*, 155, 257-265.
- TAYLOR, E. J., GLOSTER, T. M., TURKENBURG, J. P., VINCENT, F., BRZOSOWSKI, A. M., DUPONT, C., SHARECK, F., CENTENO, M. S., PRATES, J. A., PUCHART, V., FERREIRA, L. M., FONTES, C. M., BIELY, P. & DAVIES, G. J. 2006. Structure and activity of two metal ion-dependent acetylxyylan esterases involved in plant cell wall degradation reveals a close similarity to peptidoglycan deacetylases. *Journal of Biological Chemistry*, 281, 10968-10975.
- TERRAPON, N., LOMBARD, V., GILBERT, H. J. & HENRISSAT, B. 2015. Automatic prediction of polysaccharide utilization loci in Bacteroidetes species. *Bioinformatics*, 31, 647-655.
- TEWS, I., TERWISSCHA VAN SCHELTINGA, A. C., PERRAKIS, A., WILSON, K. S. & DIJKSTRA, B. W. 1997. Substrate-assisted catalysis unifies two families of chitinolytic enzymes. *Journal of the American Chemical Society*, 119, 7954-7959.
- TEWS, I., VINCENTELLI, R. & VORGIAS, C. E. 1996. N-Acetylglucosaminidase (chitobiase) from *Serratia marcescens*: gene sequence, and protein production and purification in *Escherichia coli*. *Gene*, 170, 63-67.
- THADATHIL, N. & VELAPPAN, S. P. 2014. Recent developments in chitosanase research and its biotechnological applications: A review. *Food Chemistry*, 150, 392-399.
- THOMAS, F., HEHEMANN, J.-H., REBUFFET, E., CZIZEK, M. & MICHEL, G. 2011. Environmental and gut *Bacteroidetes*: The food connection. *Frontiers in Microbiology*, 2.
- TJALSMA, H., ANTELMANN, H., JONGBLOED, J. D., BRAUN, P. G., DARMON, E., DORENBOS, R., DUBOIS, J. Y., WESTERS, H., ZANEN, G., QUAX, W. J., KUIPERS, O. P., BRON, S., HECKER, M. & VAN DIJL, J. M. 2004. Proteomics of protein secretion by *Bacillus subtilis*: separating the "secrets" of the secretome. *Microbiology and Molecular Biology Reviews*, 68, 207-233.
- TORATANI, T., SHOJI, T., IKEHARA, T., SUZUKI, K. & WATANABE, T. 2008. The importance of chitobiase and N-acetylglucosamine (GlcNAc) uptake in N, N'-diacetylchitobiose [(GlcNAc)₂] utilization by *Serratia marcescens* 2170. *Microbiology*, 154, 1326-1332.
- TSCHUMI, A., GRAU, T., ALBRECHT, D., REZWAN, M., ANTELMANN, H. & SANDER, P. 2012. Functional analyses of mycobacterial lipoprotein diacylglyceryl transferase and

REFERENCES

- comparative secretome analysis of a mycobacterial lgt mutant. *Journal of Bacteriology*, 194, 3938-3949.
- TSIGOS, I., MARTINO, A., KAFETZOPOULOS, D. & BOURIOTIS, V. 2000. Chitin deacetylases: new, versatile tools in biotechnology. *Trends in biotechnology*, 18, 305-312.
- UCHIYAMA, T., KATOONO, F., NIKAIKIDOU, N., NONAKA, T., SUGIYAMA, J. & WATANABE, T. 2001. Roles of the exposed aromatic residues in crystalline chitin hydrolysis by chitinase A from *Serratia marcescens* 2170. *Journal of Biological Chemistry*, 276, 41343-41349.
- UEDA, K., ISHIKAWA, S., ITAMI, T. & ASAI, T. 1952. Studies on the aerobic mesophilic cellulose-decomposing bacteria. 5-2. Taxonomical study on genus *Pseudomonas*. *Journal of the Agricultural Chemical Society of Japan*, 26, 35-41.
- VAAJE-KOLSTAD, G., FORSBERG, Z., LOOSE, J. S. M., BISSARO, B. & EIJSINK, V. G. H. 2017. Structural diversity of lytic polysaccharide monoxygenases. *Current Opinion in Structural Biology*, 44, 67-76.
- VAAJE-KOLSTAD, G., HORN, S. J., SØRLIE, M. & EIJSINK, V. G. H. 2013. The chitinolytic machinery of *Serratia marcescens*-a model system for enzymatic degradation of recalcitrant polysaccharides. *FEBS Journal*, 280, 3028-3049.
- VAAJE-KOLSTAD, G., HORN, S. J., VAN AALTEN, D. M. F., SYNSTAD, B. & EIJSINK, V. G. H. 2005. The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation. *Journal of Biological Chemistry*, 280, 28492-28497.
- VAAJE-KOLSTAD, G., WESTERENG, B., HORN, S. J., LIU, Z., ZHAI, H., SØRLIE, M. & EIJSINK, V. G. H. 2010. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*, 330, 219-222.
- VAIKUNTAPU, P. R., RAMBABU, S., MADHUPRAKASH, J. & PODILE, A. R. 2016. A new chitinase-D from a plant growth promoting *Serratia marcescens* GPS5 for enzymatic conversion of chitin. *Bioresource Technology*, 220, 200-207.
- VAN AALTEN, D. M. F., KOMANDER, D., SYNSTAD, B., GÅSEIDNES, S., PETER, M. G. & EIJSINK, V. G. H. 2001. Structural insights into the catalytic mechanism of a family 18 exo-chitinase. *Proceedings of the National Academy of Sciences, USA*, 98, 8979-8984.
- VAN AALTEN, D. M. F., SYNSTAD, B., BRURBERG, M. B., HOUGH, E., RIISE, B. W., EIJSINK, V. G. H. & WIERENGA, R. K. 2000. Structure of a two-domain chitotriosidase from *Serratia marcescens* at 1.9-Å resolution. *Proceedings of the National Academy of Sciences, USA*, 97, 5842-5847.
- VAN DEN BURG, H. A., HARRISON, S. J., JOOSTEN, M. H., VERVOORT, J. & DE WIT, P. J. 2006. *Cladosporium fulvum* Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection. *Molecular Plant-Microbe Interactions*, 19, 1420-1430.
- VAN EIJK, M., VAN ROOMEN, C. P., RENKEMA, G. H., BUSSINK, A. P., ANDREWS, L., BLOMMAART, E. F., SUGAR, A., VERHOEVEN, A. J., BOOT, R. G. & AERTS, J. M. 2005. Characterization of human phagocyte-derived chitotriosidase, a component of innate immunity. *International Immunology*, 17, 1505-1512.
- VAN ESSE, H. P., BOLTON, M. D., STERGIPOULOS, I., DE WIT, P. J. & THOMMA, B. P. 2007. The chitin-binding *Cladosporium fulvum* effector protein Avr4 is a virulence factor. *Molecular Plant-Microbe Interactions*, 20, 1092-1101.
- VAN SCHELTINGA, A. C. T., KALK, K. H., BEINTEMA, J. J. & DIJKSTRA, B. W. 1994. Crystal structures of hevamine, a plant defence protein with chitinase and lysozyme activity, and its complex with an inhibitor. *Structure*, 2, 1181-1189.
- VÁRNAI, A., SIIKA-AHO, M. & VIKARI, L. 2013. Carbohydrate-binding modules (CBMs) revisited: reduced amount of water counterbalances the need for CBMs. *Biotechnology for Biofuels*, 6, 30.
- VERGUNST, A. C., VAN LIER, M. C. M., DEN DULK-RAS, A., GROSSE STÜVE, T. A., OUWEHAND, A. & HOOYKAAS, P. J. J. 2005. Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of *Agrobacterium*. *Proceedings of the National Academy of Sciences, USA*, 102, 832-837.

- VIEGAS, A., BRÁS, N. F., CERQUEIRA, N. M., FERNANDES, P. A., PRATES, J. A., FONTES, C. M., BRUIX, M., ROMAO, M. J., CARVALHO, A. L. & RAMOS, M. J. 2008. Molecular determinants of ligand specificity in family 11 carbohydrate binding modules—an NMR, X-ray crystallography and computational chemistry approach. *FEBS Journal*, 275, 2524-2535.
- VIENS, P., DUBEAU, M.-P., KIMURA, A., DESAKI, Y., SHINYA, T., SHIBUYA, N., SAITO, A. & BRZEZINSKI, R. 2015. Uptake of chitosan-derived D-glucosamine oligosaccharides in *Streptomyces coelicolor* A3(2). *FEMS Microbiology Letters*, 362, fnv048.
- VINCENT, F., MOLIN, D. D., WEINER, R. M., BOURNE, Y. & HENRISSAT, B. 2010. Structure of a polyisoprenoid binding domain from *Saccharophagus degradans* implicated in plant cell wall breakdown. *FEBS Letters*, 584, 1577-1584.
- VOLLMER, W. & TOMASZ, A. 2000. The pgdA gene encodes for a peptidoglycan N-acetylglucosamine deacetylase in *Streptococcus pneumoniae*. *Journal of Biological Chemistry*, 275, 20496-20501.
- VOLLMER, W. & TOMASZ, A. 2002. Peptidoglycan N-acetylglucosamine deacetylase, a putative virulence factor in *Streptococcus pneumoniae*. *Infection and Immunity*, 70, 7176-7178.
- VON HEIJNE, G. 1990. The signal peptide. *The Journal of Membrane Biology*, 115, 195-201.
- VÅRUM, K. M., ANTHONSEN, M. W., GRASDALEN, H. & SMIDSRØD, O. 1991a. ¹³C-Nmr studies of the acetylation sequences in partially N-deacetylated chitins (chitosans). *Carbohydrate Research*, 217, 19-27.
- VÅRUM, K. M., ANTOHONSEN, M. W., GRASDALEN, H. & SMIDSRØD, O. 1991b. Determination of the degree of N-acetylation and the distribution of N-acetyl groups in partially N-deacetylated chitins (chitosans) by high-field n.m.r. spectroscopy. *Carbohydrate Research*, 211, 17-23.
- WATANABE, T., KIMURA, K., SUMIYA, T., NIKAIDOU, N., SUZUKI, K., SUZUKI, M., TAIYOJI, M., FERRER, S. & REGUE, M. 1997. Genetic analysis of the chitinase system of *Serratia marcescens* 2170. *Journal of Bacteriology*, 179, 7111-7117.
- WAWRZKIEWICZ, M., BARTCZAK, P. & JESIONOWSKI, T. 2017. Enhanced removal of hazardous dye form aqueous solutions and real textile wastewater using bifunctional chitin/lignin biosorbent. *International Journal of Biological Macromolecules*, 99, 754-764.
- WESTERENG, B., ISHIDA, T., VAAJE-KOLSTAD, G., WU, M., EJSINK, V. G. H., IGARASHI, K., SAMEJIMA, M., STÅHLBERG, J., HORN, S. J. & SANDGREN, M. 2011. The putative endoglucanase PcGH61D from *Phanerochaete chrysosporium* is a metal-dependent oxidative enzyme that cleaves cellulose. *PLoS one*, 6, e27807.
- WILLIAMS, N. 2008. Biofuel debate deepens. *Current Biology*, 18, R891-R892.
- WILLIAMS, S. J. & WITHERS, S. G. 2000. Glycosyl fluorides in enzymatic reactions. *Carbohydrate Research*, 327, 27-46.
- WILSON, M. M. & BERNSTEIN, H. D. 2016. Surface-exposed lipoproteins: An emerging secretion phenomenon in Gram-negative bacteria. *Trends in Microbiology*, 24, 198-208.
- WINKLER, A., DOMINGUEZ-NUÑEZ, J., ARANAZ, I., POZA-CARRIÓN, C., RAMONELL, K., SOMERVILLE, S. & BERROCAL-LOBO, M. 2017. Short-chain chitin oligomers: Promoters of plant growth. *Marine Drugs*, 15, 40.
- XIAO, G., YING, S.-H., ZHENG, P., WANG, Z.-L., ZHANG, S., XIE, X.-Q., SHANG, Y., ST. LEGER, R. J., ZHAO, G.-P., WANG, C. & FENG, M.-G. 2012. Genomic perspectives on the evolution of fungal entomopathogenicity in *Beauveria bassiana*. *Scientific Reports*, 2, 483.
- YAMAGUCHI, K., YU, F. & INOUE, M. 1988. A single amino acid determinant of the membrane localization of lipoproteins in *E. coli*. *Cell*, 53, 423-432.
- YE, W., LEUNG, M. F., XIN, J., KWONG, T. L., LEE, D. K. L. & LI, P. 2005. Novel core-shell particles with poly(n-butyl acrylate) cores and chitosan shells as an antibacterial coating for textiles. *Polymer*, 46, 10538-10543.
- YOUNES, I., HAJJI, S., FRACHET, V., RINAUDO, M., JELLOULI, K. & NASRI, M. 2014. Chitin extraction from shrimp shell using enzymatic treatment. Antitumor, antioxidant and antimicrobial activities of chitosan. *International Journal of Biological Macromolecules*, 69, 489-498.

REFERENCES

- YOUNES, I., HAJJI, S., RINAUDO, M., CHAABOUNI, M., JELLOULI, K. & NASRI, M. 2016. Optimization of proteins and minerals removal from shrimp shells to produce highly acetylated chitin. *International Journal of Biological Macromolecules*, 84, 246-253.
- YOUNES, I. & RINAUDO, M. 2015. Chitin and chitosan preparation from marine sources. Structure, properties and applications. *Marine Drugs*, 13, 1133-1174.
- ZAKARIASSEN, H., AAM, B. B., HORN, S. J., VÅRUM, K. M., SØRLIE, M. & EIJSINK, V. G. H. 2009. Aromatic residues in the catalytic center of chitinase A from *Serratia marcescens* affect processivity, enzyme activity, and biomass converting efficiency. *Journal of Biological Chemistry*, 284, 10610-10617.
- ZHAO, Y., PARK, R. D. & MUZZARELLI, R. A. 2010. Chitin deacetylases: properties and applications. *Marine Drugs*, 8, 24-46.
- ZHU, Z., ZHENG, T., HOMER, R. J., KIM, Y.-K., CHEN, N. Y., COHN, L., HAMID, Q. & ELIAS, J. A. 2004. Acidic mammalian chitinase in asthmatic Th2 inflammation and IL-13 pathway activation. *Science*, 304, 1678-1682.

Paper I

Structure and function of a CE4 deacetylase isolated from a marine environment.

Tuveng, T. R., Rothwiler, U., Udata, G., Vaaje-Kolstad, G., Smalås, A. & Eijsink, V. G. H. 2017. Submitted to PlosOne.

1 **Structure and function of a CE4 deacetylase isolated from a marine**
2 **environment**

3 Tina Rise Tuveng¹, Ulli Rothweiler², Gupta Udatha², Gustav Vaaje-Kolstad¹, Arne Smalås², Vincent G.
4 H. Eijsink¹

5 1. Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences
6 (NMBU), P.O. Box 5003, 1432 Ås, Norway

7 2. The Norwegian Structural Biology Centre, Department of Chemistry, The Arctic University of
8 Norway, N-9037 Tromsø, Norway

9

10 **Corresponding author:** Vincent Eijsink, vincent.eijsink@nmbu.no

11

12 **Keywords:** Chitin, deacetylase, carbohydrate esterase, acetylation, chitosan

13 **Abbreviations:** **AMAC**, 2-aminoacridone; **CE**, Carbohydrate esterase; **CDA**, Chitin deacetylase; **CHOS**,
14 chito-oligosaccharides; **GlcN** or **D**, glucosamine; **GlcNAc** or **A**, *N*-acetyl glucosamine

15 **Abstract**

16 Chitin, a polymer of $\beta(1-4)$ -linked *N*-acetylglucosamine found in e.g. arthropods, is a valuable
17 resource that may be used to produce chitosan and chitooligosaccharides, two compounds with
18 considerable industrial and biomedical potential. Deacetylating enzymes may be used to tailor the
19 properties of chitin and its derived products. Here, we describe a novel CE4 enzyme originating from
20 a marine *Arthrobacter* species (*ArCE4A*). Crystal structures of this novel deacetylase were
21 determined, with and without bound chitobiose [(GlcNAc)₂], and refined to 2.1 Å and 1.6 Å,
22 respectively. In-depth biochemical characterization showed that *ArCE4A* has broad substrate
23 specificity, with higher activity against longer oligosaccharides. Mass spectrometry-based
24 sequencing of reaction products generated from a fully acetylated pentamer showed that internal
25 sugars are more prone to deacetylation than the ends. These enzyme properties are discussed in the
26 light of the structure of the enzyme-ligand complex, which adds valuable information to our still
27 rather limited knowledge on enzyme-substrate interactions in the CE4 family.

28

29 1. Introduction

30 Today there is a focus on the shift from a fossil-based economy to a greener economy based on
31 renewable resources such as biomass. Chitin, an insoluble polymer of β -1,4 linked *N*-
32 acetylglucosamine (GlcNAc), is considered as the second most abundant biomass on earth, and
33 occurs in large amounts in different ecosystems, for example in the exoskeleton of crustaceans and
34 insects. Many microorganisms can utilize chitin as an energy source and exploration of
35 metagenomics information from chitin-rich ecosystems is thus likely to reveal enzymes with activity
36 against chitin.

37 The production of chitosan [partially deacetylated chitin consisting of GlcNAc and
38 glucosamine (GlcN)] and chitooligosaccharides (CHOS, i.e. homo- or hetero-oligosaccharides of GlcN
39 and GlcNAc) from chitin is of considerable industrial interest. However, the extraction of chitin from
40 e.g. shrimp shells and the subsequent production of chitosan and CHOS involves the use of harsh
41 chemicals that are not environmentally friendly [1, 2]. Therefore, it is desirable to replace one or
42 more of the chemical processing steps used today with enzymatic processes. The degree of
43 polymerization (DP) and the fraction of acetylation (F_A) are well known determinants of the
44 physicochemical and biological properties of chitosan and CHOS. In addition, the pattern of
45 acetylation (P_A) is believed to have impact on the properties of chitosan and CHOS [3]. The potential
46 applications of chitosan and CHOS are numerous (reviewed in e.g. [4] and [5]), which is in part due
47 to their biocompatibility.

48 Deacetylases acting on chitin (CDAs) occur in carbohydrate esterase family 4 (CE4) of the
49 CAZy database (www.cazy.org) [6]. CE4 enzymes are capable of removing acetyl groups in chitin,
50 chitosan, and CHOS, thus converting GlcNAc (or A) units to GlcN (or D) units. Enzymes in the CE4
51 family may also act on peptidoglycan [7, 8] and acetyl xylan [9]. The use of CDAs could in principle
52 allow tailoring of both the fraction and pattern of acetylation in chitosan and CHOS [10-12]. For
53 example, Hamer et al. used two different deacetylases (NodB from *Rhizobium* sp. GRH2 and VcCDA
54 from *Vibrio cholerae*) to produce CHOS containing two deacetylated sugars in their non-reducing

55 ends [12]. They could do so because NodB specifically deacetylates the non-reducing end, while
56 VcCDA specifically deacetylates the sugar next to the non-reducing end [12]. Notably, most
57 characterized CE4 deacetylases show a broader substrate specificity [13-15], deacetylating several
58 positions in CHOS, chitin, chitosan, and acetyl xylan.

59 Despite their abundance in Nature and a plethora of (potential) roles in biology and industry,
60 available structural information for CE4 enzymes remains limited, and information on enzyme-
61 substrate interactions is scarce. In 2014, Andrés et al. described structures of VcCDA in complex with
62 chitobiose and chitotriose. Based on this landmark study, these authors proposed that the pattern
63 of acetylation in the products of different CE4 enzymes is determined by variable loops near the
64 catalytic center that affect the accessibility of subsites in the binding cleft [16].

65 In an attempt to discover novel CDAs, we have searched a collection of bacterial genomes
66 and metagenomes for members of the CE4 family starting from existing annotations based on the
67 Enzyme Commission classification system [17]. Bioinformatic tools were utilized to select the most
68 promising candidates, resulting in one candidate for cloning, expression and in-depth
69 characterization. X-ray crystallography yielded two structures, one for the substrate free protein and
70 one for a complex with (GlcNAc)₂ bound in the active site. This novel CDA has an open active site (in
71 contrast to VcCDA) and the structure with substrate is the first structure of a complex for this type
72 of deacetylase. We also elucidated the substrate specificities of this deacetylase to gain insight into
73 its potential use for tailoring patterns of acetylation in CHOS.

74

75 **2. Materials and methods**

76 **2.1 Selection of candidates**

77 An internal collection of annotated bacterial genomes and metagenomes (~300 Mb of sequence
78 data), supplemented with metagenomics data from an Intestinal Microbiota Project [18] and from
79 the HOTS vertical ocean depth project, was searched for potential chitin deacetylases, i.e. enzymes
80 annotated with E.C. number 3.5.1.41. The resulting candidate proteins (64 in total) were subjected

81 to further bioinformatic investigations to select the most promising candidates, as described in the
82 Results and Discussion section.

83 **2.2 Cloning and protein production**

84 Synthetic gene encoding the selected protein (without signal peptide) with an N-terminal His6-Ala-
85 Gly-tag and sequence optimized for expression in E.coli, were ordered from GenScript (NJ, USA),
86 amplified by PCR and cloned into the pNIC-CH [19] vector utilizing Ligation Independent Cloning [20].
87 The synthetic gene encoded an N-terminal His-tag and contained its normal stop codon (meaning
88 the C-terminal His-tag encoded by this vector was not exploited). The plasmid containing the gene
89 of interest was transformed into chemically competent BL21 Star cells by heat shock. Transformants
90 were cultured in 2 ml LB medium supplemented with kanamycin (50 µg/ml) and a colony PCR type
91 of method was performed to check for correct plasmid size. Cultures for strains containing plasmids
92 with correct sizes were further cultivated by adding more LB medium and kanamycin, after which
93 plasmids were isolated using the plasmid purification kit from Macherey-Nagel GmbH & Co (Düren,
94 Germany), followed by sequencing of the inserted gene at GATC Biotech (Constance, Germany) using
95 Sanger sequencing.

96 Protein expression was started by growing a 5 ml pre-culture (LB with 50 µg/ml kanamycin,
97 overnight, 37 °C) which was used to inoculate 0.5 L TB-medium supplemented with kanamycin (50
98 µg/ml) and containing 0.011% Antifoam 204 (Sigma, Steinheim, Germany), followed by incubation
99 at 37 °C in a Harbinger system (LEX-48 Bioreactor, Harbinger biotech, Markham, Canada). At
100 $OD_{600}=0.6$, the culture was induced with IPTG (final concentration 0.2 mM) and incubation was
101 continued over night at 30 °C before harvesting the cells by centrifugation. The cell pellet was
102 resuspended in 20 ml 20 mM Tris-HCl, 150 mM NaCl, 10 mM imidazole, pH 8.0. Before sonication
103 (28% amplitude with a pulse of 5 seconds on, 10 seconds off for 10 minutes), DNaseI (final
104 concentration 1.4 µg/ml) and PMSF (final concentration 0.1 mM) were added. The sonicated sample
105 was centrifuged and the supernatant was filtered (0.45 µm), before protein purification by nickel
106 affinity chromatography using a HisTrap HP 5 ml column (GE Healthcare Life Sciences, Uppsala,

107 Sweden) connected to an Äkta pure system (GE Healthcare Life Sciences, Uppsala, Sweden). A
108 stepwise imidazole gradient ending at 500 mM imidazole was used to elute bound protein. After
109 checking the presence and purity of the protein by SDS-PAGE, relevant fractions were pooled and
110 the protein solution was concentrated, with concomitant buffer exchange to 20 mM Tris-HCl, 100
111 mM NaCl, pH 8.0, using Amicon Ultra-15 centrifugal filters with 10 000 NMWL (Merck Millipore, Cork,
112 Ireland). The protein concentration was measured with the Bradford micro assay (Bio-Rad, CA, USA).

113

114 **2.3 Structure determination**

115 The protein solution (10 mg/ml) was mixed (1:1) with the crystallization solution (100 mM MES pH6.5
116 15-18% PEG 3350) for a final drop size of 4 μ l. Crystallization was done in 24 well hanging drop plates.
117 Rod shaped crystals appeared within 1-2 days at room temperature. For the cocrystallization
118 experiments the protein solution (10 mg/ml) was treated with 1 mM EDTA (to prevent the
119 catalysis) prior to the addition of (GlcNAc)₄. Crystals were cryo-protected in the crystallization
120 solution modified to include 30% ethylene glycol and flash cooled in liquid nitrogen.

121 X-ray diffraction data were collected at the European Synchrotron Radiation Facility ESRF
122 Grenoble, France (collection statistics are summarized in Table 1). The images were integrated using
123 the XDS [21] and XDSapp [22] software . The structures were solved by molecular replacement with
124 Phaser [23] using the structure of SpPgdA, a peptidoglycan deacetylase from *Streptococcus*
125 *pneumoniae* (PDB id: 2C1G; [24]) as search model for 5LFZ and, subsequently, using 5LFZ as search
126 model for 5LGC. The structures were refined by iterative cycles of PHENIX [25] and the CCP4 program
127 REFMAC5 [26, 27] followed by the manual refitting of residues and ligands into the electron-density
128 between the refinement cycles and placement of water molecules using Coot v.0.7.2 [28]. PRODRG
129 [29] was used to generate the cif file for chitobiose.

130

131

132

133 **Table 1. Crystallographic data and model statistics for the two structures.**

Dataset PDB code	ArCE4A-Ni ²⁺ 5LFZ	ArCE4A-(GlcNAc) ₂ 5LCG
Data collection		
Source	ESRF ID29	ESRF ID29
Detector	Dectris Pilatis 6M	Dectris Pilatis 6M
Wavelength	0.97625	0.97625
No. of frames	1800	1350
Oscillation range per frame	0.1	0.1
Diffraction data		
Space group	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁
Unit cell parameters	a=39.09 Å b=56.77 Å c=76.86 Å	a=40.49 Å b=56.41 Å c=82.42 Å
No of measurements	155223	50017
Unique reflections	46454	11021
Resolution range	34.84 - 1.56	36.3-2.09
Completeness	99.7 (95.7)	93.9 (71.3)
Observed R factor (%)	4.1 (111)	5.2 (29.6)
Average I/sigma	12.82 (0.95)	17.6 (3.7)
Refinement		
Resolution limits	34.84 - 1.56	36.3-2.09
Free reflections	8.03%	5.15%
No. of protein atoms	1602	1558
No. of heterogen atoms	1	29
No. of waters	107	40
R factor (overall/free) (%)	0.178, 0.205	0.185, 0.223
Wilson B factor	24.6	33.3
R.m.s.d		
Bonds	0.005	0.006
Angles	0.957	0.930
Ramachandran		
Favored (%)	98	98
Allowed (%)	2	2
Outliers	0	0

134

135 **2.4 Activity assays**

136 Reaction mixtures for determination of enzyme activity contained 2 mM or 5 mg/ml substrate, 10
137 μM CoCl₂ and 300 nM enzyme in 50 mM Tris-HCl, pH 8.0. Reaction mixtures were incubated at 37 °C,
138 using a thermomixer with shaking at 600 rpm. Reactions were quenched by adding acetonitrile to a
139 final concentration of 50% (v/v). *N*-acetylglucosamine (GlcNAc) was purchased from Sigma- Aldrich
140 (Steinheim, Germany), while acetylated oligomers [(GlcNAc)₂₋₆] were purchased from MegaZyme
141 (Bray, Ireland). Alpha-chitin extracted from *Pandalus borealis* was from Seagarden (Avaldsnes,
142 Norway) and β-chitin extracted from squid pen was purchased from France Chitin (Batch 20140101,

143 Orange, France). Aspen acetyl xylan and chitosan ($F_A=0.64$) were a kind gifts from Bjørge Westereng
144 and BioCHOS AS (Ås, Norway), respectively. Quantification of released acetate was done by ion
145 chromatography using a Dionex ICS3000 system with suppressed conductivity detection and
146 equipped with a Dionex IonPac AS11 organic acid column, using the following gradient: 0 - 8 min, 1
147 mM KOH; 8 - 9 min, from 1 to 60 mM KOH; 9 – 16 min, 60 mM KOH; 16 – 16.1 min, from 60 to 1 mM
148 KOH; 16.1 – 22 min, 1 mM KOH. The flow rate was 0.375 ml/min. The amount of released acetate
149 was quantified using acetic acid [glacial, anhydrous (Merck, Damstadt, Germany)] as standard.
150 Operation of the Dionex ICS3000 system and processing of chromatograms were performed using
151 the Chromeleon 7 software (Dionex Corp.).

152

153 **2.5 AMAC labeling and sequencing of chito-oligomers**

154 Products generated by the deacetylase from (GlcNAc)₅ were labeled with 2-aminoacridone (AMAC)
155 (Sigma- Aldrich, Steinheim, Germany) as previously described by Bahrke et al. [30] and labeled
156 products were purified using a C18 column (Starata C18E, , Phenomenex, CA, US) as described by
157 Morelle et al. [31], with one deviation: instead of lyophilizing the labeled samples, the reaction
158 products were dried by vacuum centrifugation. The labeled products were re-dissolved in 50 µl 50 %
159 MeOH and analyzed using a LTQ-Velos Pro ion trap mass spectrometer (Thermo Scientific, Bremen,
160 Germany) connected to an Ultimate 3000 RS HPLC (Dionex, CA, USA). This setup was used for direct
161 injection without a column. The pump delivered 200 µl/min of 0.03 µM formic acid in 70%
162 acetonitrile and data was acquired for 24 seconds after injection. For the MS, the capillary voltage
163 was set to 3.5 kV and the scan range was m/z 150-2000 using two micro scans. The automatic gain
164 control was set to 10,000 charges and a maximum injection time of 20 milliseconds. For
165 fragmentation of desired precursor masses by MS2, the normalized collision energy was set to 37
166 and three micro scans were used. The data were recorded with Xcalibur version 2.2.

167

168 **3. Results and Discussion**

169 3.1 Selection of candidate CDAs from metagenome data

170 Deacetylases in CAZy family CE4 contain five conserved motifs containing residues that are important
171 for the catalytic activity [24, 32]: motif 1, T(F/Y)DD; motif 2, H(S/T)xxH; motif 3, R(P/x)PY; motif 4,
172 DxxD(W/Y); motif 5, LxH. The second aspartate in motif 1 coordinates a metal ion, preferably Co^{2+}
173 [24, 33], together with two histidines in motif 2. The first aspartate in motif 1 is believed to act as a
174 base during catalysis, activating a water molecule to carry out a nucleophilic attack on the carbon in
175 the scissile C-N bond. The histidine in motif 5, thought to be protonated, could promote C-N breaking
176 by acting as an acid protonating the leaving amino-sugar. The backbone of motif 3, in particular of
177 the tyrosine, is involved in stabilizing the oxyanion intermediate that is formed during catalysis [16,
178 24, 34]. Motif 3 and 4 each form one side of a shallow active site groove (Blair et al., 2005). Notably,
179 proteins may receive a CE4 annotation without possessing all these five motifs and such CE4 enzymes
180 are not likely to be active [35].

181 The initial search of the annotated bacterial genomes and metagenomes yielded 64 protein
182 sequences (annotated as EC 3.5.1.41), 48 of which belonged to CAZy family CE4. Each sequence was
183 manually inspected to check for the presence of all five sequence motifs, leaving 24 proteins. The
184 genes for 8 of these 24 proteins did not seem complete, leaving 16 candidates. Considering that
185 chitin would occur extracellularly, the next filter applied was the presence of a clear signal peptide,
186 as predicted by SignalP 4.1 [36]. This filtering step left 5 candidates. At this point, probable multi-
187 domain proteins (4 candidates) were excluded to increase the chances of successful expression. This
188 left one candidate protein, for which a structural model was built using Swiss-Model [37-39] to verify
189 for potential anomalies in or near the catalytic center. This novel CDA is the subject of the remaining
190 part of this report. It is interesting to note that, after using this rather straightforward approach, 48
191 CE4 sequences only yielded one candidate CDA. Obviously, the discarded CE4s, without signal
192 peptide and/or containing multi-domain proteins, could include active CDAs.

193 The selected CDA is 246 amino acids long, with a predicted signal peptide running from
194 amino acid number 1 to 31. The protein originates from the Gram-positive bacterium *Arthrobacter*

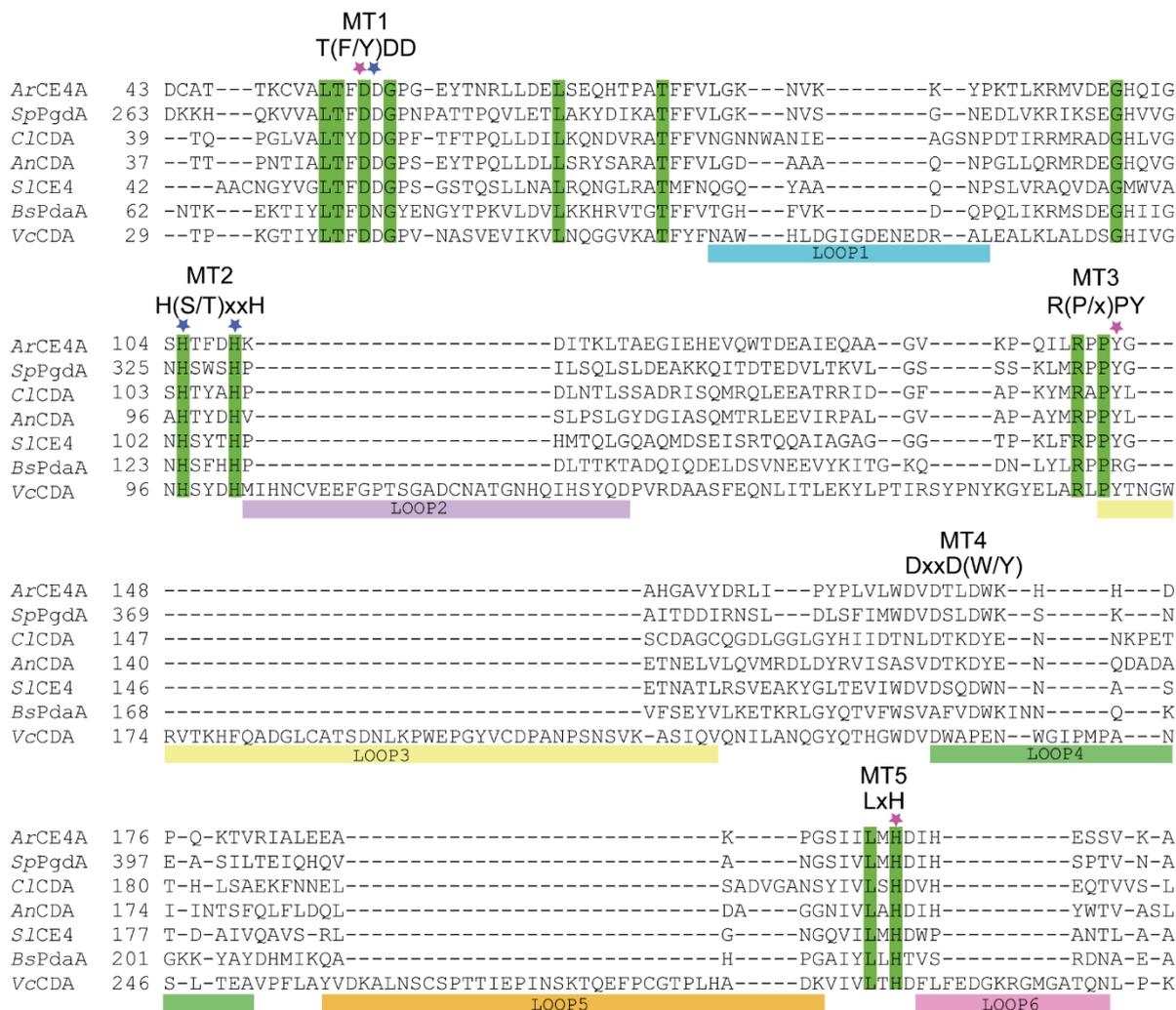
195 sp. AW19M34-1, which was isolated from a Tunicate located at 77 meters depth in Vestfjorden,
196 Norway. Tunicates secrete a chitinous peritrophic membrane [40, 41] and *Arthrobacter* species are
197 known for their ability to grow on chitin and for secretion of chitinases [42]. In line with commonly
198 used nomenclature for CAZymes the CDA was named *ArCE4A*. The gene sequence has been
199 deposited in the European Nucleotide Archive under Accession number LT630322
200 (<http://www.ebi.ac.uk/ena/data/view/LT630322>).

201

202 **3.2 Structure determination**

203 Two structures of *ArCE4A* were obtained by x-ray crystallography, one with (PDB id: 5LGC) and one
204 without (PDB id: 5LFZ) a (GlcNAc)₂ ligand, at 2.1 Å and 1.6 Å resolution, respectively (Table 1). The
205 protein has a somewhat deformed (β/α)₈ barrel topology (Fig. 1a) that is characteristic for CE4
206 proteins [15, 16, 24, 33, 34]. The structure of *ArCE4A* without (GlcNAc)₂ comprises residues 42-241,
207 meaning that no structural information was obtained for ten N-terminal residues (32-41) and five C-
208 terminal residues (242-246). Note that both the N- and the C-terminus are located on the opposite
209 side of the protein, relative to the catalytic center (Fig. 1a). The structure contains a Ni²⁺ ion
210 coordinated by Asp56, His105 and His109 (Fig. 1b), which comprise the metal binding triad that is
211 conserved in CE4 proteins. The Ni²⁺ most likely originates from the protein purification by nickel
212 affinity chromatography. The Ni²⁺ ion is in an octahedral arrangement, involving three water ligands
213 and the metal binding Asp-His-His triad. It has been proposed that one of these water molecules,
214 coordinated by Asp55, is the catalytic water acting as a nucleophile during catalysis [24].

215



216

217 **Fig. 1. Structure of ArCE4A determined by X-ray crystallography.** (a) Cartoon representation of the
 218 protein showing the disrupted (β/α)₈ barrel topology. The N- and C-terminus of the protein are
 219 marked and the metal ion in the active site is shown as a brown sphere, with the metal coordinating
 220 triad in sticks. (b) The His-His-Asp metal binding triad and the catalytic base (in sticks, PDB id: 5LFZ)
 221 with the Ni²⁺ ion as brown sphere. The Ni²⁺ ion shows octahedral coordination involving three amino
 222 acids and three water molecules (red spheres); interactions are shown as black dashed lines with
 223 distances in Å. The water molecule interacting with Asp55 is proposed to act as a nucleophile
 224 attacking the carbonyl carbon in the acetyl group. (c) Electron density map of the (GlcNAc)₂ ligand.
 225 This illustrates the lack of electron density for the remainder of the tetramer used in the co-
 226 crystallization. (d) ArCE4A in complex with (GlcNAc)₂ (PDB id: 5LCG) showing active site with the
 227 ligand bound in subsites 0 and +1 (grey carbons). Residues involved in substrate binding and catalysis
 228 are shown as sticks (purple carbons). Interactions between the protein and the substrate are shown
 229 as dashed lines in pink with distances in Å. (e) Superposition of the two structures (5LFZ in green,
 230 5LCG in blue) showing the active site cleft, and how the Ni²⁺ ion (brown sphere) and the three water
 231 molecules (red spheres) in 5LFZ are located relative to (GlcNAc)₂ in 5LCG. Interactions between the

232 Ni²⁺ ion and the substrate and ligand are shown as dashed black lines. Interactions between the
233 proposed nucleophilic water and Asp55 and the carbonyl carbon in the acetyl group are shown as
234 pink dashed lines. (f) Superposition of *AnCE4A* (purple carbons) and *C/CDA* (green carbons; PDB id:
235 2IW0 [34]), showing the extra loop containing Trp79 and nearby Phe53 (in sticks) in *C/CDA* in what
236 could be subsite -2. Subsites occupied by the ligand are labeled 0 and +1.

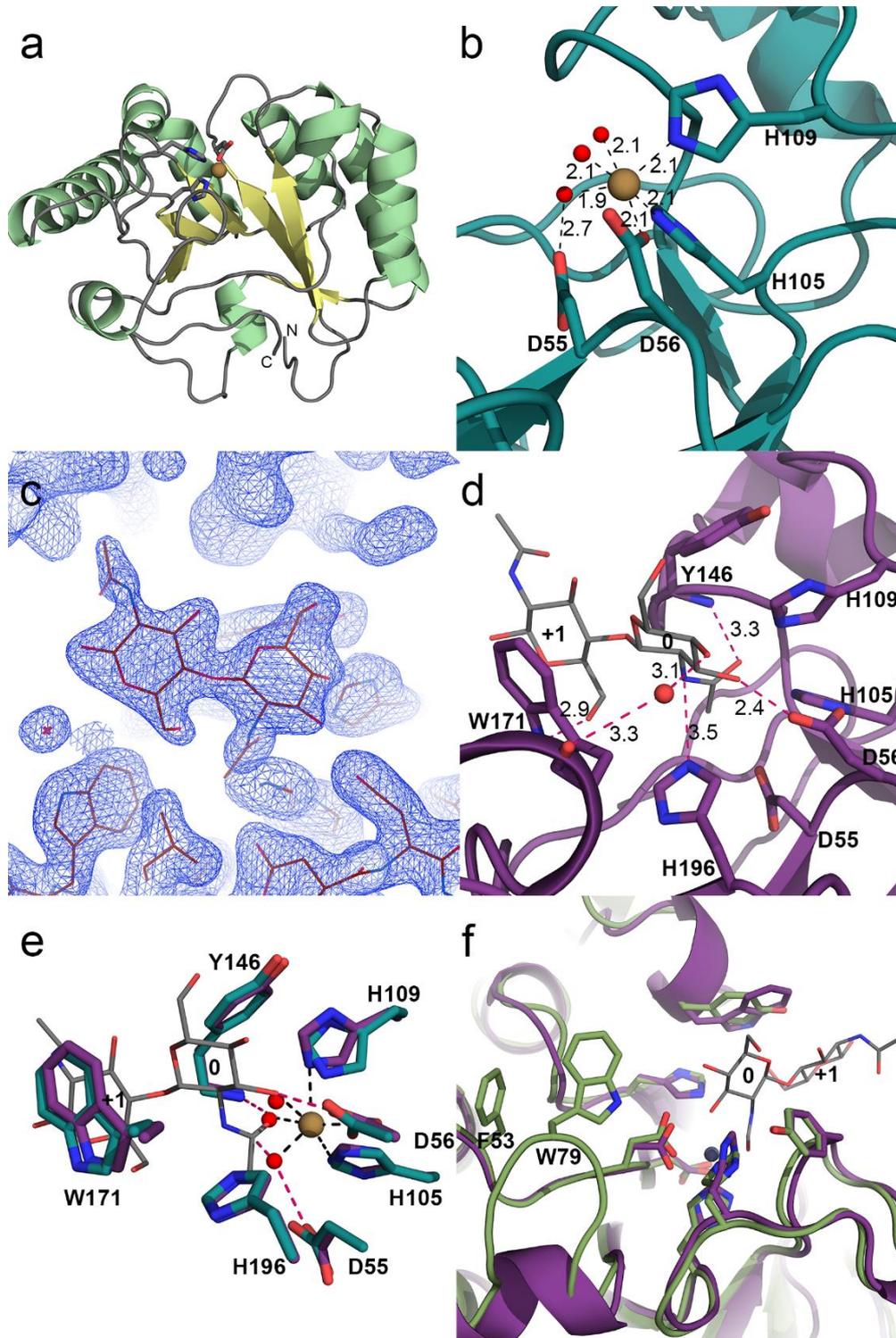
237

238 The structure with a bound ligand covers amino acids 41 to 239 and the ligand density (Fig.
239 1c) was refined as a GlcNAc dimer, occupying subsite 0 and +1 (Fig. 1d). From the four sugars of the
240 (GlcNAc)₄ that was used in the co-crystallization experiments, only two could be modeled into the
241 electron density. Apparently, the other two sugars are not stabilized by any protein-substrate
242 interactions and adopt multiple orientations/conformations that cannot be resolved in the electron
243 density map at this resolution. Fig. 1d shows that Trp171 in motif 4 stacks with the sugar bound in
244 subsite +1 forming one side of a shallow substrate-binding groove. Binding of the sugar in the +1
245 subsite seems to be dominated by this stacking interaction, whereas the acetyl group of this sugar is
246 not involved in interactions with the enzyme (Fig. 1d). The sugar bound in subsite 0 has multiple
247 interactions with the enzyme. The hydroxyl-group at C3 interacts with the metal ion and makes a
248 hydrogen bond with Asp56 (Fig. 1d and e), while the hydroxyl-group at C4 of the sugar bound in
249 subsite 0 seems to have an indirect interaction with the backbone carboxyl of Trp171 through a water
250 molecule (Fig. 1d). The backbone amide of Tyr146, thought to stabilize the oxyanion intermediate
251 by interacting with the oxygen atom of the acetyl group is located at 3.3 Å of this oxygen (Fig. 1d).
252 The Nε nitrogen of His196 in motif 5, thought to facilitate departure of the sugar, is located at 3.5 Å
253 from the nitrogen atom in the acetamido group (Fig. 1d), a distance not unlike the distances
254 proposed in previous docking studies (3.7 Å; [24, 34]). Asp55 in motif 1, expected to activate the
255 nucleophilic water is not making any direct interactions with the sugar in subsite 0. No water
256 molecules could be refined in the active site cleft in the structure with the (GlcNAc)₂ ligand, probably
257 due to the lack of a metal ion. Superposition of the two *ArCE4A* structures (Fig. 1e) reveals that the
258 water molecule coordinated by Asp55 in the substrate-free structure (Fig. 1b) indeed has a position

259 that could allow it acting as a nucleophile during catalysis. The other two water molecules, which
260 coordinate the metal ion in the substrate-free enzyme (Fig. 1b), occupy the same position as the
261 oxygens of the acetyl group and the hydroxyl on carbon 3 in of the sugar bound in subsite 0. It is
262 worth noting that the superposition (Fig. 1e) shows little difference in the conformation of the above-
263 mentioned amino acids.

264 A structure based sequence alignment with other known deacetylases (Fig. 2) shows that
265 there is high sequence similarity in the conserved motifs that are characteristic for deacetylases in
266 family CE4. However, there is some variation, which could correlate with differences in substrate
267 specificity, which are not all mapped yet, but are known to exist and be considerable. For example,
268 *BsPdaA* is an *N*-acetylmuramic acid deacetylase with no activity against CHOS [7], whereas *VcCDA*
269 only deacetylates CHOS on the sugar next to the non-reducing end. The structure of *VcCDA* so far
270 was the only available structure of a CE4 CDA in complex with its true substrate [16]. As shown in
271 Fig. 2, *VcCDA* is special in that it contains several long insertions, which are loops that cover the active
272 site and tailor this enzyme's ability to interact with its substrate [16]. *ArCE4A* and other CE4s proteins
273 acting on CHOS have active sites that are more open. Based on biochemical data, Hekmat et al. (2003)
274 proposed that *C/CDA*, having an open active site similar to *ArCE4A*, has four subsites, -2, -1, 0, and +1
275 [13]. The structure of *C/CDA* was solved by Blair et al. [34] and based on *in silico* docking of (GlcNAc)₃
276 they concluded that the sugar in subsite -1 has no interactions with the protein. Blair et al. further
277 pointed out that a tryptophan (Trp79) located in an insertion in loop 1 that is absent in *ArCE4A* (Fig.
278 1f and 2) could create a -2 subsite [34]. A phenylalanine (Phe53, Fig. 1f) located near the flexible loop
279 with Trp79 could possibly also be involved in substrate binding in subsite -2 of *C/CDA*. *ArCE4A* is more
280 open in the potential subsite -2 region (Fig. 1f) without any obvious residues to make interactions
281 with a bound sugar. Interestingly, while the protein was co-crystallized with (GlcNAc)₄ only two
282 GlcNAc units were observed. This suggests high flexibility of the rest of the ligand, which is in line
283 with the notion that *ArCE4A* has only two clear subsites, 0 and +1. Another noteworthy difference is
284 the tyrosine in *C/CDA* (Tyr173) in stead of a tryptophan in *ArCE4A* (Trp171) in motif 4 [DxxD(W/Y),

285 Fig. 1f). Of the 54 CE4 proteins listed in CAZy as characterized only *C*/CDA [34] and *An*CDA [15] have
 286 a tyrosine in motif 4.
 287



288
 289 **Fig. 2. Structure-based sequence alignment of CE4 deacetylases.** The structure-based sequence
 290 alignment was obtained using PyMod 1.0 [43]. Fully conserved residues are shown on a green

291 background. The asterisks indicate residues involved in metal binding (blue) and in catalysis (pink).
292 MT1-5 indicate the five conserved motifs in CE4 deacetylases. Colored horizontal bars indicate the
293 different loops described by Andrés et al. [16]. The deacetylases included in the alignment are:
294 *SpPgdA*, PDB id(2C1G [24]; *C/CDA*, PDB id 2IW0 [34]; *AnCDA*, PDB id 2Y8U [15]; *S/CE4*, PDB id 2CC0
295 [33]; *BsPdaA*, PDB id 1W17 [44]; *VcCDA*, PDB id 4NY2 [16]. For clarity, the alignment only shows the
296 sequence area of the five motifs and the loops.

297

298 **3.3 Enzymatic activity and substrate specificity**

299 Functional features of *ArCE4A* were investigated by testing the enzyme's activity against different
300 substrates and by sequence analysis of generated products. Table 2 shows the deacetylating activity
301 of *ArCE4A* for different substrates. For CHOS substrates, the apparent rate constant increased with
302 increasing DP up to (GlcNAc)₅, for which *ArCE4A* has a higher apparent rate against (0.18 s⁻¹)
303 compared to (GlcNAc)₆ (0.07 s⁻¹). A similar pattern of activity against CHOS was observed for *AnCDA*
304 [15]. *ArCE4A* did not deacetylate GlcNAc, and the activity against (GlcNAc)₂ was very low. Next to
305 CHOS, *ArCE4A* deacetylates chitosan, chitin and acetyl xylan (Table 2).

306

307

308

309

310

311

312

313

314

315 **Table 2. Activity of ArCE4A against different substrates.** The substrate was incubated with 300 nM
 316 ArCE4A for 30 min at 37 °C, and released acetic acid was measured by ion exchange HPLC.

Substrate	Substrate concentration	ASAR ^a (μM)	Average acetic acid release (μM)	CV%	Deacetylation degree (%)	Acetic acid released (nmol/min)	Apparent rate constant (s ⁻¹)
GlcNAc	2 mM	2000	0.0	0.0	0.000	0.00	0.00
(GlcNAc) ₂	2 mM	4000	0.1	14.5	0.003	0.00	0.00
(GlcNAc) ₃	2 mM	6000	11.9	2.6	0.20	0.04	0.02
(GlcNAc) ₄	2 mM	8000	39.8	2.8	0.50	0.13	0.07
(GlcNAc) ₅	2 mM	10000	95.8	7.7	0.96	0.32	0.18
(GlcNAc) ₆	2 mM	12000	39.4	0.2	0.33	0.13	0.07
Chitosan ^b	5 mg/ml	16000	85.4	1.4	0.53	0.28	0.16
α-chitin	5 mg/ml	24600 ^d	0.8	6.4	0.003	0.00	0.00
β-chitin	5 mg/ml	24600 ^d	1.4	1.9	0.006	0.001	0.00
Acetyl xylan ^c	5 mg/ml	9000	1696.7	3.1	18.9	5.66	3.14

317 ^aASAR: amount of substrate expressed as the concentration of acetyl groups. ^bThe degree of acetylation was
 318 64%. ^cMW_{avg} = 2800, degree of acetylation roughly estimated to be around 50% by MALDI-TOF. ^dAssuming
 319 one acetylation per sugar unit.

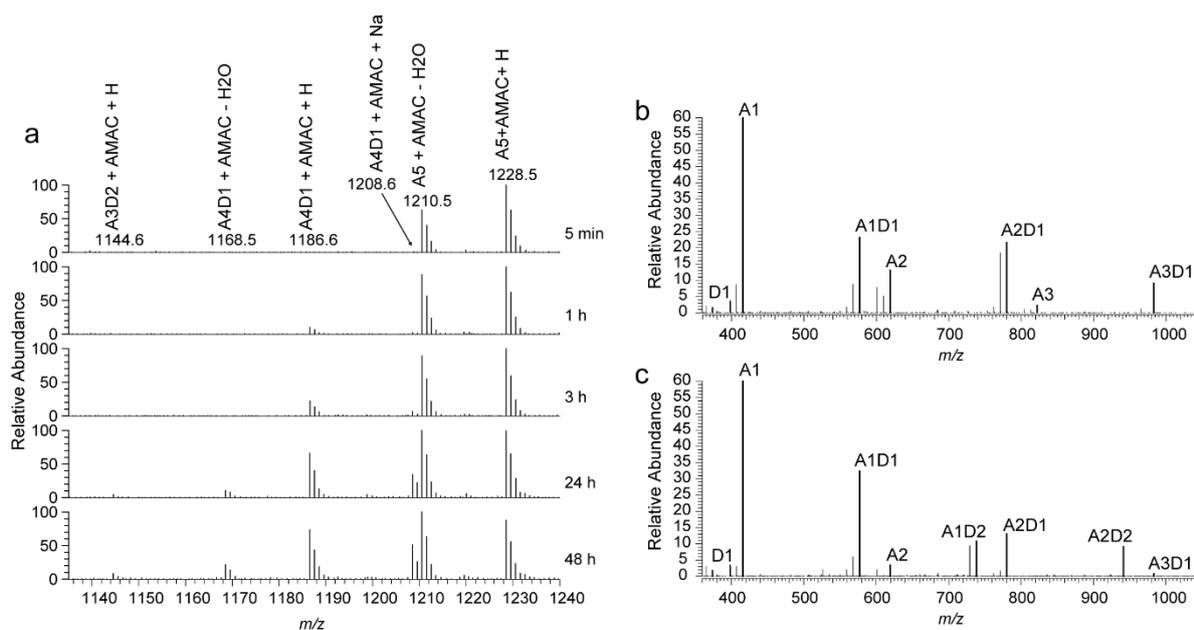
320

321 It is well known that CE4 enzymes tend to have broad substrate specificities. For example,
 322 enzymes classified as peptidoglycan deacetylases can deacetylate chito-oligomers [24, 32]. Likewise,
 323 CE4 enzymes known as acetylxylan esterases can deacetylate chitosan and CHOS [32, 45]. However,
 324 comparative information on rates is scarce. A recently described putative fungal CDA (*AnCDA*)
 325 showed in general higher rates for various substrates [15], compared to *ArCE4A*. Both *AnCDA* and
 326 *ArCE4A* are clearly most active towards acetylxylan and should thus perhaps, based on the available
 327 data, be classified as acetylxylan esterases [33, 45]. A further quantitative comparison of the activity
 328 of known CDAs towards chitinous substrates and acetylated plant polysaccharides such as
 329 acetylxylan would be of interest and could perhaps yield more insight into the true biological function
 330 of these enzymes.

331 Of the CHOS tested, *ArCE4A* showed highest activity against (GlcNAc)₅, and, therefore, this
 332 substrate was used for investigation of the position of deacetylation. The reducing ends of reaction
 333 products were labeled with AMAC and the resulting samples were analyzed using mass spectroscopy.

334 MS1 spectra of AMAC-labeled products obtained at different reaction times (Fig. 3a) show the initial
335 appearance of mono-deacetylated products (m/z 1186.6) and the subsequent appearance of
336 products with two deacetylations (m/z 1144.6) after 24 hours. The peaks for mono- and di-
337 deacetylated products were isolated and subjected to fractionation in MS2 experiments (Fig. 3b and
338 c). Although a signal corresponding to GlcN-AMAC (m/z 374) is visible, a signal at m/z 416,
339 corresponding to GlcNAc-AMAC, dominates in the MS2 spectra, indicating that the reducing end was
340 hardly deacetylated. The MS2 spectrum for the mono-deacetylated product (Fig. 3b) shows no signal
341 that would indicate deacetylation of the non-reducing end (i.e. no A4-AMAC signal), indicating that
342 the non-reducing end is not preferred for deacetylation. This may seem contradictory to the binding
343 mode of the (GlcNAc)₂ ligand seen in the structure where the non-reducing end is bound in subsite
344 0. It should be noted, however that the structure only shows part of the used substrate, (GlcNAc)₄,
345 and that it is thus not certain whether the chain “end” seen in the structure really is a chain end. The
346 fact that no non-reducing end deacetylation is observed in Fig. 3b may be taken to indicate that there
347 must be some substrate affinity beyond subsites 0 and +1, in particular in what would be -1 and -2
348 subsites. Notably, the presence of a weak signal for A3D1 in Fig. 3c, showing MS2 data for the double
349 deacetylated product, shows that deacetylation of the non-reducing end did occur. The relative
350 intensity of this signal is low, indicating that the non-reducing end is less preferred for deacetylation
351 compared to the middle sugars of the pentamer.

352



353

354 **Fig. 3. Mass spectrometric analysis of products generated from (GlcNAc)₅.** Reaction products
 355 generated upon treating (GlcNAc)₅ with ArCE4A were labeled with AMAC and analyzed by MS. (a)
 356 MS1 spectra of AMAC labeled reaction products at different reaction time points, showing
 357 appearance of mono- and di-deacetylated products. (b) Representative MS2 spectrum for the A4D1
 358 peak at *m/z* 1186 from MS1 spectra (1h reaction time). (c) Representative MS2 spectrum for the
 359 A3D2 peak at *m/z* 1144 from MS1 spectra (24h reaction time). Bold signals in (b) and (c) correspond
 360 to the mass of the indicated CHOS plus AMAC and hydrogen. Reaction mixtures contained 2 mM A5
 361 and 300 nM ArCE4A and were incubated at 37 °C.

362

363 The signals in Fig. 3b show that the first deacetylation happens at all three internal positions.
 364 Although quantitative interpretation of the MS spectra is not very reliable, the data do seem to
 365 suggest that deacetylation near the reducing end is most frequent (suggested by the strong A1D1
 366 signal). The products with two deacetylations seem to be dominated by deacetylation of the sugar
 367 next to the reducing end and of either of the two other internal sugars. The active site of C/CDA bears
 368 resemblance to that of ArCE4A (see Fig. 1f) and the kinetics of this enzyme have been studied in
 369 detail. For C/CDA acting on (GlcNAc)₄, the first deacetylation is fast, while the subsequent
 370 deacetylations are slower [13]. It was also shown that C/CDA deacetylates the reducing much more
 371 slowly than all other positions [13]. Our results indicate that, like in the case of C/CDA, the reducing
 372 end is less preferred by ArCE4A. This conclusion coincides with the structural data for the enzyme-

373 substrate complex, showing a strong binding interaction in the +1 subsite. This suggests that *ArCE4A*
374 prefers a sugar bound in the +1 subsite for optimal activity, and thus will not be very active on
375 reducing ends.

376 It should be noted that *ArCE4A* showed very low activity against (GlcNAc)₂ (Table 1), which
377 suggests that occupation of more than two subsites, i.e. beyond subsite 0 and +1, is beneficial for
378 activity. Currently available data do not allow a prediction of what additional interactions could
379 benefit catalysis. Studies with *VcCDA*, which, notably, has a very differently shaped catalytic center
380 (see above), suggested that substrate-binding could lead to conformational changes, which in the
381 case of *ArCE4A* could lead to interactions that we cannot detect in the current data.

382

383 **4. Concluding remarks**

384 In this study, we present structural and functional data for *ArCE4A*, including the first structural data
385 for a complex between a low-specificity CE4 enzyme with an open active site and a substrate. While
386 our motivation for this work was to develop enzymes for chitin processing, it is not certain that
387 deacetylation of GlcNAc is the true biological function of *ArCE4A*. If chitin were the natural substrate
388 one would perhaps expect a higher activity against chitin, chitosan and CHOS compared to acetyl
389 xylan (Table 3). A similar trend in substrate specificity was observed for *AnCDA*, which is thought to
390 be a fungal chitin deacetylase [15]. Interestingly, xylan is found in the cell wall of some marine algae
391 [46], and it is therefore conceivable that certain marine bacteria may benefit from the ability to
392 deacetylate this substrate. The broad substrate specificity observed for *ArCE4A* and other CE4s [15,
393 45] is intriguing, and more comparable studies are needed to fully understand the substrate
394 specificity.

395 The crystal structure of *ArCE4A* in complex with (GlcNAc)₂ provides a deeper understanding
396 of how CE4 enzymes interact with their substrates, especially CE4s with an open active site, which
397 are common in Nature. The structural data suggest that there are relatively few interactions between
398 the substrate and the enzyme beyond subsites 0 and +1. The interaction in subsite +1 involves a

399 tryptophane and is thus not very sugar specific, whereas more specific interactions in the form of
400 hydrogen bonds occur in subsite 0. This interaction pattern is compatible with the observed broad
401 specificity of the enzyme. It should be noted, however, that the activity of *ArCE4A* against (GlcNAc)₂
402 is low compared to other (longer) substrates, suggesting that unknown interactions, perhaps
403 involving conformational changes, take place upon substrate binding (e.g. loop rearrangements
404 [16]). Still, it is conceivable that a seemingly short and open substrate binding groove is an intentional
405 feature of these enzymes in order to fit different substrates in the active site. Structural data for
406 *ArCE4A* in complex with longer substrate and different substrates would be of great interest and will
407 be useful for better understanding the functionality of the CE4s. Such additional information may
408 eventually also create possibilities for using these enzymes, or engineered variants thereof, to
409 produce chitosans and CHOS with defined patterns of acetylation.

410

411 **Acknowledgements**

412 We would like to thank Dr. Jane W. Agger and Dr. Magnus Øverlie Arntzen for technical assistance
413 with HPLC and MS analysis.

414 **References**

- 415 1. Hackman R. Studies on chitin I. Enzymic degradation of chitin and chitin esters. Aust J Biol
416 Sci. 1954;7(2):168-78. doi: 10.1071/BI9540168.
- 417 2. Vårum KM, Ottøy MH, Smidsrød O. Acid hydrolysis of chitosans. Carbohydr Polym.
418 2001;46(1):89-98. doi: 10.1016/S0144-8617(00)00288-5.
- 419 3. Rinaudo M. Chitin and chitosan: Properties and applications. Prog Polym Sci. 2006;31(7):603-
420 32. doi: 10.1016/j.progpolymsci.2006.06.001.
- 421 4. Aam BB, Heggset EB, Norberg AL, Sørli M, Vårum KM, Eijsink VGH. Production of
422 chitooligosaccharides and their potential applications in medicine. Mar Drugs.
423 2010;8(5):1482-517. doi: 10.3390/md8051482.
- 424 5. Xia W, Liu P, Zhang J, Chen J. Biological activities of chitosan and chitooligosaccharides. Food
425 Hydrocolloids. 2011;25(2):170-9. doi: 10.1016/j.foodhyd.2010.03.003.
- 426 6. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-
427 active enzymes database (CAZy) in 2013. Nucleic Acids Res. 2014;42(D1):D490-5. Epub
428 2013/11/26. doi: 10.1093/nar/gkt1178.
- 429 7. Fukushima T, Kitajima T, Sekiguchi J. A polysaccharide deacetylase homologue, PdaA, in
430 *Bacillus subtilis* acts as an *N*-acetylmuramic acid deacetylase in vitro. J Bacteriol.
431 2005;187(4):1287-92. doi: 10.1128/JB.187.4.1287-1292.2005.
- 432 8. Bui NK, Turk S, Buckenmaier S, Stevenson-Jones F, Zeuch B, Gobec S, et al. Development of
433 screening assays and discovery of initial inhibitors of pneumococcal peptidoglycan
434 deacetylase PgdA. Biochem Pharmacol. 2011;82(1):43-52. doi: 10.1016/j.bcp.2011.03.028.
- 435 9. Biely P, Côté GL, Kremnický L, Greene RV, Dupont C, Kluepfel D. Substrate specificity and
436 mode of action of acetylxylan esterase from *Streptomyces lividans*. FEBS Lett. 1996;396(2-
437 3):257-60. doi: 10.1016/0014-5793(96)01080-0.

- 438 10. Cord-Landwehr S, Melcher RLJ, Kolkenbrock S, Moerschbacher BM. A chitin deacetylase from
439 the endophytic fungus *Pestalotiopsis* sp. efficiently inactivates the elicitor activity of chitin
440 oligomers in rice cells. *Sci Rep.* 2016;6:38018. doi: 10.1038/srep38018.
- 441 11. Naqvi S, Cord-Landwehr S, Singh R, Bernard F, Kolkenbrock S, El Gueddari NE, et al. A
442 recombinant fungal chitin deacetylase produces fully defined chitosan oligomers with novel
443 patterns of acetylation. *Appl Environ Microbiol.* 2016;82(22):6645-55. doi:
444 10.1128/AEM.01961-16.
- 445 12. Hamer SN, Cord-Landwehr S, Biarnes X, Planas A, Waegeman H, Moerschbacher BM, et al.
446 Enzymatic production of defined chitosan oligomers with a specific pattern of acetylation
447 using a combination of chitin oligosaccharide deacetylases. *Sci Rep.* 2015;5:8716. doi:
448 10.1038/srep08716.
- 449 13. Hekmat O, Tokuyasu K, Withers SG. Subsite structure of the endo-type chitin deacetylase
450 from a Deuteromycete, *Colletotrichum lindemuthianum*: an investigation using steady-state
451 kinetic analysis and MS. *Biochem J.* 2003;374(2):369-80. doi: 10.1042/bj20030204.
- 452 14. Tang M-C, Nisole A, Dupont C, Pelletier JN, Waldron KC. Chemical profiling of the deacetylase
453 activity of acetyl xylan esterase A (AxeA) variants on chitooligosaccharides using hydrophilic
454 interaction chromatography–mass spectrometry. *J Biotechnol.* 2011;155(2):257-65. doi:
455 10.1016/j.jbiotec.2011.06.041.
- 456 15. Liu Z, Gay LM, Tuveng TR, Agger JW, Westereng B, Mathiesen G, et al. Structure and function
457 of a broad-specificity chitin deacetylase from *Aspergillus nidulans* FGSC A4. *Sci Rep.*
458 2017;7(1):1746. doi: 10.1038/s41598-017-02043-1.
- 459 16. Andrés E, Albesa-Jové D, Biarnés X, Moerschbacher BM, Guerin ME, Planas A. Structural basis
460 of chitin oligosaccharide deacetylation. *Angew Chem Int Ed.* 2014;53(27):6882-7. doi:
461 10.1002/anie.201400220.

- 462 17. Webb EC. Enzyme nomenclature 1992. Recommendations of the Nomenclature Committee
463 of the International Union of Biochemistry and Molecular Biology on the Nomenclature and
464 Classification of Enzymes: Academic Press; 1992.
- 465 18. Ni Y, Li J, Panagiotou G. A molecular-level landscape of diet-gut microbiome interactions:
466 toward dietary interventions targeting bacterial genes. *MBio*. 2015;6(6):e01263-15.
- 467 19. Savitsky P, Bray J, Cooper CD, Marsden BD, Mahajan P, Burgess-Brown NA, et al. High-
468 throughput production of human proteins for crystallization: the SGC experience. *J Struct*
469 *Biol*. 2010;172(1):3-13. doi: 10.1016/j.jsb.2010.06.008.
- 470 20. Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids*
471 *Res*. 1990;18(20):6069-74.
- 472 21. Kabsch W. Xds. *Acta Crystallogr Sect D Biol Crystallogr*. 2010;66(2):125-32. doi:
473 10.1107/S0907444909047337.
- 474 22. Krug M, Weiss MS, Heinemann U, Mueller U. XDSAPP: a graphical user interface for the
475 convenient processing of diffraction data using XDS. *J Appl Crystallogr*. 2012;45(3):568-72.
476 doi: 10.1107/S0021889812011715.
- 477 23. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser
478 crystallographic software. *J Appl Crystallogr*. 2007;40(4):658-74. doi:
479 10.1107/S0021889807021206.
- 480 24. Blair DE, Schuttelkopf AW, MacRae JI, van Aalten DM. Structure and metal-dependent
481 mechanism of peptidoglycan deacetylase, a streptococcal virulence factor. *Proc Natl Acad*
482 *Sci U S A*. 2005;102(43):15429-34. doi: 10.1073/pnas.0504339102.
- 483 25. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a
484 comprehensive Python-based system for macromolecular structure solution. *Acta*
485 *Crystallogr Sect D Biol Crystallogr*. 2010;66(2):213-21. doi: 10.1107/S0907444909052925.

- 486 26. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. REFMAC5
487 for the refinement of macromolecular crystal structures. *Acta Crystallogr Sect D Biol*
488 *Crystallogr.* 2011;67(4):355-67. doi: 10.1107/S0907444911001314.
- 489 27. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4
490 suite and current developments. *Acta Crystallogr Sect D Biol Crystallogr.* 2011;67(4):235-42.
491 doi: 10.1107/S0907444910045749.
- 492 28. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta*
493 *Crystallogr Sect D Biol Crystallogr.* 2010;66(4):486-501. doi: 10.1107/S0907444910007493.
- 494 29. Schüttelkopf AW, Van Aalten DM. PRODRG: a tool for high-throughput crystallography of
495 protein–ligand complexes. *Acta Crystallogr Sect D Biol Crystallogr.* 2004;60(8):1355-63.
- 496 30. Bahrke S, Einarsson JM, Gislason J, Haebel S, Letzel MC, Peter-Katalinic J, et al. Sequence
497 analysis of chitooligosaccharides by matrix-assisted laser desorption ionization postsorce
498 decay mass spectrometry. *Biomacromolecules.* 2002;3(4):696-704. doi:
499 10.1021/bm020010n.
- 500 31. Morelle W, Page A, Michalski JC. Electrospray ionization ion trap mass spectrometry for
501 structural characterization of oligosaccharides derivatized with 2-aminobenzamide. *Rapid*
502 *Commun Mass Spectrom.* 2005;19(9):1145-58. doi: 10.1002/rcm.1900.
- 503 32. Caufrier F, Martinou A, Dupont C, Bouriotis V. Carbohydrate esterase family 4 enzymes:
504 substrate specificity. *Carbohydr Res.* 2003;338(7):687-92. doi: 10.1016/s0008-
505 6215(03)00002-8.
- 506 33. Taylor EJ, Gloster TM, Turkenburg JP, Vincent F, Brzozowski AM, Dupont C, et al. Structure
507 and activity of two metal ion-dependent acetylxyln esterases involved in plant cell wall
508 degradation reveals a close similarity to peptidoglycan deacetylases. *J Biol Chem.*
509 2006;281(16):10968-75. doi: 10.1074/jbc.M513066200.

- 510 34. Blair DE, Hekmat O, Schüttelkopf AW, Shrestha B, Tokuyasu K, Withers SG, et al. Structure
511 and mechanism of chitin deacetylase from the fungal pathogen *Colletotrichum*
512 *lindemuthianum*. *Biochemistry*. 2006;45(31):9416-26. doi: 10.1021/bi0606694.
- 513 35. Urch JE, Hurtado-Guerrero R, Brosson D, Liu Z, Eijsink VG, Texier C, et al. Structural and
514 functional characterization of a putative polysaccharide deacetylase of the human parasite
515 *Encephalitozoon cuniculi*. *Protein Sci*. 2009;18(6):1197-209. doi: 10.1002/pro.128.
- 516 36. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides
517 from transmembrane regions. *Nat Methods*. 2011;8(10):785-6. doi: 10.1038/nmeth.1701.
- 518 37. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based
519 environment for protein structure homology modelling. *Bioinformatics*. 2006;22(2):195-201.
520 Epub 2005/11/23. doi: 10.1093/bioinformatics/bti770.
- 521 38. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, et al. SWISS-MODEL:
522 modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic*
523 *Acids Res*. 2014;42(Web Server issue):W252-8. doi: 10.1093/nar/gku340.
- 524 39. Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T. Protein structure homology
525 modeling using SWISS-MODEL workspace. *Nat Protocols*. 2009;4(1):1-13. doi:
526 10.1038/nprot.2008.197.
- 527 40. Peters W. Chitin in tunicata. *Experientia*. 1966;22(12):820-1. doi: 10.1007/Bf01897437.
- 528 41. Jeuniaux C, editor Distribution and quantitative importance of chitin in animals. *Proceedings*
529 *of the First International Conference on Chitin/Chitosan; 1978; Cambridge (Massachusetts)*
530 *USA, 1978.*
- 531 42. Manucharova N, Vlasenko A, Zenova G, Dobovol'skaya T, Stepanov A. Methodological
532 aspects of assessing chitin utilization by soil microorganisms. *Biol Bull*. 2008;35(5):549-53.
533 PubMed PMID: 18956744.

- 534 43. Bramucci E, Paiardini A, Bossa F, Pascarella S. PyMod: sequence similarity searches, multiple
535 sequence-structure alignments, and homology modeling within PyMOL. BMC
536 Bioinformatics. 2012;13(4):S2. doi: 10.1186/1471-2105-13-s4-s2.
- 537 44. Blair DE, van Aalten DMF. Structures of *Bacillus subtilis* PdaA, a family 4 carbohydrate
538 esterase, and a complex with *N*-acetyl-glucosamine. FEBS Lett. 2004;570(1-3):13-9. doi:
539 10.1016/j.febslet.2004.06.013.
- 540 45. Puchart V, Gariépy M-C, Shareck F, Dupont C. Identification of catalytically important amino
541 acid residues of *Streptomyces lividans* acetylxylan esterase A from carbohydrate esterase
542 family 4. Biochim Biophys Acta. 2006;1764(2):263-74. doi:
543 dx.doi.org/10.1016/j.bbapap.2005.11.023.
- 544 46. Popper ZA, Tuohy MG. Beyond the green: Understanding the evolutionary puzzle of plant
545 and algal cell walls. Plant Physiol. 2010;153(2):373-83. doi: 10.1104/pp.110.158055.

Paper II

Proteomic investigation of the secretome of *Cellvibrio japonicus* during growth on chitin.

Tuveng, T. R., Arntzen, M. Ø., Bengtsson, O., Gardner, J. G., Vaaje-Kolstad, G. & Eijsink, V. G. H. 2016. *Proteomics*, 16, 1904-1914.*

*There is an error in Fig. 2A where four proteins lack LipoP annotation. A figure where all the LipoP annotation is visible is included at the end of the paper.

RESEARCH ARTICLE

Proteomic investigation of the secretome of *Cellvibrio japonicus* during growth on chitin

Tina Rise Tuveng¹, Magnus Øverlie Arntzen¹, Oskar Bengtsson¹, Jeffrey G. Gardner², Gustav Vaaje-Kolstad¹ and Vincent G.H. Eijsink¹

¹ Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences (NMBU), Aas, Norway

² Department of Biological Sciences, University of Maryland – Baltimore County, Baltimore, MD, USA

Studies of the secretomes of microbes grown on insoluble substrates are important for the discovery of novel proteins involved in biomass conversion. However, data in literature and this study indicate that secretome samples tend to be contaminated with cytoplasmic proteins. We have examined the secretome of the Gram-negative soil bacterium *Cellvibrio japonicus* using a simple plate-based culturing technique that yields samples with high fractions (60–75%) of proteins that are predicted to be secreted. By combining this approach with label-free quantification using the MaxLFQ algorithm, we have mapped and quantified proteins secreted by *C. japonicus* during growth on α - and β -chitin. Hierarchical clustering of the detected protein quantities revealed groups of up-regulated proteins that include all five putative *C. japonicus* chitinases as well as a chitin-specific lytic polysaccharide monooxygenase (CjLPMO10A). A small set of secreted proteins were co-regulated with known chitin-specific enzymes, including several with unknown catalytic functions. These proteins provide interesting targets for further studies aimed at unraveling the enzymatic machineries used by *C. japonicus* for recalcitrant polysaccharide degradation. Studies of chitin degradation indicated that *C. japonicus* indeed produces an efficient chitinolytic enzyme cocktail. All MS data have been deposited in the ProteomeXchange with the dataset identifier PXD002843 (<http://proteomecentral.proteomexchange.org/dataset/PXD002843>).

Received: October 28, 2015

Revised: April 5, 2016

Accepted: May 9, 2016

Keywords:

Cellvibrio japonicus / Chitin / Chitinase / Label-free quantification / Microbiology / Secretome



Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

Early proteomic studies were used primarily for qualitative measures, such as assessing the presence or absence of a specific protein [1]. Now, however, MS-based proteomic approaches have evolved into a powerful tool for investigating

complex protein samples. It is now possible to concurrently identify and quantify thousands of proteins from standard tryptic digests using a variety of different analytical techniques. Label-free quantification (LFQ) has gained popularity in recent years due to its simplicity, high dynamic range and robustness. These properties are all due to the availability of sophisticated computational algorithms [2, 3]. For example, label-free proteomics has already been successfully applied to study bacterial proteomes and secretomes, addressing several biological questions (e.g. [4, 5]). Here, we have combined the power of LFQ with a recently developed method [6] that

Correspondence: Professor Vincent G.H. Eijsink, Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, 1432 Aas, Norway
E-mail: vincent.eijsink@nmbu.no

Abbreviations: AA, auxiliary activity; CAZymes, carbohydrate active enzymes; CBM, carbohydrate-binding module; GH, glycosyl hydrolase; LPMO, lytic polysaccharide monooxygenase

Colour Online: See the article online to view Figs. 2 and 3 in colour.

Significance of the study

Studies of the secretomes of bacteria growing on insoluble substrates are of biotechnological and industrial interest, but are challenging because the secretome samples tend to be contaminated by cytoplasmic proteins and, until recently, protein quantification demanded complicated culturing or post-culture labeling techniques. We have employed a novel

and simple plate-based culturing technique yielding samples that are strongly enriched for truly secreted proteins. Our data indicate that combining this culturing approach with recently developed label-free quantification techniques yields high quality secretome data for a Gram-negative bacterium, *C. japonicus*, growing on different forms of chitin.

yields high quality secretome samples for microorganisms growing on insoluble substrates.

Chitin is an insoluble linear polysaccharide comprised of $\beta(1\rightarrow4)$ linked *N*-acetyl glucosamine residues and is often considered as the second most abundant biopolymer in nature. Chitin is primarily found in the environment as a part of the cell wall of fungi and the exoskeletons of crustaceans and insects. There has been increasing interest in the processing of chitin for the production of chitosan and chitooligosaccharides (CHOS), as these metabolites have a variety of applications in medicine and as antimicrobial agents [7–9]. Current industrial processes for the isolation of chitin and the production of chitosan and CHOS are harsh and not environmentally friendly [10, 11]. Therefore, it is desirable to find other ways to process chitin, such as by utilizing enzymatic methods. One of the best studied chitinolytic bacteria is the Gram-negative *Serratia marcescens*, which is an efficient chitin degrader [12] and known to produce several chitinases [13]. However, so far, little is known about the secretomes of chitinolytic bacteria growing on chitin.

Cellvibrio japonicus is a Gram-negative soil bacterium known for its ability to degrade plant cell wall polysaccharides, and notable for a large number of carbohydrate active enzymes (CAZymes) encoded in its genome [14–17]. The CAZyme repertoire of *C. japonicus* includes several putative chitinolytic enzymes, including two auxiliary activity (AA) family 10 proteins, four carbohydrate esterase (CE) family 4 proteins, one glycosyl hydrolase (GH) family 46 chitosanase, two GH20 hexosaminidases, four GH18 chitinases and one GH19 chitinase. While the plant cell wall polysaccharide-degrading machinery of *C. japonicus* has been studied for decades, not much is known about the bacterium's ability to degrade chitin or about the enzymes involved. The capacity to process chitin seems useful for a bacterium whose primary habitat is chitin-rich soil, which contains fungi and insect remains. Notably, the majority of the biomass degrading enzymes in *C. japonicus* are secreted by a type II secretion system (T2SS) and the bacterium shows limited growth on cellulose and chitin in the absence of this secretory system [16, 18]. Studies on bacterial chitin turnover often focus on the chitinases belonging to families GH18 and GH19, as well as the GH20 beta-glucosidase that is needed to convert the primary product of chitinases, chitobiose, into *N*-acetylglucosamine [13]. Other scenarios for chitin conversion have been ob-

served that involve partial deacetylation and the subsequent action of chitosanases and other types of beta-glucosidases [19, 20]. In its native form, chitin normally occurs in complex co-polymeric structures that contain proteins, other glycans and minerals (calcium carbonate). Enzymes involved in conversion of these other compounds are potentially of interest for processing of chitin-rich biomass, but are rarely targeted in research on enzymatic chitin conversion and largely unknown.

Here, we have used LFQ proteomics to identify proteins secreted by *C. japonicus* when grown on different chitin forms as the sole carbon source. We compare secretomes from shake flask cultures and from cells grown by a novel agar plate method [6] and use the latter method to obtain cell free secretomes that are highly enriched in truly secreted proteins. In this way, we identified which of the putative chitin-active *C. japonicus* enzymes that are involved in chitin degradation. Furthermore, correlations in expression patterns revealed proteins with unknown functions that putatively contribute to the conversion of chitin-rich biomass.

2 Materials and methods

2.1 Strain and media

C. japonicus Ueda107 was grown in shaking cultures and on plates with 2% wt/vol α -chitin (extracted from *Pandalus borealis*, Seagarden, Husøyvegen 278, Karmsund Fiskerihavn, 4262 Avaldsnes, Norway), 2% wt/vol β -chitin (extracted from squid pen, Batch 20140101, France Chitin, Chemin de Porte Claire, F- 84100 Orange, France), or 0.2% wt/vol glucose (VWR International) as sole carbon source in M9 minimal medium. The M9 minimal medium was supplemented with 1 mM $MgSO_4$ and 0.1 mM $CaCl_2$ and for the plates 1% agarose was used. The plates were essentially prepared according to Bengtsson et al. [6], using a sterile Supor 200 0.2 μ m membrane (Pall Life Sciences, Port Washington, NY) to separate cells from the site of protein harvesting, with the exception that we used glass petri dishes with a diameter of 80 mm, hence reducing the volume of medium from 20 to 16 mL. Plates were incubated in a heating cabinet, whereas liquid cultures were incubated with shaking at 200 rpm, both at 30°C.

2.2 Sample preparation

During growth in shake flasks, we collected samples (4 mL) from 90 mL cultures, throughout the growth cycle, with three biological replicates for each time point. Growth was monitored by measuring OD₆₀₀; to avoid interference from the insoluble α - and β -chitin, all samples were allowed to stand for 10 min in the cuvette before the OD was measured. Sterile culture supernatants were obtained by centrifugation and filtration (0.22 μ m). Proteins were then precipitated by adding 50% (v/v) ice cold trichloroacetic acid to a final concentration of 10% (v/v), followed by thorough mixing and storage overnight at 4°C. After centrifugation for 15 min at 4000 rpm in a fixed angle rotor, the protein pellets were washed three times with 0.01 M HCl/90% acetone, air dried and dissolved in 100 μ L 20 mM Tris HCl pH 8. The protein concentration was measured (Bradford microassay, Bio-Rad, USA) before reduction with DTT (added to a final concentration of 10 mM) and subsequent alkylation (in the dark) with iodo-acetamide (added to a final concentration of 15 mM). Reduction and alkylation were carried out at room temperature, the incubation time being 30 min for each step. Trypsin (Sequencing grade, Promega, USA) was added to reach a 1:40 (w/w) ratio relative to the protein concentration. After incubation at 37°C overnight, peptides were purified using ZipTip C18 pipette tips (Merck Millipore, Cork, Ireland), dried under vacuum (Concentrator plus, Eppendorf, Denmark) and dissolved in 10 μ L 2% (v/v) ACN, 0.1% (v/v) TFA.

Secretomes from cells grown on plates were collected at different time points, with three biological parallels, as described by Bengtsson et al. [6]. It was not possible to follow the growth on plates in the same way as for the shaking flask cultures. Since *C. japonicus* cells have a strong yellow color, we used color development as an indicator for growth, and performed pre-experiments to select suitable sampling points. The sample preparation was done as described by Bengtsson et al. [6], with the exception that trypsinated samples were lyophilized, in order to concentrate the samples, and dissolved in 0.1% TFA, before purification of peptides using ZipTip C18 pipette tips (Merck Millipore, Cork, Ireland). The purified peptides were dried and dissolved as described above.

2.3 Mass spectrometry

The peptides were analyzed with two technical replicates using a nanoHPLC-MS/MS system consisting of a Dionex Ultimate 3000 RSLCnano (Thermo Scientific, Bremen, Germany) connected to a Q-Exactive hybrid quadrupole-orbitrap mass spectrometer (Thermo Scientific, Bremen, Germany) equipped with a nano-electrospray ion source. Details about the MS analysis are given in Supporting Information.

2.4 Data analysis

The raw files from MS were imported into MaxQuant [21, 22] version 1.4.1.2 and proteins were identified and quantified

using the MaxLFQ algorithm [2]. The samples were searched against a database containing the proteome of *C. japonicus* Ueda 107, downloaded from Uniprot (3711 sequences) [23], supplemented with common contaminants such as keratins, trypsin and BSA. In addition, reversed sequences of all protein entries were concatenated to the database for estimation of false discovery rates. The results from MaxQuant were further processed using Perseus (version 1.5.1.6) and subcellular location of proteins was predicted using several bioinformatics servers. The Supporting Information provides further details of the data analysis.

2.5 Collection of chitinolytic cocktails and chitin degradation studies

C. japonicus Ueda107 and *S. marcescens* BJL200 were grown in M9 medium supplemented with 1 mM MgSO₄, 0.1 mM CaCl₂ and 0.5% wt/vol β -chitin as the carbon source, at 30°C. Right after all the chitin had been utilized by the bacteria (after approximately 5 days), the culture supernatant was collected by centrifugation and sterilized by filtration (0.45 μ m). Both chitinolytic cocktails were then tested for the ability to degrade β -chitin. Standard reactions consisted of 1.5 μ g protein/mg substrate (protein concentration determined as above), 7.5 mg/mL β -chitin and 20 mM BisTris pH 6.0. Reactions were incubated in a thermomixer at 37°C and 800 rpm. All reactions were performed in triplicates. Released chitobiose and N-acetylglucosamine products were analyzed on a Dionex Ultimate 3000 as previously described by Hamre et al. [24], with the exception that the absorbance was measured at 194 nm instead of 210 nm.

To confirm that the assumed chitin-active proteins were present in the culture supernatant from *C. japonicus*, we performed a standard in-gel digestion proteomics analysis, essentially as described by Shevchenko et al. [25] and analyzed the peptides on a nanoHPLC-MS/MS system as described above. The data were analyzed using MASCOT [26]. Regarding chitin active proteins secreted by *S. marcescens*, we know from recent unpublished work, in our group and published work by Watanabe and co-workers [27, 28] that *S. marcescens* secretes all enzymes known to be involved in chitin-degradation during growth on chitin.

3 Results and discussion

C. japonicus was grown in shaking flasks (Supporting Information Fig. 1 shows growth curves) and on agar plates containing α -chitin, β -chitin, or glucose as the sole carbon source. Secretomes were collected at different time points, analyzed by high resolution LC-MS/MS, and quantified with the MaxLFQ algorithm [2]. In addition, the ability of the *C. japonicus* chitinolytic cocktail to degrade β -chitin was investigated. Key samples and their analysis are discussed in detail below.

3.1 Comparison of secretomes from shaking flask cultures and agar plates

According to the LipoP 1.0 software [29], 18% of the proteins encoded by the genome of *C. japonicus* are secreted due to the presence of SpI or SpII leader peptides. If one includes proteins predicted as secreted by the twin-arginine pathway [30] and proteins predicted as secreted by non-classical mechanisms [31], the fraction of secreted proteins increases to 29%. In all shaking flask cultures the content of cytosolic proteins was high, while the content of proteins predicted as secreted was so low (20–30% and approx. 40%, excluding or including non-classical secretion, respectively) that the protein samples can hardly be considered as secretome specific (Fig. 1). The high total number of detected proteins (Supporting Information Table 1) and the low fraction of secreted proteins (Fig. 1) in the cultures with chitin indicate that cell lysis occurs. Lower protein numbers in the faster growing cultures with glucose suggest less cell lysis but the fractions of secreted proteins were still too low to consider the protein samples derived from these cultures as secretome specific. Notably, large fractions of proteins without classical signal peptides is common in the study of bacterial secretomes [32–34]. For example, in a study of the secretome of *Staphylococcus epidermidis* grown in tryptic soy broth 80% of the proteins were predicted to be cytosolic [34]. The reason for finding apparently cytosolic proteins in secretomes is debated in the literature; cell lysis is an obvious reason, but it is also possible that some of these “cytosolic” proteins are secreted via known or unknown non-classical mechanisms [35–39]. Some proteins are known to have both a cytoplasmic function and a function outside the cell, referred to as moonlighting proteins [40].

With the aim to enrich for truly secreted proteins, we adapted the agar plate method originally developed by Bengtsson et al. [6] for analyzing fungal secretomes. As shown in Fig. 1 and Supporting Information Table 1, this approach led to a reduction in the fraction of cytosolic proteins and, consequently, to protein fractions that were highly enriched for secreted proteins, with levels reaching approximately 60 and 75%, excluding or including non-classical secretion, respectively. Based on these findings, one may speculate that the use of large solid substrates in shaking flasks causes a shearing force on the bacteria, leading to non-natural cell damage and leakage of cytosolic proteins. Another important advantage of the plate approach is that the chitin particles are present in the sample during processing prior to the LC-MS/MS analysis, hence, proteins bound to the chitin particles will be present in the final sample. In contrast, before sample preparation from shaking flask cultures, the chitin particles are separated from the secretome by centrifugation, meaning that proteins attached to the chitin particles will not be present in the final sample. Interestingly, one would expect the plate method to be better in any case, because stationary growth on plates is somewhat closer to natural growth in soil than growth in shaking flask cultures. Therefore, based on the above results

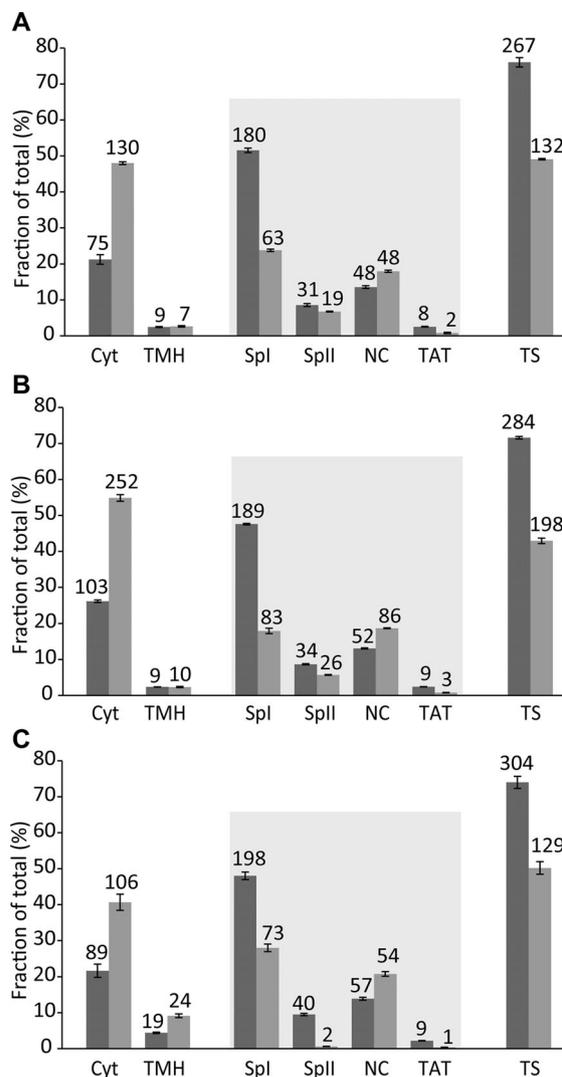


Figure 1. Numbers of detected proteins. The figure shows the relative quantities of detected proteins in shaking flask cultures (light grey) and on plates (dark grey) for the three carbon sources α -chitin (A), β -chitin (B) and glucose (C). The data are average values of three biological replicates, with error bars representing the standard deviation for the percentages. The average absolute numbers of proteins are indicated above each vertical bar. Detected proteins were classified using LipoP, PRED-TAT and SecretomeP. First LipoP and PRED-TAT were used to predict proteins as cytosolic (CYT), containing N-terminal transmembrane helices (TMH), secreted (SpI and SpII), or secreted through the twin-arginine arginine (TAT) pathway. Proteins predicted to be cytosolic by LipoP, but with a score above 0.5 in SecretomeP are classified as non-classically secreted proteins (NC). Proteins considered as secreted are highlighted by a grey box. “Total secreted (TS)” indicates the sum of all detected proteins that are predicted to be secreted; for the predicted complete *C. japonicus* proteome, this value is 29%. Three time points were analyzed for each of the six conditions (three substrates, two culturing conditions; Supporting Information Table 1). The figure only shows one time point for each condition. The time points shown, for flasks and plates, respectively, are, 96 and 144 h for α -chitin, 32 and 96 h for β -chitin, and 12 and 24 h for glucose.

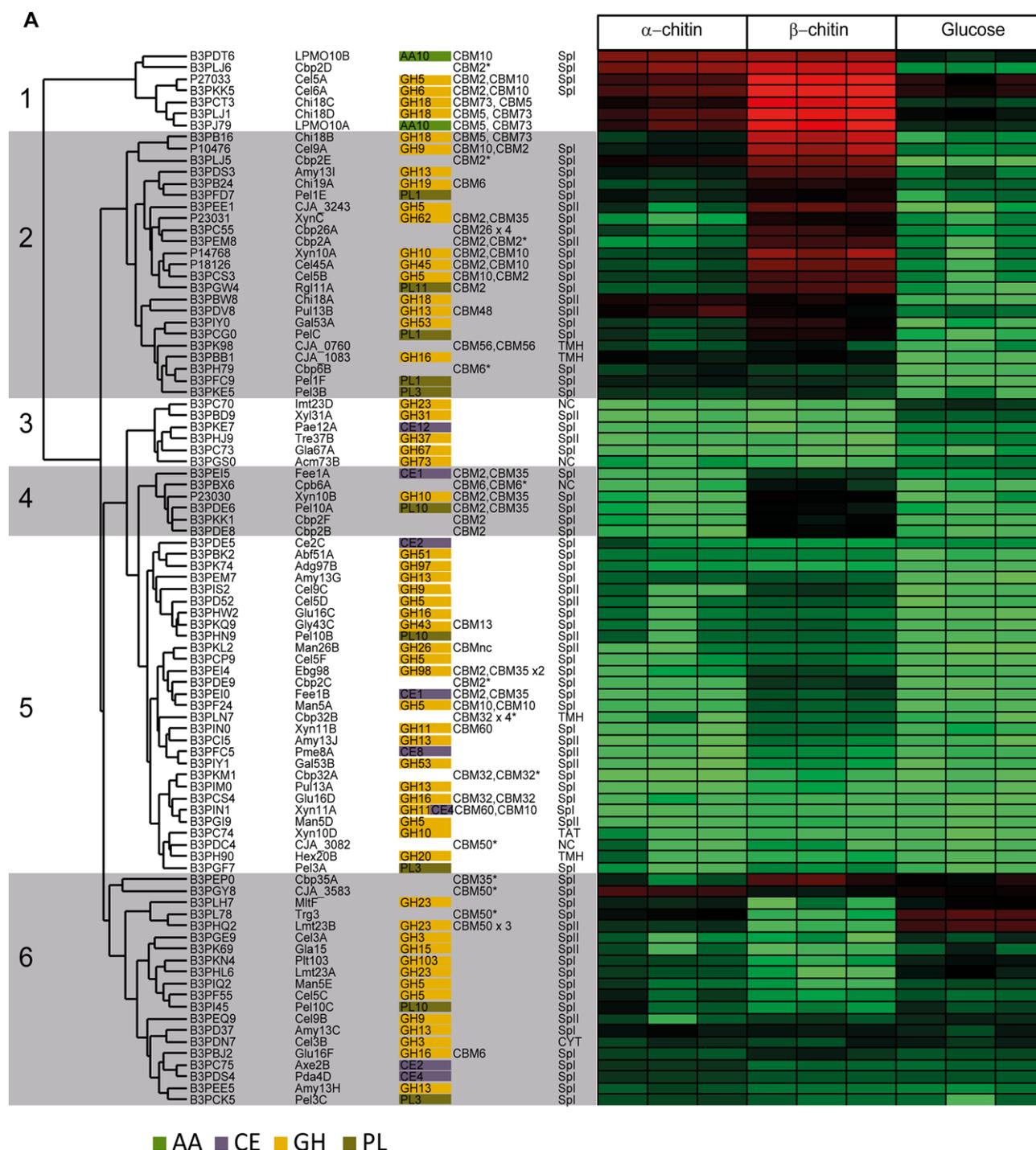


Figure 2. Heat map and intensity profiles for CAZymes. Panel A shows a heat map with identified CAZymes in the different samples (three samples for each substrate) divided into six clusters. The colors in the heat map indicate protein intensity, ranging from high (red color, MaxLFQ 5×10^{10}) to low intensity (green color, MaxLFQ 7×10^6). The Uniprot ID, protein name, CAZY family (auxiliary activity (AA), carbohydrate esterase (CE), glycosyl hydrolase (GH), polysaccharide lyase (PL)), and predicted cellular localization (abbreviated as in Fig. 1) are shown to the left of the heat map. Note that the heat map includes proteins only containing a carbohydrate-binding module (CBM) defined in CAZY; a * indicates that there are additional functional predictions for such a protein; see text. Panel B, Abundance profiles for proteins in the six clusters indicated in panel A.

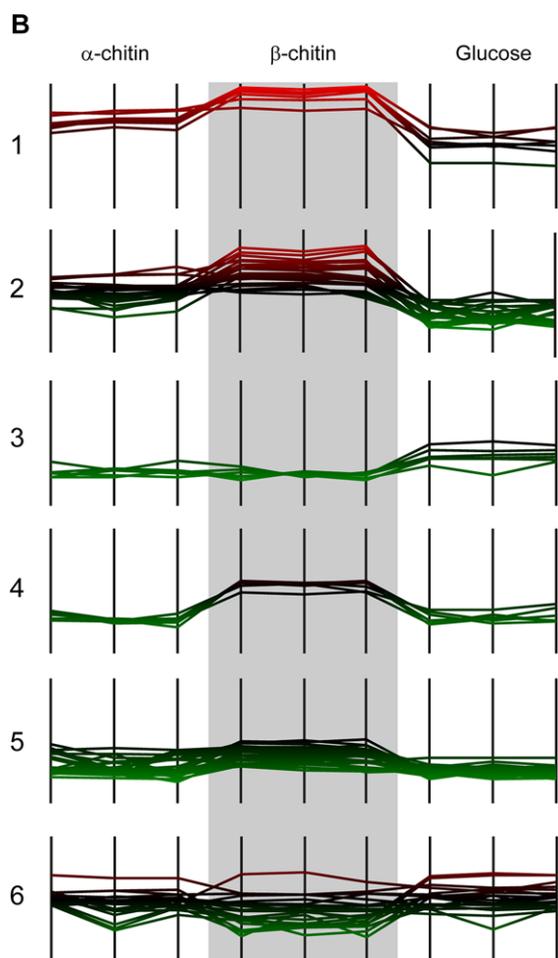


Figure 2. Continued.

and considerations, we only present data from the plate secretomes in the following sections.

Three time points were sampled for each carbon source after growth on plates and the raw data for each time point was quantified using MaxLFQ. Since we had no proper method for measuring growth on plates, we used development of the yellow color associated with the growth of *C. japonicus* as a rough indicator. Thus, the time points for harvesting varied between the carbon sources (Supporting Information Table 1), due to different growth rates, but the extend of growth (i.e. yellow color) at the first, second and third harvesting point were similar for the three substrates. For comparison of the three carbon sources, we focus on the samples from the second time point, which coincidentally, yielded similar amounts of proteins (approx. 400). A list of all proteins detected by this approach, 504 in total, including quantification data, is provided in Supporting Information Table 2.

3.2 Carbohydrate active enzymes (CAZymes)

The matrix from the comparative MaxLFQ analysis was sorted to select proteins annotated as CAZymes. Of the putative chitin-active enzymes, all the predicted chitinases (four GH18s and one GH19), the two AA10s, one of the two GH20s, and one of the four CE4 enzymes were identified, however, the GH46 chitosanase was not observed. Figure 2A shows a heat map of the CAZymes and after hierarchical clustering of the proteins and division into larger groups with similar regulation patterns, two clusters labeled 1 (seven proteins) and 2 (24 proteins) stand out. These clusters contain proteins with higher intensity in the secretomes from the α - and β -chitin samples, compared to the glucose samples (Fig. 2B). The five chitinases and the two AA10s occur in these two clusters, indicating that these clusters indeed harbor proteins related to the degradation of chitinous biomass.

At this time point (time point 2) all but one of the chitinases were detected in the glucose samples, albeit at low levels. However, at the earlier time point, time point 1, known putative chitin modifying proteins were only detected in the chitin samples (Supporting Information Table S3), strongly suggesting that *C. japonicus* indeed utilizes its putative chitinolytic machinery to degrade chitin. At this early time point, all chitinases were detected in the β -chitin samples, whereas only three, Chi18A, Chi18C and Chi18D, were detected in the α -chitin sample. The latter three enzymes may be the most important enzymes in the beginning of the chitin degradation process. Supporting Information Table 3 also shows that Hex20B (GH20) was only detected in the β -chitin sample, and at the second time point only. This enzyme cleaves dimers into monomers and it is therefore reasonable to believe that Hex20B appears later, compared to the chitinases, during the degradation of chitin. In the α -chitin secretome, Hex20B was not detected using the set criteria for identification. However, manual inspection of the data showed that this protein is present in low amounts (in one of the three biological replicates) at the second time point. Alpha-chitin is known to be more recalcitrant than β -chitin [41, 42], which could be one reason for the observed differences in the protein profiles from the different chitin samples.

The two AA10s were among the most abundant proteins and were clearly more abundant in the chitin samples compared to the glucose samples. One of these AA10s, CjLPMO10B, has previously been described by Gardner et al. [17]; it contains a CBM10 domain likely to bind to cellulose [43] and its activity against cellulose has been demonstrated [17]. The other AA10 protein (CjLPMO10A) contains a carbohydrate-binding module belonging to the CBM5 family that is known for chitin-binding [44]. In a recent study, Forsberg et al. confirmed that CjLPMO10B is strictly active on cellulose, whereas CjLPMO10A is only active on chitin. In the same study, a knockout of CjLPMO10A was shown to

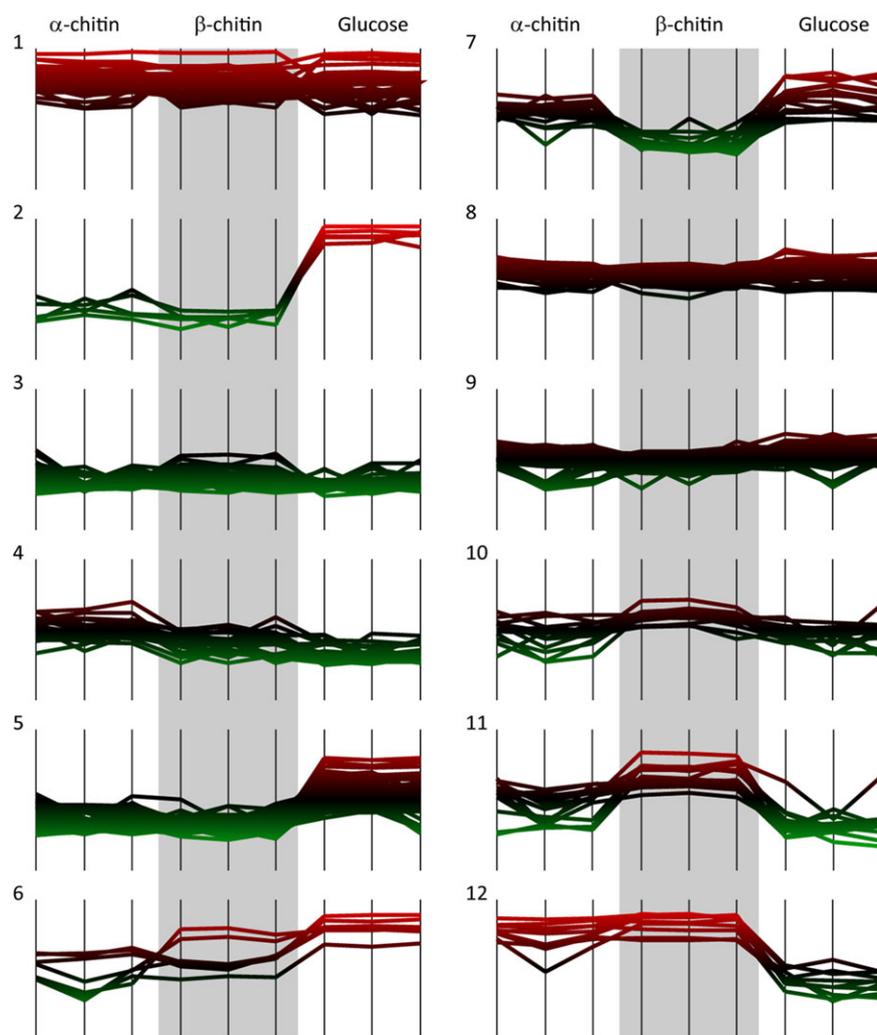


Figure 3. Intensity profiles for secreted non-CAZymes. The figure shows intensity profiles for proteins without CAZy annotation, but predicted as secreted by LipoP, PRED-TAT, or SecretomeP. Clustering is derived from the heat map in Supporting Information Fig. 2. The gradient colors indicate protein intensity, ranging from high (red color, $\text{MaxLFQ } 5 \times 10^{10}$) to low intensity (green color, $\text{MaxLFQ } 1 \times 10^6$).

cause a larger growth defect during growth on chitin compared to a *CjLPMO10B* knock out [18]. Based on these observations, the finding that *CjLPMO10B* is upregulated during growth on chitin is somewhat surprising. Perhaps upregulation of *CjLPMO10B* is part of a general response that is elicited when *C. japonicus* senses recalcitrant substrates (see below). Another explanation may be that *CjLPMO10B*, and perhaps LPMOs in general, have additional functions that yet are to be discovered.

Interestingly, many of the other CAZymes upregulated during growth on chitin (clusters 1 and 2 in Fig. 2) have putative roles in the conversion of cellulose (GH5, GH6, GH9, GH45), whereas enzymes putatively acting on other plant polysaccharides were also detected (e.g. GH10, xylan; PL1 & PL3, pectin; GH13, starch; GH16). CBM2 and CBM10 domains, known for their interactions with cellulose, are relatively abundant among the additional CAZymes. Cel5A [45] and Cel6A [17], each containing a CBM2 and a CBM10, are among the most abundant enzymes (Cluster 1 in Fig. 2). It is possible that some of the detected putative plant-cell-wall-

active enzymes in fact act on chitin, since chitin-modifying activities have been described for members of families GH5, GH9 and GH16 (www.cazy.org; [46]). Still, our data suggest that the response of *C. japonicus* to growth on chitin has a somewhat general nature in that enzymes capable of degrading a variety of recalcitrant biomasses are upregulated.

An observation of particular interest is the presence of Cbp2D in Cluster 1 and Cbp2E in Cluster 2. Cbp2D and Cbp2E contain carbohydrate-binding modules belonging to the CBM2 family whose members are known to bind both chitin and cellulose. Importantly, these proteins also contain predicted cytochrome domains for electron transfer. The observation that *cbp2D* and *cbp2E* knock-outs in *C. japonicus* showed impaired growth on cellulosic substrates led Gardner et al. [17] to speculate that these enzymes are important in oxidative cleavage of recalcitrant polysaccharides, possibly in concert with the LPMOs. Our data show co-regulation of Cbp2D with both LPMOs (Cluster 1), whereas Cbp2E is found in cluster 2 (Fig. 2) showing co-regulation with putatively chitin-specific proteins.

Table 1. List of non-CAZy secreted proteins upregulated in chitin samples. The table includes proteins found in clusters 11 and 12 in Fig. 4 and Supporting Information Fig. 2. The values given for the carbon sources are log₁₀ of MaxLFQ intensities, average of three biological replicates if not stated otherwise. Note that the SecretomeP scores suggest that all proteins predicted as CYT by LipoP in fact are secreted by non-classical pathways (i.e. SecP score >0.5). Using PRED-TAT, only two of the listed proteins were predicted to be secreted by the TAT-pathway (B3PI69 and B3PLJ7)

Cluster	Uniprot ID	Protein name	Mol. weight (kDa)	α-chitin	β-chitin	Glucose	LipoP	SecP score	
11	B3PLF1	von Willebrand factor type A domain protein	220.6	7.2	9.4	8.2	Spl	0.81	
	B3PDZ8	Uncharacterized protein	56.6	6.7	9.3	6.3 ^{a)}	Spl	0.43	
	B3PDZ9	Uncharacterized protein	29.1		9.1	6.0 ^{a)}	Spl	0.43	
	B3PI69 ^{b)}	High-potential iron-sulfur protein (HiPIP)	10.0	8.7 ^{a)}	8.7		Spl	0.87	
	B3PJI3	Putative Ig domain family	373.4	8.1 ^{a)}	8.4		Spl	0.94	
	B3PL02	Uncharacterized protein	36.8	7.9 ^{a)}	8.1		CYT	0.93	
	B3PB26	Putative lipoprotein	86.8	8.5	8.6		SpII	0.94	
	B3PFC4	Pectin methylesterase, putative, ce8	97.5	8.2	8.4		SpII	0.94	
	B3PC38	Uncharacterized protein	68.2	8.1	8.7		Spl	0.87	
	B3PI27	Uncharacterized protein	30.3	8.3	8.6	7.0 ^{a)}	Spl	0.18	
	B3PKZ2	Uncharacterized protein	74.0	8.2	8.7		CYT	0.94	
	B3PGR9	Flagellar hook-associated protein 1 FlgK	97.3	7.9	9.8	7.0 ^{a)}	CYT	0.95	
	B3PFW9	Putative lipoprotein	25.8	8.2	9.2		Spl	0.07	
	B3PD31	Putative hemolysin	241.6	8.1	9.2		Spl	0.91	
	B3PGR8	Flagellar hook-associated protein type 3 FlgL	55.7	7.9	9.2		CYT	0.96	
	12	B3PFF5	Putative lipoprotein	21.7	9.1	9.2		SpII	0.92
		B3PEH0	Uncharacterized protein	106.0	9.2	9.2	7.6 ^{a)}	CYT	0.56
B3PGR4		Flagellin	69.2	8.6	9.8	7.7	CYT	0.88	
B3PCJ3		Putative lipoprotein	80.5	10.1	10.2	8.0	SpII	0.95	
B3PGR3		Flagellar hook-associated protein 2	72.2	9.5	10.2	7.9	CYT	0.92	
B3PEH2		Uncharacterized protein	46.5	9.7	9.6	6.8	Spl	0.89	
B3PEH3		Putative lipoprotein	89.9	9.9	10.0	7.5	SpII	0.89	
B3PLJ7 ^{b)}		Conserved domain protein	42.6	9.6	10.2	7.3	CYT	0.93	
B3PGR6		Flagellin	70.5	9.3	10.2	7.2	CYT	0.96	

a) Average of two biological replicates.

b) Predicted to be secreted by the TAT-pathway using PRED-TAT

Cbp2D and Cbp2E are examples of the 16 proteins in Fig. 2 whose only CAZyme domains are one or more CBMs. Next to Cbp2D and Cbp2E, four such proteins occur in clusters 1 and 2. Next to the CBMs, B3PC55 and B3PK98 contain unknown domains amounting to approx. 95 and 71 kDa, respectively. B3PEM8 contains putative PKD/chitinase domains between its two CBM2 domains, according to Interpro [47]. MEROPS annotation [48] showed that B3PH79 contains a 21 kDa unassigned protease domain next to its CBM6 (see below).

Figure 2 shows that the great majority of the CAZymes, including all enzymes in Cluster 1, contain an Spl signal peptide for secretion. The second next abundant proteins are putative lipoproteins containing an SpII signal peptide. The occurrence of putative lipoproteins in secretome samples is commonly observed [38, 49, 50]. Four of the proteins in Fig. 2, including two members of Cluster 2 and Hex20B (in cluster 5), are predicted to contain an N-terminal trans-membrane helix (TMH) by LipoP. However, the TMHMM 2.0 server

[51, 52] predicts that only two of these proteins contain an N-terminal TMH. Since such a TMH sequence has properties similar to N-terminal signal sequences, it is actually possible that these proteins are secreted.

3.3 Other secreted proteins

In order to discover new proteins putatively involved in chitin degradation, it is important to investigate non-CAZy proteins found in the secretome. The matrix from the MaxLFQ comparative analysis of the three carbon sources was sorted to contain non-CAZy proteins, predicted to be secreted by LipoP 1.0 (Spl or SpII) [29], PRED-TAT (Tat) [30] or SecretomeP 2.0 (non-classically) [31]. The intensity profiles for different clusters (Fig. 3) derived from hierarchical clustering (Supporting Information Fig. 2) show that three clusters stand out: 2 (five proteins), 11 (15 proteins) and 12 (nine proteins). Cluster 11

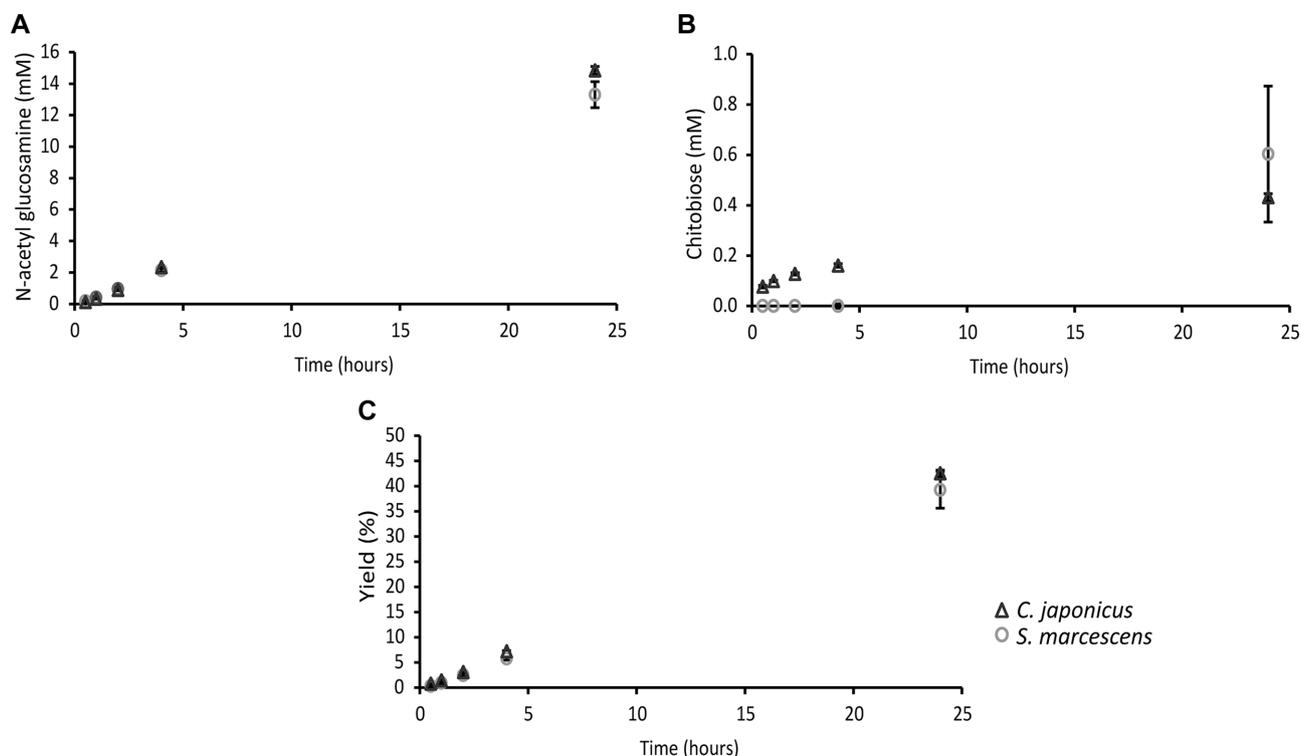


Figure 4. Product profiles and solubilization yields during chitin degradation. Panel A and B show the production of *N*-acetyl glucosamine and chitobiose over time during degradation of β -chitin with culture supernatants from *C. japonicus* and *S. marcescens*. Panel C shows the total solubilization yields. All data points are the average of three independent experiments, with error bars representing standard deviation.

and 12 show profiles with elevated protein intensities for the α - and/or β -chitin samples, while cluster 2 shows clearly elevated intensities for the glucose sample. Clusters 11 and 12 show similar expression profiles as clusters 1 and 2 of the CAZymes (Fig. 2), suggesting that these proteins may be important for growth on chitin biomass. In cluster 11 and 12, there are 24 proteins (Table 1), many of which with unknown or putative function. Notably, this is a rather low number of proteins, which could be targeted for further studies.

Considering the fact that chitin often occurs in protein-containing co-polymeric structures, we were particularly interested in detected proteases. Thirty-one of the in total 504 detected proteins received an annotation in MEROPS (Supporting Information Table 2). Twelve of these are not secreted and/or are predicted as non-peptidase homologues. Of the remaining 19, 12 are unassigned, whereas seven are assigned to a known MEROPS family (Supporting Information Table 4). Six of the putative proteases are perhaps of interest, either because they are upregulated on chitin (B3PIU45, B3P100, B3PD75, B3PJ05) albeit often to low extent (Supporting Information Table 4) or because they contain a CBM (B3PH79, B3PL78), which may suggest activity on carbohydrate-associated substrates. Interestingly, the CBM6 containing protein combines both; it is part of cluster 2 in Fig. 2, meaning that this protein is clearly upregulated dur-

ing growth on chitin. The CBM50 (B3PL78) containing putative protease in cluster 6 of Fig. 2, is not upregulated during growth on chitin.

3.4 Chitin-degrading capacity of *C. japonicus*

C. japonicus and *S. marcescens* were grown in a medium containing β -chitin and the culture supernatant was collected for investigation of its chitin degrading ability. The culture supernatants were collected when almost all of the chitin had been utilized by the bacteria, to avoid enzyme loss by adsorption to the substrate. Proteomic analysis of the *C. japonicus* culture supernatant confirmed the presence of the four GH18s, the two AA10s, the GH19 and a GH20 (Hex20B) (data not shown), implying that the supernatant should be able to degrade chitin. Indeed, the chitinolytic cocktail of *C. japonicus* was able to degrade β -chitin (Fig. 4). The *C. japonicus* and *S. marcescens* supernatants performed similarly, converting the β -chitin to *N*-acetyl glucosamine and a minor fraction of chitobiose produced. After 24 h of incubation the total solubilization yield obtained with the *C. japonicus* cocktail was 43%, compared to 39% for the *S. marcescens* cocktail (Fig. 4C). This shows that the two bacterial enzyme cocktails degrade the β -chitin in a similar way.

4 Concluding remarks

In this quantitative proteomics study, we have analyzed the secretome of *C. japonicus* and demonstrated that this bacterium is able to grow on α - and β -chitin utilizing the putative chitin-degrading proteins encoded in its genome. The data showed clear up-regulation of the four GH18 chitinases, the only GH19 chitinase and CjLPMO10A, a chitin-specific LPMO. The use of a novel agar plate-based method for bacterial growth and sample preparation yielded improved sampling of secretome-specific proteins. Thus, the limited number of secreted proteins found to be co-regulated with known chitin-specific enzymes, in particular those with not-yet annotated catalytic functions, provide an interesting target list for further studies aimed at discovering new elements of the enzyme machinery used by *C. japonicus* for converting biomass.

This research was supported by the Research Council of Norway through the MarPol Project 221576. The proteomics data has been deposited to the ProteomeXchange consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository [53] with the data set identifier PXD002843.

The authors have declared no conflict of interest.

5 References

- [1] Nahnsen, S., Bielow, C., Reinert, K., Kohlbacher, O., Tools for label-free peptide quantification. *Mol. Cell. Proteomics* 2013, 12, 549–556.
- [2] Cox, J., Hein, M. Y., Luber, C. A., Paron, I. et al., Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* 2014, 13, 2513–2526.
- [3] Weisser, H., Nahnsen, S., Grossmann, J., Nilse, L. et al., An automated pipeline for high-throughput label-free quantitative proteomics. *J. Proteome Res.* 2013, 12, 1628–1644.
- [4] Müller, J. E. N., Litsanov, B., Bortfeld-Miller, M., Trachsel, C. et al., Proteomic analysis of the thermophilic methylotroph *Bacillus methanolicus* MGA3. *Proteomics* 2014, 14, 725–737.
- [5] Wen, Y.-T., Wang, J.-S., Tsai, S.-H., Chuan, C.-N. et al., Label-free proteomic analysis of environmental acidification-influenced *Streptococcus pyogenes* secretome reveals a novel acid-induced protein histidine triad protein A (HtpA) involved in necrotizing fasciitis. *J. Proteomics* 2014, 109, 90–103.
- [6] Bengtsson, O., Arntzen, M. Ø., Mathiesen, G., Skaugen, M., Eijsink, V. G. H., A novel proteomics sample preparation method for secretome analysis of *Hypocrea jecorina* growing on insoluble substrates. *J. Proteomics* 2016, 131, 104–112.
- [7] Aam, B. B., Heggset, E. B., Norberg, A. L., Sorlie, M. et al., Production of chitooligosaccharides and their potential applications in medicine. *Marine Drugs* 2010, 8, 1482–1517.
- [8] Crini, G., Badot, P.-M., Application of chitosan, a natural aminopolysaccharide, for dye removal from aqueous solutions by adsorption processes using batch studies: A review of recent literature. *Progress Polymer Science* 2008, 33, 399–447.
- [9] Kong, M., Chen, X. G., Xing, K., Park, H. J., Antimicrobial properties of chitosan and mode of action: A state of the art review. *Int. J. Food Microbiol.* 2010, 144, 51–63.
- [10] Hackman, R., Studies on chitin I. Enzymic degradation of chitin and chitin esters. *Aust. J. Biol. Sci.* 1954, 7, 168–178.
- [11] Vårum, K. M., Ottøy, M. H., Smidsrød, O., Acid hydrolysis of chitosans. *Carbohydrate Polymers* 2001, 46, 89–98.
- [12] Monreal, J., Reese, E. T., The chitinase of *Serratia marcescens*. *Can. J. Microbiol.* 1969, 15, 689–696.
- [13] Vaaje-Kolstad, G., Horn, S. J., Sorlie, M., Eijsink, V. G., The chitinolytic machinery of *Serratia marcescens*—a model system for enzymatic degradation of recalcitrant polysaccharides. *FEBS J.* 2013, 280, 3028–3049.
- [14] Ueda, K., Ishikawa, S., Itami, T., Asai, T., Studies on the aerobic mesophilic cellulose-decomposing bacteria. 5-2. Taxonomical study on genus *Pseudomonas*. *J. Agric. Chem. Soc. Jpn* 1952, 26, 35–41.
- [15] DeBoy, R. T., Mongodin, E. F., Fouts, D. E., Tailford, L. E. et al., Insights into plant cell wall degradation from the genome sequence of the soil bacterium *Cellvibrio japonicus*. *J. Bacteriol.* 2008, 190, 5455–5463.
- [16] Gardner, J. G., Keating, D. H., Requirement of the type II secretion system for utilization of cellulosic substrates by *Cellvibrio japonicus*. *Appl. Environ. Microbiol.* 2010, 76, 5079–5087.
- [17] Gardner, J. G., Crouch, L., Labourel, A., Forsberg, Z. et al., Systems biology defines the biological significance of redox-active proteins during cellulose degradation in an aerobic bacterium. *Mol. Microbiol.* 2014, 94, 1121–1133.
- [18] Forsberg, Z., Nelson, C. E., Dalhus, B., Mekasha, S. et al., Structural and functional analysis of a lytic polysaccharide monooxygenase important for efficient utilization of chitin in *Cellvibrio japonicus*. *J. Biol. Chem.* 2016, 291, 7300–7312.
- [19] Tanaka, T., Fukui, T., Fujiwara, S., Atomi, H., Imanaka, T., Concerted action of diacetylchitobiose deacetylase and exo-beta-D-glucosaminidase in a novel chitinolytic pathway in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J. Biol. Chem.* 2004, 279, 30021–30027.
- [20] Hunt, D. E., Gevers, D., Vahora, N. M., Polz, M. F., Conservation of the chitin utilization pathway in the *Vibrionaceae*. *Appl. Environ. Microbiol.* 2008, 74, 44–51.
- [21] Cox, J., Mann, M., MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 2008, 26, 1367–1372.
- [22] Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A. et al., Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* 2011, 10, 1794–1805.
- [23] Consortium, T. U., UniProt: a hub for protein information. *Nucleic Acids Res.* 2015, 43, D204–D212.
- [24] Hamre, A. G., Lorentzen, S. B., Våljamäe, P., Sørli, M., Enzyme processivity changes with the extent of recalcitrant

- polysaccharide degradation. *FEBS Lett.* 2014, *588*, 4620–4624.
- [25] Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., Mann, M., In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols* 2006, *1*, 2856–2860.
- [26] Cottrell, J. S., London, U., Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999, *20*, 3551–3567.
- [27] Watanabe, T., Kimura, K., Sumiya, T., Nikaidou, N. et al., Genetic analysis of the chitinase system of *Serratia marcescens* 2170. *J. Bacteriol.* 1997, *179*, 7111–7117.
- [28] Suzuki, K., Suzuki, M., Taiyoji, M., Nikaidou, N., Watanabe, T., Chitin binding protein (CBP21) in the culture supernatant of *Serratia marcescens* 2170. *Biosc. Biotechnol. Biochem.* 1998, *62*, 128–135.
- [29] Juncker, A. S., Willenbrock, H., Von Heijne, G., Brunak, S. et al., Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci.* 2003, *12*, 1652–1662.
- [30] Bagos, P. G., Nikolaou, E. P., Liakopoulos, T. D., Tsirigos, K. D., Combined prediction of Tat and Sec signal peptides with hidden Markov models. *Bioinformatics* 2010, *26*, 2811–2817.
- [31] Bendtsen, J., Jensen, L., Blom, N., von Heijne, G., Brunak, S., Feature based prediction of non-classical protein secretion. *Protein Eng. Des. Sel.* 2004, *17*, 349–356.
- [32] Adav, S. S., Cheow, E. S. H., Ravindran, A., Dutta, B., Sze, S. K., Label free quantitative proteomic analysis of secretome by *Thermobifida fusca* on different lignocellulosic biomass. *J. Proteomics* 2012, *75*, 3694–3706.
- [33] Hirose, I., Sano, K., Shioda, I., Kumano, M. et al., Proteome analysis of *Bacillus subtilis* extracellular proteins: a two-dimensional protein electrophoretic study. *Microbiology* 2000, *146*, 65–75.
- [34] Siljamaki, P., Varmanen, P., Kankainen, M., Sukura, A. et al., Comparative exoprotein profiling of different *Staphylococcus epidermidis* strains reveals potential link between non-classical protein export and virulence. *J. Proteome Res.* 2014, *13*, 3249–3261.
- [35] Bendtsen, J., Kiemer, L., Fausboll, A., Brunak, S., Non-classical protein secretion in bacteria. *BMC Microbiol.* 2005, *5*, 58.
- [36] Götz, F., Yu, W., Dube, L., Prax, M., Ebner, P., Excretion of cytosolic proteins (ECP) in bacteria. *Int. J. Med. Microbiol.* 2015, *305*, 230–237.
- [37] Kulp, A., Kuehn, M. J., Biological functions and biogenesis of secreted bacterial outer membrane vesicles. *Annu. Rev. Microbiol.* 2010, *64*, 163–184.
- [38] Tjalsma, H., Antelmann, H., Jongbloed, J. D., Braun, P. G. et al., Proteomics of protein secretion by *Bacillus subtilis*: separating the “secrets” of the secretome. *Microbiol. Mol. Biol. Rev.* 2004, *68*, 207–233.
- [39] Wang, G., Chen, H., Xia, Y., Cui, J. et al., How are the non-classically secreted bacterial proteins released into the extracellular milieu? *Curr. Microbiol.* 2013, *67*, 688–695.
- [40] Jeffery, C. J., Moonlighting proteins. *Trends Biochem. Sci.* 1999, *24*, 8–11.
- [41] Blackwell, J., Structure of β -chitin or parallel chain systems of poly- β -(1 \rightarrow 4)-N-acetyl-D-glucosamine. *Biopolymers* 1969, *7*, 281–298.
- [42] Minke, R., Blackwell, J., The structure of α -chitin. *J. Mol. Biol.* 1978, *120*, 167–181.
- [43] Bolam, D. N., Ciruela, A., McQueen-Mason, S., Simpson, P. et al., *Pseudomonas* cellulose-binding domains mediate their effects by increasing enzyme substrate proximity. *Biochem. J.* 1998, *331*, 775–781.
- [44] Akagi, K.-i., Watanabe, J., Hara, M., Kezuka, Y. et al., Identification of the substrate interaction region of the chitin-binding domain of *Streptomyces griseus* chitinase C. *J. Biochem.* 2006, *139*, 483–493.
- [45] Ferreira, L. M., Hazlewood, G. P., Barker, P. J., Gilbert, H. J., The cellodextrinase from *Pseudomonas fluorescens* subsp. *cellulosa* consists of multiple functional domains. *Biochem. J.* 1991, *279* (Pt 3), 793–799.
- [46] Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., Henrissat, B., The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 2014, *42*, D490–D495.
- [47] Mitchell, A., Chang, H. Y., Daugherty, L., Fraser, M. et al., The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res.* 2015, *43*, D213–D221.
- [48] Rawlings, N. D., Waller, M., Barrett, A. J., Bateman, A., MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 2014, *42*, 503–509.
- [49] Tschumi, A., Grau, T., Albrecht, D., Rezwani, M. et al., Functional analyses of mycobacterial lipoprotein diacylglycerol transferase and comparative secretome analysis of a mycobacterial lgt mutant. *J. Bacteriol.* 2012, *194*, 3938–3949.
- [50] Fazzini, R. A. B., Levican, G., Parada, P., *Acidithiobacillus thiooxidans* secretome containing a newly described lipoprotein Licanantase enhances chalcopyrite bioleaching rate. *Appl. Microbiol. Biotechnol.* 2011, *89*, 771–780.
- [51] Sonnhammer, E. L., Von Heijne, G., Krogh, A., A hidden Markov model for predicting transmembrane helices in protein sequences, in: Glasgow, J., Major, T. L. F., Lathrop, R., Sankoff, D., and Sensen, C. (Eds.), *Int. Conf. Intell. Syst. Mol. Biol.*, AAAI Press, Motreal 1998, 175–182.
- [52] Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 2001, *305*, 567–580.
- [53] Vizcaino, J. A., Côté, R. G., Csordas, A., Dianes, J. A. et al., The PRoteomics IDentifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* 2013, *41*, D1063–D1069.

Proteomic investigation of the secretome of *Cellvibrio japonicus* during growth on chitin

Tina Rise Tuveng, Magnus Øverlie Arntzen¹, Oskar Bengtsson¹, Jeffrey G. Gardner², Gustav Vaaje-Kolstad¹, Vincent G.H. Eijsink¹

¹Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, 1432 Aas, Norway

²Department of Biological Sciences, University of Maryland – Baltimore County, Baltimore, MD, USA

Supporting information

Figures and tables

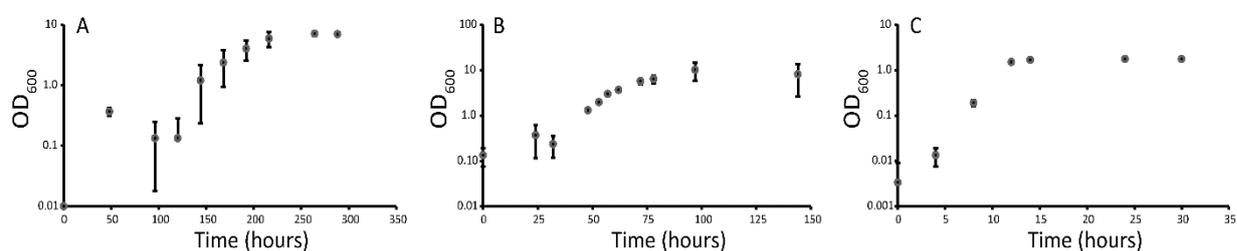


Figure S1. Growth curves. The figure shows growth curves for *C. japonicus* on different carbon sources in shaking flask cultures. The points represent the average of three biological replicates with error bars representing the standard deviation. A) α -chitin, B) β -chitin, and C) glucose.

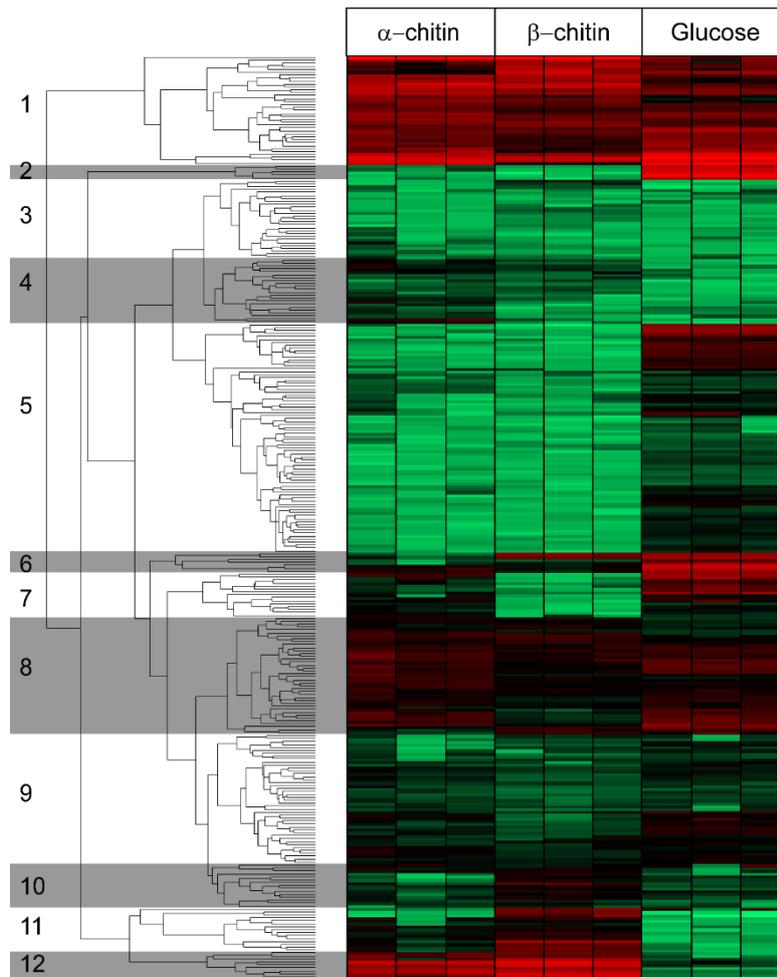


Figure S2. Heat map of secreted proteins. These proteins are predicted as secreted, including non-classical secretion, and divided into 12 clusters. CAZymes are not included (see main text). The colors in the heat map indicate protein intensity, ranging from high (red color, MaxLFQ 5×10^{10}) to low intensity (green color, MaxLFQ 1×10^6). Cluster-wise intensity profiles for each protein is shown in Figure 3, and details of the proteins found in clusters 11 and 12 are provided in Table 1.

Table S1. Numbers of detected proteins in various samples. The numbers (#) of proteins and fraction (%) of the total number of proteins (average of three biological replicates) predicted to be secreted by the Type II secretion system (Spl, SpII, and TAT), predicted to be secreted in a non-classical (NC) manner, predicted to contain a TMH, or predicted as cytosolic (Cyt) are shown. “Total secreted” equals the sum of Spl, SpII, TAT and NC. For the three different substrates, three time points are shown for both shake flask cultures and plates. The predicted numbers and fractions for the complete proteome of *C. japonicus*, are also shown.

Carbon source	Time (hours)	Flasks						Plates						Total predicted proteome	
		96		144		192		96		144		192		#	%
α-chitin	Total	#	%	#	%	#	%	#	%	#	%	#	%	#	%
	Cyt	130	48	570	54	703	55	51	24	75	21	118	26	2200	59
	TMH	7	3	47	4	84	7	6	3	9	3	17	4	431	12
	Spl	63	24	202	19	240	19	105	48	180	51	204	46	505	14
	SpII	19	7	73	7	84	7	17	8	31	9	37	8	143	4
	NC	48	18	149	14	152	12	32	15	48	14	65	14	405	11
	TAT	2	1	9	1	12	1	5	2	8	2	11	2	27	1
	TS	132	49	433	41	488	38	159	73	267	76	316	70	1080	29
β-chitin	Total	#	%	#	%	#	%	#	%	#	%	#	%	#	%
	Cyt	252	55	836	59	798	58	62	20	103	26	232	43	2200	59
	TMH	10	2	63	4	79	6	8	3	9	2	10	2	431	12
	Spl	83	18	266	19	257	19	156	51	189	48	189	35	505	14
	SpII	26	6	80	6	81	6	24	8	34	9	34	6	143	4

	NC	86	19	163	11	140	10	51	17	52	13	67	12	405	11
	TAT	3	1	10	1	11	1	8	2	9	2	9	2	27	1
	TS	198	43	519	37	490	36	239	77	284	72	300	55	1080	29
	Time (hours)	8		12		24		10		24		30			
Glucose	Total	#	%	#	%	#	%	#	%	#	%	#	%	#	%
		128		259		262		161		412		580		3711	
	Cyt	52	41	106	41	122	47	36	22	89	22	196	34	2200	59
	TMH	3	3	24	9	23	9	9	6	19	5	22	4	431	12
	Spl	42	33	73	28	66	25	72	45	198	48	229	40	505	14
	SpII	7	5	2	1	3	1	15	10	40	10	46	8	143	4
	NC	23	18	54	21	47	18	25	15	57	14	76	13	405	11
	TAT	1	1	1	0	2	1	4	2	9	2	10	2	27	1
TS	73	57	129	50	117	45	116	72	304	74	361	62	1080	29	

Table S2 is available from the publisher's website through the following link:

<http://onlinelibrary.wiley.com/doi/10.1002/pmic.201500419/full>

Table S3. Presence of putative chitin-active enzymes at two time points. The table shows the presence of assumed chitin active enzymes (indicated by a X) at the first and the second time point of sampling (Table S1) for the three carbon sources, α -chitin (α), β -chitin (β) and glucose (g).

Time point	Substrate	Chi18A	Chi18B	Chi18C	Chi18D	Chi19	Pda4D	LPMO 10A	LPMO 10B	Hex20B
		B3PBW8	B3PB16	B3PCT3	B3PLJ1	B3PB24	B3PDS4	B3PJ79	B3PDT6	B3PH90
1	α	X		X	X			X	X	
	β	X	X	X	X	X	X	X	X	
	g									
2	α	X	X	X	X	X	X	X	X	
	β	X	X	X	X	X	X	X	X	X
	g		X	X	x	X	X	X	X	

Table S4. MEROPS annotation of secreted proteins. List of assumed secreted proteins with possible peptidase activity for all carbon sources. The values given under the various carbon sources are log₁₀ of LFQ intensities; they are the average of three biological replicates if not stated otherwise.

Uniprot ID	Protein name	Mol. weight [kDa]	α-chitin	β-chitin	Glucose	LipoP	Cazy	SecP score	MEROPS
B3PLG6	MucD	49.6	8.5	8.1	8.2	Spl		0.65	unassigned
B3PKN6	D-alanyl-D-alanine carboxypeptidase family	44.3		7.8	8.2	Spl		0.18	unassigned
B3PIU5	Endo-1,4-beta-xylanase B	33.9	8.6	7.6 ^a		Spl		0.91	unassigned
B3PI85	Trypsin domain protein	50.5	7.7	8.4	8.4	Spl		0.92	unassigned
B3PI00	Lipase	36.5	8.6	8.8	7.9	Spl		0.77	unassigned
B3PH79	Carbohydrate binding protein, putative, cbp6B	48.3	8.3	8.2		Spl	CBM6	0.90	unassigned
B3PGW9	Beta-lactamase	59.9			8.9	CYT		0.94	unassigned
B3PGD9	Peptidase family M48 family (EC 3.4.-.-)	45.8			8.3	SplI		0.80	unassigned
B3PCZ3	Peptidase, M23/M37 family	30.9	8.2	7.7	8.7	Spl		0.22	unassigned
B3PCW9	Peptidase family M48 family (EC 3.4.-.-)	55.1	7.8 ^a	7.9 ^a	8.2	Spl		0.12	unassigned
B3PC44	Beta-lactamase	43.4			7.9	Spl		0.36	unassigned
B3PBC3	Uncharacterized protein	22.8			7.7 ^a	Spl		0.10	unassigned
B3PD75	Peptidase family S41B family	54.9	8.0	8.5	7.8	CYT		0.95	S41.012
B3PFM0	Phospholipase/carboxylesterase (EC 3.1.-.-)	73.8	8.5	8.4	8.6	Spl		0.56	S09.A77
B3PJ05	Uncharacterized protein	37.5		7.6		Spl		0.77	S01.472
B3PGV7	Peptidase, M23/M37 family	46.2	7.7		8.1	Spl		0.25	M23.950
B3PL78	Peptidase, M23/M37 family	52.0	8.7	7.1	9.5	Spl	CBM50	0.78	M23.009
B3PDW3	Peptidase, M16 (Pitriylisin) family (EC 3.4.24.-)	107.6	7.6 ^a	7.6	9.1	SplI		0.58	M16.001
B3PI83	YbcL	20.1	8.5	8.1	8.2	Spl		0.92	I51.003

^a Average of 2 biological replicates

Proteomic investigation of the secretome of *Cellvibrio japonicus* during growth on chitin

Tina Rise Tuveng, Magnus Øverlie Arntzen¹, Oskar Bengtsson¹, Jeffrey G. Gardner², Gustav Vaaje-Kolstad¹, Vincent G.H. Eijssink¹

¹Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, 1432 Aas, Norway

²Department of Biological Sciences, University of Maryland – Baltimore County, Baltimore, MD, USA

Supporting information

Materials and methods

S1.1 Mass spectrometry details

Samples were loaded onto a trap column (Acclaim PepMap100, C₁₈, 5 µm, 100 Å, 300 µm i.d. x 5 mm, Thermo Scientific) and backflushed onto a 50 cm analytical column (Acclaim PepMap RSLC C₁₈, 2 µm, 100 Å, 75 µm i.d., Thermo Scientific). At the start, the columns were in 96 % solution A [0.1 % (v/v) formic acid], 4% solution B [80 % (v/v) ACN, 0.1 % (v/v) formic acid]. Peptides were eluted using a 120 min gradient developing from 4 to 12 % (v/v) solution B in 2 minutes, 12 to 43 % (v/v) in 93 minutes and finally to 60% B in 6 minutes before the wash phase at 90 % B, all at a flow rate of 300 nL/min. In order to isolate and fragment the 10 most intense peptide precursor ions at any given time throughout the chromatographic elution, the Q-Exactive mass spectrometer was operated in data-dependent mode to switch automatically between orbitrap-MS and higher-energy collisional dissociation (HCD) orbitrap-MS/MS acquisition. The selected precursor ions were then excluded for repeated fragmentation for 20 seconds. The resolution was set to R=70,000 and R=35,000 for MS and MS/MS, respectively. For optimal acquisition of MS/MS spectra, automatic-gain-control target values were set to 1,000,000 charges and a maximum injection time of 128 ms.

S1.2 Data analysis

During analysis in MaxQuant protein N-terminal acetylation, oxidation of methionine, conversion of glutamine to pyro glutamic acid, and deamination of asparagine and glutamine were used as variable modifications, while carbamidomethylation of cysteine residues was used as a fixed modification.

Trypsin was used as digestion enzyme and two missed cleavages were allowed. To increase the number of identified peptides the 'match between runs' feature of MaxQuant was enabled with default parameters, in order to transfer identifications between samples based on accurate mass and retention time [1]. The settings were such that transfer of peptides was only allowed between samples from the same carbon source. All identifications were filtered in order to achieve a protein false discovery rate (FDR) of 1%.

The protein group file from MaxQuant was loaded in Perseus (version 1.5.1.6), with LFQ intensities columns as expressions and protein names, and ID (Uniprot code) columns as text. We reduced the matrix by removing proteins categorized as only identified by site, reverse, or as a contaminant. We used both unique and razor peptides for quantification and at least two ratio counts were required for a quantification to be considered valid. Furthermore, we required the protein to be quantified in at least two of the three replicates on at least one substrate. LFQ intensities were log₁₀ transformed and missing values were replaced by imputation from a normal distribution (width of 0.2 and downshifted 1.9 standard deviations from the original distribution) in total matrix mode. We performed manual inspection of histograms to ensure that the imputed values did not affect the normality of the distributions and that the values were located in the lower end of the original histograms, i.e. close to the detection limit of the mass spectrometer. Hierarchical clustering and heat map generation were done with Euclidean distance measure and average linkage. We then manually selected protein clusters based on similarity of expression patterns.

To predict sub-cellular localization, we used the LipoP 1.0 server (<http://www.cbs.dtu.dk/services/LipoP/>) [2], and the PRED-TAT software (<http://www.compgen.org/tools/PRED-TAT/>) [3]. Proteins were considered as secreted if predicted to be cleaved by signal peptidase I (Spi) or signal peptidase II (SpII) by the LipoP server, or if predicted to have a twin-arginine signal peptide by the PRED-TAT software. Proteins annotated as cytosolic (CYT) by the LipoP server were analyzed using the SecretomeP 2.0 server (<http://www.cbs.dtu.dk/services/SecretomeP/>) [4] to predict non-classically secreted proteins. To

be considered as non-classically secreted the score from SecretomeP must be 0.5 or higher. Proteins predicted to contain N-terminal trans membrane helices (TMH) by LipoP were also analyzed in the TMHMM 2.0 server (<http://www.cbs.dtu.dk/services/TMHMM/>) [5, 6]. Annotation of predicted activity on carbohydrates for the *C. japonicus* genome was downloaded from the CAZy database (<http://www.cazy.org/>) [7]. Prediction of peptidases was done using the MEROPS database (<http://merops.sanger.ac.uk/index.shtml>) [8].

References

- [1] Cox, J., Hein, M. Y., Lubner, C. A., Paron, I., *et al.*, Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & Cellular Proteomics* 2014, *13*, 2513-2526.
- [2] Juncker, A. S., Willenbrock, H., Von Heijne, G., Brunak, S., *et al.*, Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein science* 2003, *12*, 1652-1662.
- [3] Bagos, P. G., Nikolaou, E. P., Liakopoulos, T. D., Tsirigos, K. D., Combined prediction of Tat and Sec signal peptides with hidden Markov models. *Bioinformatics* 2010, *26*, 2811-2817.
- [4] Bendtsen, J., Jensen, L., Blom, N., von Heijne, G., Brunak, S., Feature based prediction of non-classical protein secretion. *Protein Eng Des Sel* 2004, *17*, 349 - 356.
- [5] Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 2001, *305*, 567 - 580.
- [6] Sonnhammer, E. L., Von Heijne, G., Krogh, A., A hidden Markov model for predicting transmembrane helices in protein sequences. in: Glasgow, J., Littlejohn, T., Major, F., Lathrop, R., Sankoff, D. and Sensen, C. (Ed.), *Int. Conf. Intell. Syst. Mol. Biol.*, pp. 175-182, Motreal, Canada 1998, AAAI Press.
- [7] Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., Henrissat, B., The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research* 2014, *42*, D490-495.
- [8] Rawlings, N. D., Waller, M., Barrett, A. J., Bateman, A., MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research* 2013, *42*, D490-D49.

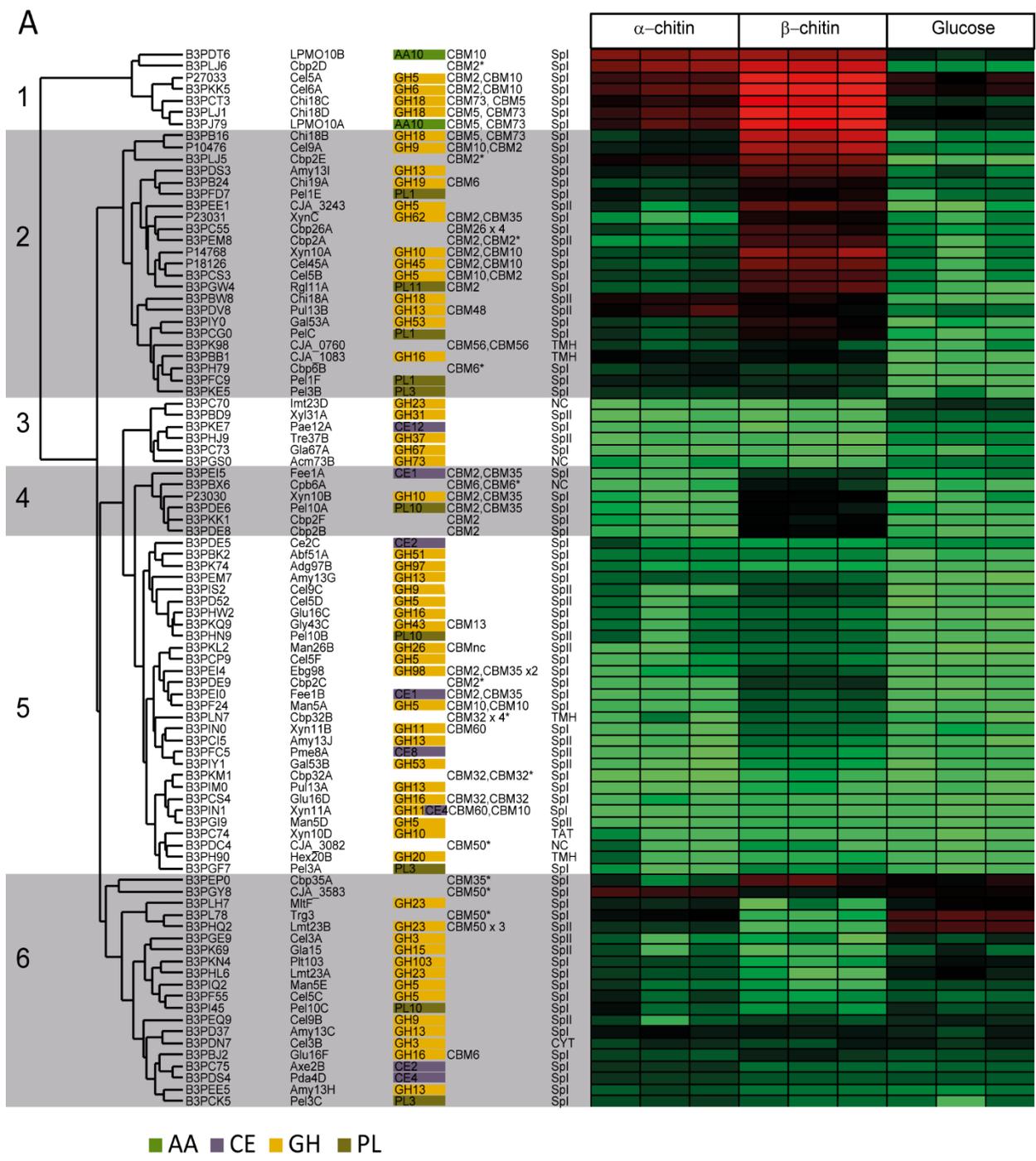


Figure 2A corrected. See original manuscript for figure legend.

Paper III

Chitin degradation by *Cellvibrio japonicus* is dependent on the non-redundant *CjChi18D* chitinase.

Monge, E., Tuveng, T. R., Vaaje-Kolstad, G., Eijsink, V. G. H. & Gardner, J. G. 2017.
Manuscript.

47 **ABSTRACT**

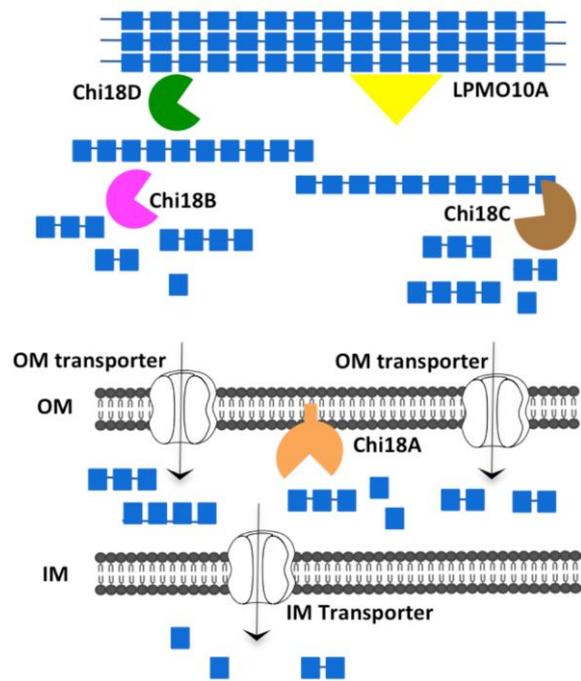
48 Understanding the strategies used by bacteria to degrade polysaccharides constitutes an
49 invaluable tool for biotechnological applications. Bacteria are major mediators of
50 polysaccharide degradation in nature, however the complex mechanisms used to detect,
51 degrade, and consume these substrates are not well understood, especially for recalcitrant
52 polysaccharides such as chitin. It has been previously shown that the model bacterial
53 saprophyte *Cellvibrio japonicus* is able to catabolize chitin, but little is known about the
54 enzymatic machinery underlying this capability. Previous analysis of the *C. japonicus*
55 genome and proteome indicated the presence of four family 18 Glycoside Hydrolase (GH18)
56 enzymes, and studies of the proteome indicated that all are involved in chitin utilization. Using
57 a combination of *in vitro* and *in vivo* approaches, we have studied the roles of these four
58 chitinases in chitin conversion. Genetic analyses showed that only the *CjChi18D* enzyme is
59 essential for the degradation of chitin substrates. Biochemical characterization of the four
60 enzymes showed functional differences and synergistic effects during chitin degradation,
61 indicating non-redundant functionalities. In accordance with the genetic data, *CjChi18D* was
62 the most efficient chitin degrader. Transcriptomic studies revealed complex regulation of the
63 chitin degradation machinery of *C. japonicus* and confirmed the importance of *CjChi18D* and
64 a previously characterized chitin-active LPMO, *CjLPMO10A*. This combination of *in vitro*
65 and *in vivo* approaches provides enhanced understanding of the initial stages of chitin
66 degradation by bacteria.

67 **INTRODUCTION**

68 Chitin, a linear polymer of $\beta(1-4)$ -linked *N*-acetylglucosamine, is the second most abundant
69 polysaccharide on earth after cellulose and a major source of fixed carbon and nitrogen,
70 especially in the oceans (1). The rate of chitin degradation has been calculated to match the
71 rate of synthesis resulting in little net accumulation of chitin in the environment, suggesting
72 that the microorganisms that derive nutrition from this polysaccharide are able to degrade it
73 with high efficiency (2). Moreover, chitin is increasingly being recognized as an important
74 feedstock for the production of renewable chemicals, and chitooligosaccharides are being
75 investigated for several biomedical applications (3-6). As a consequence, the strategies used
76 by microbes for efficient chitin bioconversion are gaining interest, as current chemical
77 methods used for industrial processing of chitin are inefficient and wasteful (7, 8). The
78 combination of environmental importance and industrial/medical relevance has renewed
79 interest in the bioconversion of chitin (9, 10).

80 Depolymerization of chitin-rich biomasses has been studied predominately from
81 biochemical and structural perspectives, and as a consequence there is considerable
82 understanding on certain mechanistic aspects of enzymatic degradation (11). Chitin
83 deconstruction shares many features with cellulose degradation, which is not surprising given
84 that chitin has a crystalline structure similar to that of cellulose and differs chemically by an
85 *N*-acetyl substitution at the C2 carbon (12). Specifically, chitin depolymerization is achieved
86 through a concerted effort of endo- and exo-acting enzymes to reduce polymer length to short
87 oligosaccharides that are then converted to *N*-acetylglucosamine monomers by *N*-
88 acetylhexosaminidases for entry into cellular metabolism [Fig 1, (6)]. The action of lytic
89 polysaccharide monooxygenases in aerobic microbes promotes chitin degradation by acting
90 on crystalline regions of the substrate (13-15). It is worth noting that chitin and fragments
91 thereof may be enzymatically deacetylated by carbohydrate esterases and that some
92 microorganisms use chitin conversion strategies that include deacetylation of chitin-derived
93 chitobiose (16, 17). Partially deacetylated polymeric chitin is called chitosan, which may be
94 broken down by most chitinases [e.g. (18)], by chitosanases and specialized *N*-
95 acetylhexosaminidases.

96



97

98 **Fig. 1. Proposed model for chitin utilization in *Cellvibrio japonicus*.** *Cj*Chi18D (green semicircle)
 99 and *Cj*LPMO10A (yellow triangle) work together to disrupt the crystalline structure of chitin, while
 100 *Cj*Chi18B (magenta semicircle) and *Cj*Chi18C (brown semicircle) are acting on the more accessible
 101 chitin fibers to produce chitooligosaccharides which are taken up into the periplasm. *Cj*Chi18A helps
 102 generating GlcNAc and this lipoprotein may be acting in the periplasmic space (as shown here) or
 103 may be located on the outer side of the outer membrane and act outside the cell (see text for more
 104 discussion).

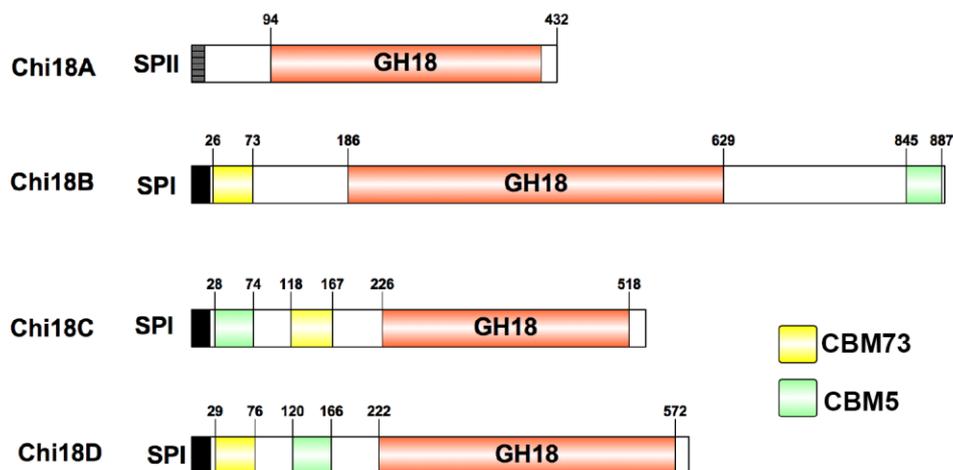
105

106 Physiological and genetic aspects of chitin degradation, especially for terrestrial
 107 bacteria, are not well characterized. One reason for this lack of knowledge is the multiplicity
 108 of chitinolytic enzymes in many, but not all, chitin-degrading bacteria, which makes
 109 functional analysis of the individual enzymes challenging. While several model organisms
 110 have been used to study chitin degradation (19-21), one that is emerging as a powerful model,
 111 due the available systems biology tools, is the saprophytic Gram-negative bacterium
 112 *Cellvibrio japonicus* (22, 23). This bacterium is a potent chitin degrader (13), and has a
 113 predicted suite of nine genes encoding proteins possibly involved in chitin degradation (24).
 114 Specifically, the *C. japonicus* genome is predicted to encode four GH18 chitinases, one GH19
 115 chitinase, two *N*-acetylhexosaminidases (GH20), one chitosanase (GH46), and one chitin-
 116 specific LPMO [AA10 (24)]. Secretome analysis (25) has shown that the four GH18s and, to
 117 a lesser extent, the GH19, are upregulated during growth on chitin. Generally, the GH18s and
 118 LPMOs are considered key enzymes for conversion of crystalline chitin. The role of GH19s

119 is not clear, however, we cannot rule out a role in chitin metabolism (26, 27).

120 The possession of large numbers of Carbohydrate Active Enzymes (CAZymes) (28)
 121 is a hallmark feature of *C. japonicus*. There is increasing evidence that the CAZymes of *C.*
 122 *japonicus* that belong to the same GH family are not functionally redundant, but have unique
 123 physiological functions (14, 29). In the current study, the physiological role of the four *C.*
 124 *japonicus* GH18 chitinases was determined focusing on the initial stages of chitin
 125 degradation. Sequence analysis indicates that *CjChi18B*, *CjChi18C*, and *CjChi18D* are
 126 secreted enzymes containing two carbohydrate-binding modules (CBMs), whereas *CjChi18A*
 127 is a single domain protein that likely is membrane anchored (Fig. 2; see below). Through
 128 RNAseq analysis we determined that these four chitinases were highly up-regulated while
 129 chitin was being used as a carbon source. Combinatorial mutational analyses determined that
 130 only *CjChi18D* is essential for chitin degradation. Further, biochemical characterization of
 131 the catalytic domains of these four GH18 chitinases showed that *CjChi18D* is substantially
 132 more active towards insoluble chitin than the other GH18 chitinases, underpinning its
 133 importance in chitin degradation. The functional insight into the initial stages of chitin
 134 degradation by saprophytic bacteria that is provided in this study has the potential to
 135 accelerate the development of industrial bioconversion processes for chitins.

136



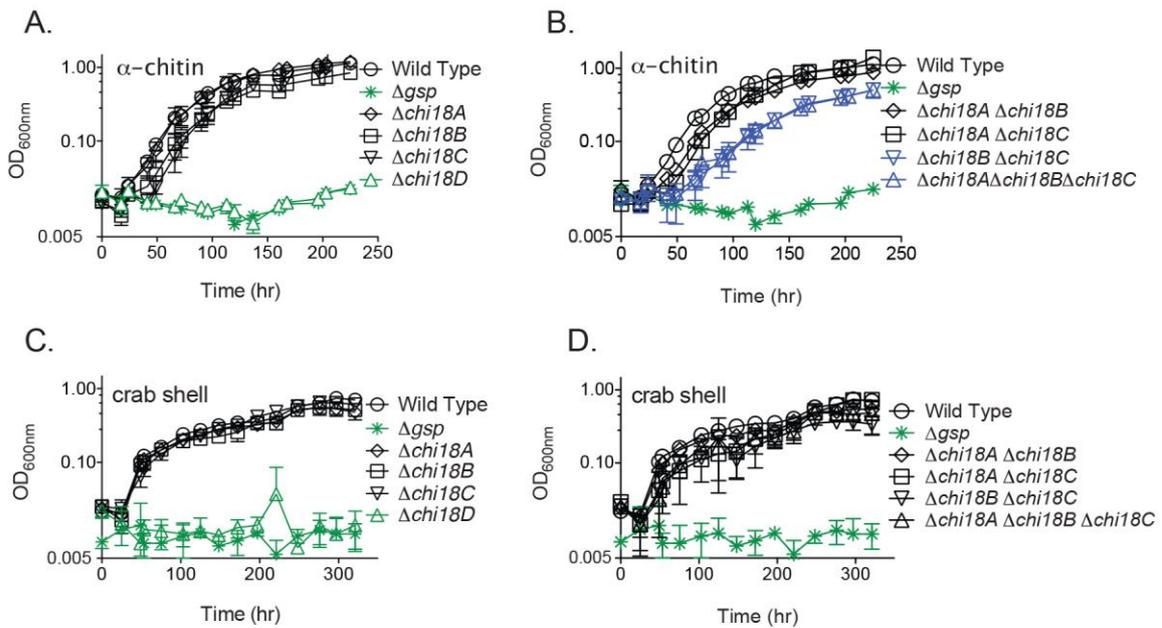
137

138 **Fig. 2. Diverse Architecture of the Family GH18 chitinases of *C. japonicus*.** CAZy domain
 139 representation of the family GH18 chitinases of *C. japonicus*. The indicated domains are as follows:
 140 GH18, family GH18 catalytic domain; CBM73, carbohydrate-binding module family 73 domain;
 141 CBM5, carbohydrate-binding module family 5 domain; SPI, signal peptide, type I; SPII, signal
 142 peptide, type II.

143 **RESULTS**

144 ***CjChi18D* is essential for the degradation of α -chitin.** According to previous sequence
145 analysis (24), *C. japonicus* possess four GH18 chitinases (*CjChi18A*, *CjChi18B*, *CjChi18C*
146 and *CjChi18D*). A recent proteomic study (25) detected the four enzymes in the secretome of
147 *C. japonicus* growing on α - and β -chitin, which suggested the importance of these enzymes
148 for efficient chitin degradation. To elucidate the physiological relevance of these enzymes,
149 we generated a suite of GH18 deletion mutants and assessed their fitness on insoluble chitin
150 substrates, including an environmentally relevant substrate, crab shell.

151 Wild type and GH18 deletions mutant strains all grew well in defined media with
152 either glucose (Glc) or *N*-acetylglucosamine (GlcNAc) as the sole source of carbon (Fig S1).
153 A Δ *gsp* strain deficient in the Type II Secretion System that is thought to drive secretion of
154 the chitinolytic machinery (see below) was also able to grow on these sugars, as was
155 previously shown (13, 30). When the single deletion mutants were grown using insoluble α -
156 chitin or crab shell as the sole carbon source distinct phenotypes emerged (Fig 3A & 3C).
157 When α -chitin was the sole source of carbon, the Δ *chi18D* mutant was unable to grow. The
158 other three single deletion strains (Δ *chi18A*, Δ *chi18B* and Δ *chi18C*) had similar growth rates
159 and maximum optical densities (OD) as the wild type (Table S1A). Interestingly, the Δ *chi18B*
160 and Δ *chi18C* single mutant strains had substantially more protracted lag phases than the wild
161 type strain when using α -chitin as carbon source (Fig 3A). When using crab shells as the only
162 source of carbon, the Δ *chi18D* mutant was also unable to grow while the single mutants of
163 Δ *chi18A*, Δ *chi18B* and Δ *chi18C* displayed growth similar to wild type (Fig 3C & Table S1B).
164



165

166

Fig. 3. Growth of *C. japonicus* mutants on chitin. Family GH18 deletion mutants were grown on MOPS minimal medium supplemented with 0.25 % α-chitin (A-B) or 4 % crab shell (C-D). Panels A and C show single mutants; panels B and D show multiple mutants. All experiments were performed in biological triplicates; error bars represent standard deviations. These growth experiments were performed simultaneously, but are separated into multiple panels for clarity. As a consequence, the control strains (wild type and Δgsp) are repeated in each panel.

172

173

174

175

176

177

178

179

180

181

182

183

184

185

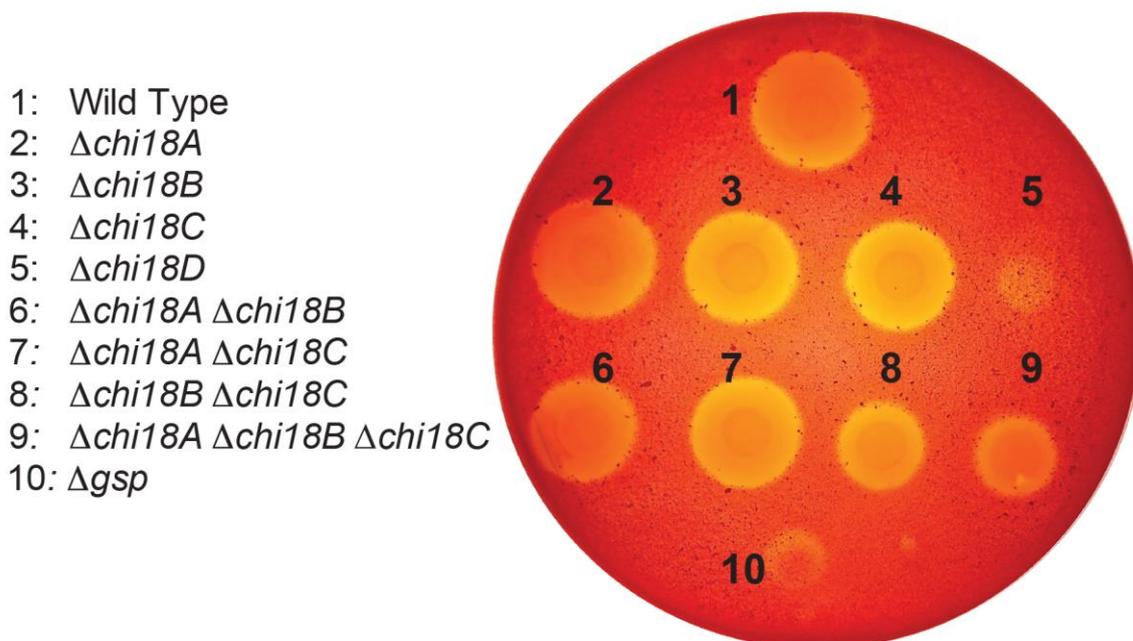
186

To test for synergy between GH18 chitinases, we generated all combinations of double deletion mutants, and then assessed growth using α-chitin or crab shells as the sole carbon source. The double deletion Δchi18B Δchi18C resulted in a slower growth rate and a longer lag phase when grown on α-chitin in comparison to either of the single deletion mutants or the wild type (Fig 3B). The growth rate of the Δchi18B Δchi18C double mutant was reduced 37 % compared to wild type (Table S1A). The Δchi18A Δchi18B double mutant grew like the single deletion mutant of chi18B while the Δchi18A Δchi18C double mutant grew like the single deletion mutant of chi18C. The triple mutant of Δchi18A Δchi18B Δchi18C recapitulated the growth defect observed in the double mutant of Δchi18B Δchi18C in terms of the growth rate, lag phase and the maximum OD (Fig 3B & Table S1A).

Interestingly, the double mutants exhibited wild type-like phenotypes when grown using crab shells as the sole carbon source (Fig 3D). The growth rates of the Δchi18B and Δchi18C single mutants and the Δchi18B Δchi18C double mutant were similar to the wild type on crab shells. It is worth noting that for all tested strains, the growth rate on the crab

187 shell substrate was reduced compared to α -chitin. The results of experiments with β -chitin
 188 (Fig. S2) were very similar to the results obtained with α -chitin.

189 ***CjChi18D* is a potent secreted chitinase.** Sequence alignment of the catalytic domains of the
 190 four chitinases, generally showed large sequence variation. *CjChi18A* and *CjChi18D* share
 191 higher sequence identity (32 %), while *CjChi18C* and *CjChi18D* show the lowest identity (19
 192 %). As predicted by LipoP (31) and SignalP (32) software tools, all four of the *C. japonicus*
 193 GH18 chitinases have a signal sequence, however *CjChi18A* has a SPaseII-cleaveable
 194 sequence and is predicted to be an outer membrane associated lipoprotein. The *CjChi18B*,
 195 *CjChi18C* and *CjChi18D* enzymes have a SPaseI-cleaved sequence (31) and are predicted to
 196 be secreted. The domain structure of the four GH18 chitinases of *C. japonicus* is summarized
 197 in Fig 2. The LipoP/SignalP software predictions are in alignment with a proteomics study
 198 that showed the occurrence of all four GH18 chitinases in the secretome (25). To further
 199 assess the contribution of individual GH18 enzymes as effectors of chitin degradation, we
 200 used our suite of GH18 mutants to assess secreted activity using colloidal chitin plate assays
 201 (Fig 4).



202
 203 **Fig. 4. Growth of *C. japonicus* mutants on chitin-containing plates.** The strains were grown on a
 204 plate that contained MOPS with 1.5 % agar, 2 % colloidal chitin and 0.2 % glucose. After incubation
 205 at 30°C for 5 days, the plates were stained with Congo Red. This experiment was conducted in
 206 triplicate

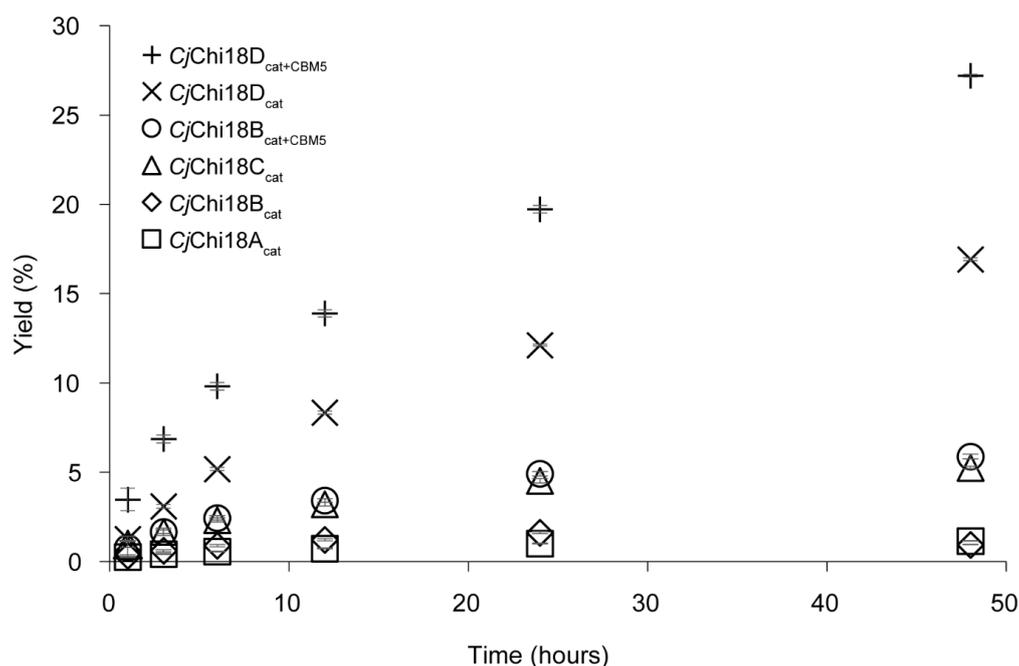
207

208 The wild type strain generated a robust zone of clearing while a Δgsp secretion-
209 deficient mutant generated no zone of clearing (Table S2). We observed a similarly striking
210 phenotype with the $\Delta chi18D$ single mutant strain, which also generated no zone of clearing.
211 The remaining GH18 single mutants displayed approximately wild type zones of clearing.
212 Interestingly, the $\Delta chi18B \Delta chi18C$ double mutant had a zone of clearing that was ~50 %
213 smaller than the wild type. The $\Delta chi18A \Delta chi18B \Delta chi18C$ triple mutant displayed a zone of
214 clearing similar to the $\Delta chi18B \Delta chi18C$ double mutant. The reduced chitin-degrading
215 capacities of the $\Delta chi18B \Delta chi18C$ double mutant and the $\Delta chi18A \Delta chi18B \Delta chi18C$ triple
216 mutant are in agreement with the observed growth defects of these strains when grown on α -
217 chitin (*vide supra*).

218 **The GH18 chitinases have different activities towards chitin and (GlcNAc)₆.** To further
219 investigate the features of the four GH18 chitinases, they were cloned, expressed and purified
220 for biochemical characterization. Despite massive efforts, soluble full length multi-domain
221 chitinases *CjChi18B*, *CjChi18C*, and *CjChi18D* could not be obtained and comparative
222 biochemical analysis was therefore primarily conducted with overexpressed catalytic
223 domains. All four chitinases were able to degrade α -chitin, although with greatly varying
224 efficiency that was clearly highest for *CjChi18D* (Fig 5). As expected (33), versions of
225 *CjChi18B* and *CjChi18D* containing their CBM5 gave higher yields compared to their
226 respective catalytic domains. The catalytic domain of *CjChi18C* gave higher yields than the
227 catalytic domains of *CjChi18A* and *CjChi18B*, but all were poor in chitin degradation
228 compared to *CjChi18D*.

229

230



231

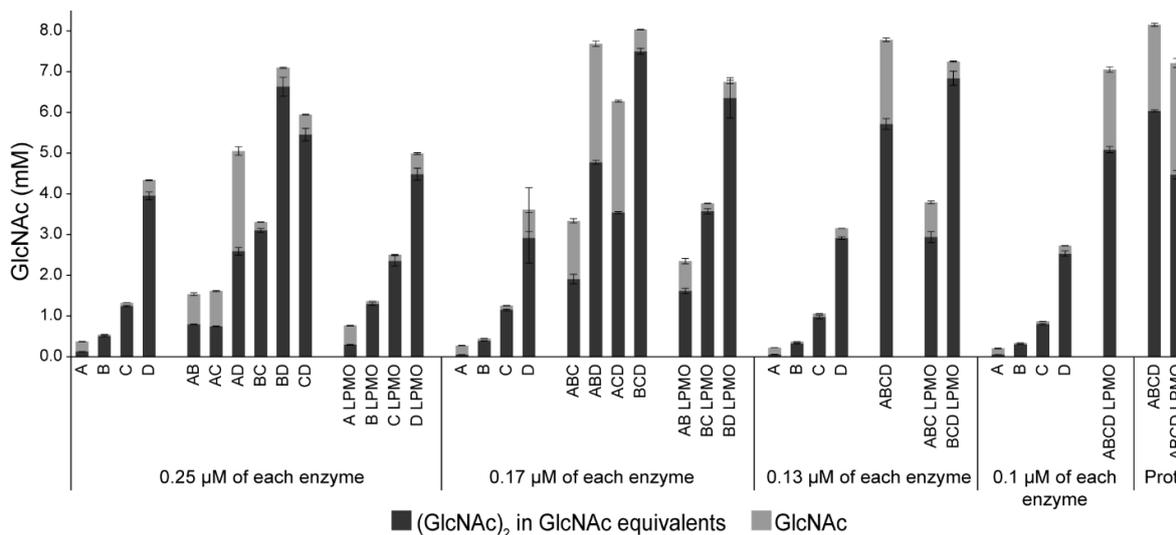
232 **Figure 5. Degradation of α -chitin.** Degradation of α -chitin (15 g/L) at 30 °C was tested in 20 mM
 233 BisTris pH 6.5, 0.1 mg/ml BSA. The enzyme concentration was 0.5 μ M and samples were taken at
 234 different time points. The yield refers to the degree of chitin solubilization. The α -chitin used contained
 235 6.43 % ash and 5.42 % moist and this was taken into account when calculating yields. Each reaction
 236 was performed in triplicates; standard deviations are shown as error bars, but are difficult to see since
 237 they were low and are partly covered by the symbols.

238

239 Analysis of products generated during degradation of α -chitin (Figure S3) showed that
 240 *CjChi18A*_{cat} initially produces both GlcNAc and (GlcNAc)₂. The product profiles for the
 241 other variants (*CjChi18B*_{cat}, *CjChi18B*_{cat+CBM5}, *CjChi18C*_{cat}, *CjChi18D*_{cat}, and
 242 *CjChi18D*_{cat+CBM5}) showed mainly (GlcNAc)₂ and smaller amounts of GlcNAc (Fig S3) as is
 243 usual for GH18 chitinases. Surprisingly, the product profile of *CjChi18B*_{cat} was different from
 244 *CjChi18B*_{cat+CBM5} after 48h, where GlcNAc was the dominating product for *CjChi18B*_{cat},
 245 while (GlcNAc)₂ was the dominating product for *CjChi18B*_{cat+CBM5} (Fig S3). For *CjChi18A*_{cat},
 246 only GlcNAc was detected after 48 hours, indicating an *N*-acetylhexosaminidase-like activity
 247 for this enzyme.

248 Synergy experiments (Fig. 6) confirmed the dominating role of *CjChi18D* in chitin
 249 conversion and revealed synergistic effects for various enzyme combinations, such as when
 250 mixing *CjChi18A*_{cat} with *CjChi18B*_{cat} or when mixing *CjChi18B*_{cat} with *CjChi18D*_{cat}. These
 251 synergistic effects indicate that the various chitinases must have different functionalities. As
 252 expected on the basis of the data presented above, the presence of *CjChi18A*_{cat} shifted the

253 product profile towards a higher GlcNAc/(GlcNAc)₂ ratio. The effect of adding *Cj*LPMO10A
 254 was generally small, but seemed slightly larger when combined with the individual catalytic
 255 domains of *Cj*Chi18A, *Cj*Chi18B, and *Cj*Chi18C, compared to *Cj*Chi18D. While *Cj*Chi18D
 256 alone was clearly the most powerful individual enzyme, the highest product formations were
 257 obtained upon combining *Cj*Chi18D with at least two of the other chitinases (Fig. 6).
 258



259

260 **Figure 6. Synergy experiments.** The catalytic domains of all chitinases and the *Cj*LPMO10A (13)
 261 were mixed in different ways to investigate possible synergistic effects in reactions with α -chitin (15
 262 g/L). The total enzyme load was 0.5 μ M with equal amount of each enzyme (concentrations given in
 263 the figure), except in the reactions where the ratio was determined by protein quantification data from
 264 a previous proteomics study [“Prot”; (25)]. In the “Prot” reaction without *Cj*LPMO10A the ratio was;
 265 21% *Cj*Chi18A, 4 % *Cj*Chi18B, 22 % *Cj*Chi18C, and 52 % *Cj*Chi18D. In the “Prot” reaction with
 266 *Cj*LPMO10A the ratio was: 14 % *Cj*Chi18A, 3 % *Cj*Chi18B, 15 % *Cj*Chi18C, 35 % *Cj*Chi18D, and
 267 33 % *Cj*LPMO10A. The *Cj*LPMO10A was Cu²⁺-saturated before use. Reaction mixtures were
 268 incubated at 30 °C in 20 mM BisTris pH 6.5, 0.1 mg/ml BSA, and samples were taken after 24 hours.
 269 In reactions with *Cj*LPMO10A, 0.5 mM ascorbate was added as external electron donor. Production
 270 of GlcNAc and (GlcNAc)₂ was quantified and (GlcNAc)₂ is given in GlcNAc equivalents. Three
 271 parallel reactions were done for each condition and standard deviations are shown as error bars.
 272 Reaction mixtures that contained the LPMO showed minor amounts of oxidized (GlcNAc)₂, which
 273 were not quantified.

274

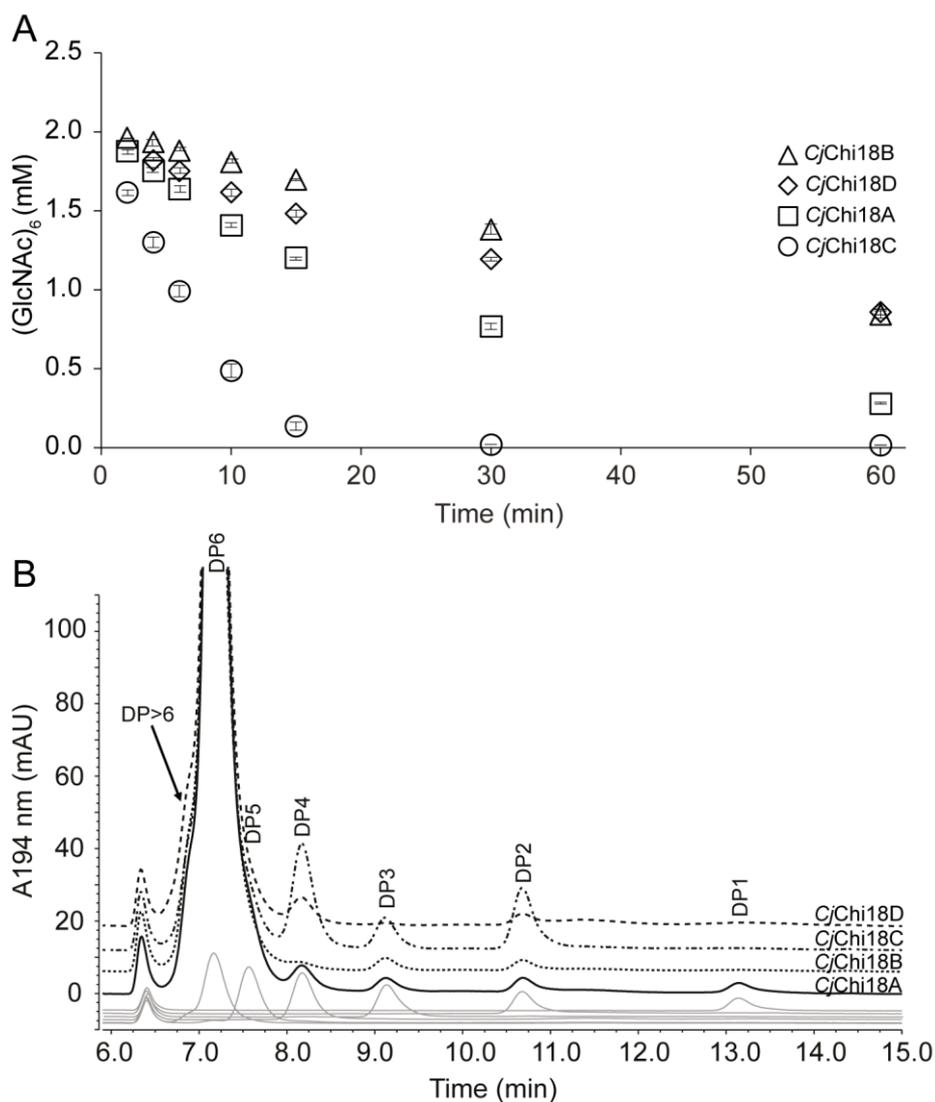
275 We have previously performed a quantitative proteomics study of *C. japonicus* grown
 276 on chitin (25) and hypothesized that the quantified levels of the four GH18 chitinases and
 277 *Cj*LPMO10A could represent an optimal ratio between the enzymes. The use of this ratio
 278 indeed resulted in high product formation, but these were not higher than the product

279 formations obtained with some of the other efficient enzyme combinations (Fig 6).
280 Interestingly, under the assay conditions used here, replacing chitinases with the LPMO led
281 to reduced solubilization.

282 To further look into the substrate specificity of the chitinases, their ability to degrade
283 (GlcNAc)₆ was investigated (Fig 7A). The results showed that *CjChi18C_{cat}* had the highest
284 activity against (GlcNAc)₆ with an initial rate of disappearance of the substrate of 300 ± 8
285 min^{-1} . *CjChi18A_{cat}* had an initial rate of $118 \pm 1 \text{ min}^{-1}$, while *CjChi18D_{cat}* and *CjChi18B_{cat}* had
286 initial rates of 44.5 ± 2.6 and $27.0 \pm 0.8 \text{ min}^{-1}$, respectively. It is worth noting that the ranking
287 of the apparent enzyme efficiencies is strongly substrate-dependent.

288 The product profile obtained for the enzymes shortly after initiation of the (GlcNAc)₆
289 hydrolysis reactions (2 minutes reaction time) show striking differences between the
290 chitinases (Fig 7B). *CjChi18A_{cat}* yielded all possible product types [(GlcNAc)₁₋₅], where
291 GlcNAc was the dominating product ($97.8 \pm 7.3 \mu\text{M}$) followed by (GlcNAc)₂ ($62.3 \pm 2.5 \mu\text{M}$),
292 (GlcNAc)₃ ($47.7 \pm 1.7 \mu\text{M}$) and (GlcNAc)₄ ($33.9 \pm 1.5 \mu\text{M}$) (quantification of (GlcNAc)₅ was
293 not possible). *CjChi18B_{cat}* produced mainly (GlcNAc)₂ and (GlcNAc)₃, with minor amounts
294 of (GlcNAc)₄. *CjChi18C_{cat}* produced (GlcNAc)₂₋₄, while *CjChi18D_{cat}* produced mainly
295 (GlcNAc)₂ and (GlcNAc)₄. Chitooligosaccharides longer than (GlcNAc)₆ were observed in
296 the *CjChi18D_{cat}* reaction (Figs 7B and S4), indicating that this enzyme has transglycosylating
297 activity. Reactions using (GlcNAc)₂ as substrate showed rapid conversion by *CjChi18A_{cat}*,
298 whereas only trace amounts of monomer were detected for the other three chitinase (results
299 not shown).

300



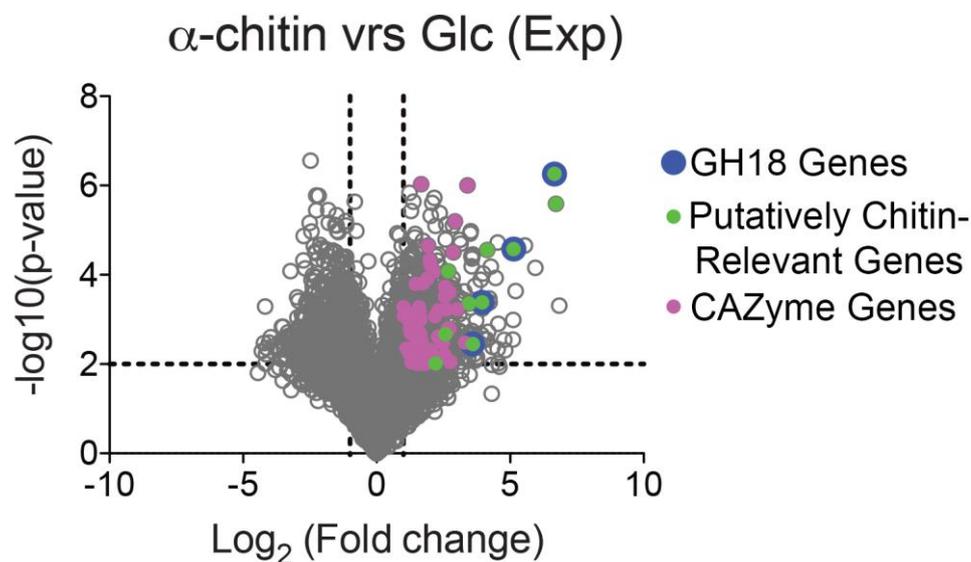
301

302 **Figure 7. Hydrolysis of (GlcNAc)₆.** (A): Hydrolysis of (GlcNAc)₆ over time. The slopes of the linear
 303 parts of these curves were used to calculate the initial rates. Standard deviations are shown as error
 304 bars. (B): Chromatograms showing the product profile obtained 2 minutes after mixing chitinases with
 305 substrate. DP1-6 represent (GlcNAc)₁₋₆. Chromatograms for various standards are shown as grey lines
 306 at the bottom. These experiments were done with the catalytic domains of the chitinases. The reactions
 307 contained 2 mM (GlcNAc)₆, 10 mM BisTris pH 6.5, 0.1 mg/ml BSA and 50 nM enzyme, and were
 308 done in triplicates..

309

310 ***C. japonicus* GH18 chitinase genes are highly expressed during chitin utilization.** To
 311 obtain further insight into the regulation of chitinase expression and to complement the
 312 previous secretome study (25), we used RNAseq to determine changes in gene expression
 313 during growth on α -chitin consumption, relative to growth on glucose. Analysis of samples
 314 from the exponential growth phase revealed significant up-regulation of 73 CAZymes (Table
 315 S3A). The seven most strongly upregulated genes were all related to chitin conversion,

316 including the LPMO (6.72 log₂ fold change), a hexosaminidase (4.14 log₂ fold change), the
 317 *nag9A* gene (3.47 log₂ fold change) belonging to a a predicted GlcNAc utilization operon
 318 (*gluP*, CJA_1162, *nag9A*, *nagB*), and the four GH18 chitinases in the following order
 319 *chi18D* > *chi18C* > *chi18A* > *chi18B*, the log₂ fold change being 6.66, 5.13, 3.95 and 3.60,
 320 respectively (Fig 8; Table S3A). Several other genes putatively involved in chitin conversion
 321 were up-regulated, albeit to a lesser extent, including the GH19 chitinase (2.20 log₂ fold
 322 change), the other GH20 hexosaminidase (log₂ fold changes of 2.69 and 4.14 for Hex20A and
 323 Hex20B, respectively), the GH46 chitosanase and two polysaccharide deacetylases (log₂ fold
 324 changes 2.57, 1.60 and 1.41, respectively).
 325



326
 327 **Fig. 8. Differential expression of family GH18 chitinases of *C. japonicus* during exponential**
 328 **growth.** This volcano plot shows the log₂(fold change) plotted against the -log₁₀(p-value) of all
 329 expressed genes in *C. japonicus* during exponential growth on glucose compared to α-chitin in which
 330 each gray circle represents the expression of a gene. The black dashed lines indicate the significance
 331 cut-off values: -log₁₀(p-value)>2 and log₂(fold change)>1.
 332

333 Comparison of samples taken at early stationary phase yielded similar results (Table
 334 S3B; Figure S5A). This analysis revealed the up-regulation of 47 CAZymes, of which seven
 335 are implicated in chitin degradation (*lpmo10A*, *chi18D*, *chi18C*, *chi18B*, *nag9A*, *hex20A*, and
 336 *chi19A*). While the expression data, in particular for the exponential comparison, showed a
 337 strong upregulation of chitin-relevant genes, there was also up-regulation of a wide variety of
 338 CAZyme genes associated with the degradation of other polysaccharides such as starch,
 339 xylan, cellulose, pectin, arabinanan, mannan, β-glucan and xyloglucan. It thus seems that the

340 regulation of the chitin response is not as specific as the recently mapped response to xylan
341 (J.G. Gardner et al., unpublished observations), but more closely resembles the general
342 response observed for cellulose (14). When comparing the transcriptomes of the exponential
343 and the stationary phase during growth on chitin, none of the significant changes in expression
344 concerned CAZymes (Figure S5B).

345 **DISCUSSION**

346 A previous report demonstrated that *C. japonicus* can grow using both purified chitin
347 and unrefined crab shells as a sole nutrient source (13). Additionally, a recent proteomic study
348 found several secreted GH18 chitinases, suggesting a robust response to chitin (25). It has
349 remained unknown, however, how these GH18 chitinases contribute to the degradation of
350 chitin, and to what extent the four GH18 genes in the *C. japonicus* genome are equivalent. In
351 this study, through an implementation of transcriptomic, genetic, and biochemical
352 approaches, we assessed the physiological function of the four GH18 enzymes to degrade
353 chitin. The synthesis of *in vitro*, *in vivo*, and *in silico* data provides a comprehensive
354 understanding of the initial stages of chitin degradation by *C. japonicus*.

355 **The architecture of the family GH18 chitinases indicates differential biological roles.**

356 Analysis using LipoP (31) predicts that the *CjChi18A* enzyme is a lipoprotein with a signal
357 sequence cleaved by signal peptidase II (SPII). The presence of a glutamine in position +2
358 after the cleavage site of SPII further suggests that *CjChi18A* resides attached to the outer
359 membrane (34-37). It is likely that *CjChi18A* functions in degrading short
360 chitooligosaccharides, as the absence of CBMs may limit the association of this enzyme with
361 polymeric chitin (33, 38). Such a function is supported by the biochemical data for *CjChi18A*
362 showing high activity on soluble substrates and an ability to produce monomers.

363 Analysis of the *CjChi18B*, *CjChi18C* and *CjChi18D* enzymes indicated a signal
364 peptide that is cleaved by signal peptidase I (31). The growth defect on chitin observed in the
365 Δgsp mutant suggest that these three enzymes are transported to the extracellular space by the
366 type II secretion system (39). The *CjChi18B*, *CjChi18C* and *CjChi18D* enzymes all possess
367 a single CBM5 and a CBM73 domain (40), however the orientations of these domains differ
368 between the enzymes (Fig 2), which may have functional consequences. As found with many
369 CAZymes, the *CjChi18B*, *CjChi18C* and *CjChi18D* chitinases contain serine rich regions
370 separating the individual domains, which likely act as flexible linkers (41).

371 **Functional characterization of family GH18 enzymes of *Cellvibrio japonicus* reveals** 372 **divergent functionalities and the importance of *CjChi18D*.**

373 The growth data and secretion
374 assay strongly pointed to the *chi18D* gene product as essential for the degradation of chitinous
375 substrates (Figs 3, 4, & Fig S2). These results reinforce two features previously described for
376 *C. japonicus*, that secreted enzymes are essential for recalcitrant polysaccharide degradation,
377 and that single CAZymes elicit major physiological effects (30, 42). Other individual enzymes
378 did not seem essential for growth but the double deletion mutant $\Delta chi18B \Delta chi18C$ did show
reduced growth and reduced chitin conversion efficiency (Figs 3 & 4), indicating that these

379 two chitinases have partly overlapping functionalities that contribute to, but are not essential
380 for, chitin conversion.

381 Biochemical characterization of the catalytic domains revealed clear functional
382 differences (Figs 5, 7, S4, & Table S4) and also showed that the enzymes act synergistically
383 during the degradation of chitin (Fig 6). This indicates that the enzymes have different roles
384 during chitin conversion. Enzyme synergy during the course of chitin degradation has been
385 described previous for *Serratia marcescens*, and is generally ascribed to collaboration
386 between endo- and exo-acting enzymes, where endo-acting enzymes generate chain ends that
387 are substrates for exo-acting enzymes that usually show processive action [(43); see below
388 for further discussion].

389 The absence of growth defects of the multiple mutants while growing using crab shells
390 is likely a consequence of the slow overall growth rate, combined with some ability to use
391 protein in the crab shells as a nutrient source [the protein content in crab shells is usually in
392 the range of 20-40 % (8, 44)]. Still, also in this case *CjChi18D* was essential for growth,
393 indicating that chitin degradation is essential, either because the bacterium needs the sugar or
394 because degradation of chitin fibers in the crab shells is important in order for the bacterium
395 to access the protein fraction. It is well known that the chitin fibers in crustaceans are covered
396 with proteins [and vice versa; (45)]. Thus, the activity of *CjChi18D* may be important for
397 disentangling the protein for degradation by secreted proteases. In this context, it is worth
398 noting that the proteomics study on *C. japonicus* published by Tuveng et al. (25) led to
399 identification of several potential proteases that were upregulated during growth on chitin.
400 Inspection of the RNAseq data shows that one of these predicted proteases (CJA_0276, Table
401 S3A) is significantly upregulated when *C. japonicus* grows on chitin. The CJA_0276 protein
402 is predicted with an unassigned protease domain by MEROPS (46) and a CBM6 that could
403 aid in binding to chitin-rich substrates.

404 **The different roles of the GH18 chitinases in chitin hydrolysis.** Chitinolytic systems
405 hitherto described in the literature, e.g. those of *S. marcescens* (21) and *Vibrio* species (19,
406 47), usually contain non-redundant enzyme activities that play different roles in the
407 depolymerization of the recalcitrant chitin substrate. Analysis of the domain organization of
408 the *C. japonicus* chitinases, functional data, and homology modelling of their 3D structures
409 (Fig. S6) all indicate that this chitinolytic system also represents a suite of complementary
410 activities. Available data from known processive chitinases show that there are some
411 hallmarks that are specific for processive enzymes, this involves a deep substrate binding cleft
412 created by an $\alpha+\beta$ domain and other loops protruding from this domain or other parts of the

413 catalytic domain (48, 49), and conserved aromatic amino acids lining the catalytic cleft (50,
414 51). Structural modelling of the *C. japonicus* GH18 chitinases (Fig S6) and sequence
415 alignment with known processive chitinases (Fig S7), show that *CjChi18B* and *CjChi18D*
416 have these specifications. *CjChi18B* has a large $\alpha+\beta$ domain (~110 amino acids) and an
417 “tunnel-like” substrate-binding cleft, while *CjChi18D* has a much smaller $\alpha+\beta$ domain (~60
418 amino acids) and a more open active site cleft. *CjChi18A* has the $\alpha+\beta$ domain (~65 amino
419 acids), but lacks the aromatic residues, *CjChi18C* does not have any of the features
420 characteristic for processivity and is more similar to the non-processive exo-chitinase *SmChiC*
421 from *S. marcescens*, displaying an open substrate binding cleft that is characteristic for non-
422 processive glycoside hydrolases (38).

423 Although it is difficult to determine the degree of processivity from standard substrate
424 degradation experiments (see Horn et al. (52) for discussion), such experiments do provide
425 indications. During degradation of crystalline chitin, processive enzymes tend to have high
426 (GlcNAc)₂/GlcNAc ratios (18, 51). Furthermore, during degradation of (GlcNAc)₆ processive
427 enzymes will yield a high (GlcNAc)₂/(GlcNAc)₄, since the hexamer is expected to yield three
428 dimers (18). Of the GH18 chitinases of *C. japonicus*, only *CjChi18B* showed such high ratios
429 (Table S4), indicating that this enzyme is the most processive, as also suggested by the
430 structural model (Fig S6). The other three enzymes did not show clear signs of processivity.
431 *CjChi18C* displays all hallmarks of a non-processive enzyme, and *CjChi18A* is special in that
432 it primarily produces monomers, which is not really compatible with processivity considering
433 that the repeating unit in a chitin chain is a dimer. These latter two enzymes showed the
434 highest specific activities towards the hexamer, suggesting that they primarily act on soluble
435 substrates.

436 Like the processive *CjChi18B*, *CjChi18D* was not especially active towards soluble
437 substrates. The structural model of this enzyme shows some features indicative of
438 processivity, but the biochemical data do not support this mechanism. Notably, the
439 transglycosylating properties of *CjChi18D* may cause problems when interpreting the product
440 profiles obtained after chitin and (GlcNAc)₆ degradation. It remains uncertain why *CjChi18D*
441 is the superior when it comes to degrading crystalline chitin. It is worth noting that the
442 structural model of this enzyme suggest a “hybrid” character. The active site cleft is deep, but
443 not closed, as in other processive enzymes and there are aromatic residues that, while perhaps
444 not conferring a large degree of processivity, will contribute to interacting productively with
445 the crystalline material (53-55).

446 *CjChi18A* deviates strongly from the other three GH18 chitinases. Firstly, the signal
447 peptide indicates that this enzyme is immobilized at the outer membrane, similar to *FjChiB*
448 of *Flavobacterium johnsoniae* (56) and *SdChiB* from *Saccharophagus degradans* (57, 58).
449 Secondly, *CjChi18A* has an *N*-acetylhexosaminidase-like activity similar to that of GH20
450 chitobiasis. The biological role of *CjChi18A* is unclear. If the enzyme is facing the periplasm,
451 its role seems redundant as *C. japonicus* encodes two GH20 chitobiasis, which are
452 upregulated during growth on chitin (Table S3). Degradation of (GlcNAc)₂ into GlcNAc
453 monomers in the extracellular space could, on the other hand, be relevant as the presence of
454 GlcNAc might induce important cellular signaling processes and/or facilitate sugar uptake.
455 Another plausible role of having *CjChi18A* facing the outside of the cell could be in
456 controlling the (GlcNAc)₂ concentration to avoid product inhibition of the other chitinases.

457 It should be noted that we were not able to produce the full-length proteins, which
458 implies that the biochemical data do not give the full picture when it comes to enzyme
459 functionality. The full-length enzymes could give different results, especially during chitin
460 degradation, since CBMs are known to aid in the (challenging) interaction between the
461 enzymes and crystalline substrates (33). It is thought that the positioning of the CBMs
462 indicates the directionality of a chitinase [an N-terminally positioned CBM indicates enzyme
463 movement from the reducing towards the non-reducing end (49, 51, 59)] and it may be that
464 crystalline substrates are more easily depolymerized in a specific direction. Indeed, *SmChiA*,
465 having an N-terminal FnIII domain and acting from the reducing end towards the non-
466 reducing end, has been shown to move substantially faster than *SmChiB* when degrading
467 crystalline chitin (59). This could be due to a CBM-driven differences in directionality, but
468 may also be related to the topological features of the substrate binding cleft. *CjChi18D* and
469 *CjChi18C* have their CBMs on the N-terminal side of the catalytic domain, while *CjChi18B*
470 has one CBM on each side of the catalytic domain (Fig 2). While these differences in the
471 CBM organization likely affect enzyme functionality, it is reassuring that the biochemical
472 data derived for studying the catalytic domains align rather well with the mutant data and with
473 the functional inferences made from the structural models of the chitinases.

474 On the basis of the present data, it is difficult to judge whether the four enzymes are
475 endo- or exo-acting and it is worth noting that mixed endo-/exo-modes occur and possibly are
476 common (60). *CjChi18B* clearly looks exo-processive, while *CjChi18C* clearly looks endo-
477 non-processive. The situation for *CjChi18A* and *CjChi18D* is less clear; both enzymes show
478 relatively open active site clefts and are likely to have considerable endo-character.

479 The expression data show that the most up-regulated chitin-relevant enzyme of them
480 all is the chitin-specific LPMO, *CjLPMO10A*. It has been shown that this enzyme will act
481 synergistically with chitinases (13) but in the present study, the effects of adding the LPMO
482 to the chitinases was limited (Fig 6). This is likely an effect of the substrate. The substrate
483 used in this study was balled milled into fine particles, and it has previously been shown that
484 such ball milling reduces the crystallinity of the substrate, hence reducing the activity of the
485 LPMO (61).

486 **The chitinolytic machinery of *C. japonicus* is highly responsive to the presence of chitin**
487 **substrates.** There was significant up-regulation of the predicted *C. japonicus* genes encoding
488 chitin-active proteins (GH18, GH19, GH20, GH46 and AA10 families) both during
489 exponential and stationary phase (Table S3). The up-regulation of a number of other CAZyme
490 genes not involved in chitin degradation was striking, but not surprising. *C. japonicus* has
491 been shown to have both substrate sensing- and growth rate-dependent control of CAZyme
492 gene expression (14). In regards to the substrate-specific response, *C. japonicus* seems to have
493 two variations, one that is general (where diverse CAZyme genes are up-regulated), and one
494 that is specific (where only substrate-specific CAZymes genes are up-regulated). The
495 response observed for cellulose is an example of the former (14), and the response for xylan
496 is an example of the latter (J.G. Gardner et al., unpublished observations). The regulatory
497 network for the chitinolytic response appears to be similar to that for the cellulose response,
498 which is not surprising given the similarities of the substrates.

499 **Final Remarks.** The proposed model (Fig 1) for crystalline chitin degradation by *C.*
500 *japonicus*, based on *in vitro*, *in vivo*, and *in silico* analyses is that the four GH18 chitinases
501 synergistically degrade chitin. *CjChi18D* releases chitooligosaccharides from the crystalline
502 material, perhaps by the combined endo-, exo- and mildly processive actions. The other
503 enzymes likely act on the more accessible chitin fibers converting them into dimers
504 (*CjChi18B* and *CjChi18C*) and smaller chitooligosaccharides, while *CjChi18A* converts
505 chitooligosaccharides to *N*-acetylglucosamine. Next to these chitinases, *CjLPMO10A* plays
506 a major role in chitin degradation, as indicated by previously published biochemical data (13)
507 and the transcriptomic data presented here.

508 EXPERIMENTAL PROCEDURES

509 **Growing conditions.** *Cellvibrio japonicus* strains were grown on MOPS (3-(N-
510 morpholino)propanesulfonic acid)) minimal media (62) containing 0.2 % wt/vol glucose, 0.5
511 % wt/vol *N*-acetylglucosamine, 0.25 % wt/vol α -chitin from shrimp shells (Sigma, Aldrich),
512 0.2 % wt/vol β -chitin from squid pen or 4 % wt/vol milled crab shells from *Callinectes*
513 *sapidus* as the sole source of carbon. The following protocols were developed to prepare the
514 chitin containing substrates. The coarse flakes of α -chitin were sieved through the top piece
515 of a 130 mm Buchner polypropylene filter to homogenize the particle size. β -chitin from squid
516 pen was manually ground in a ceramic mortar and pestle and then passed through a plastic
517 filter with 4 mm diameter holes. For the crab shells, any membranous material was discarded
518 and the shells were then thoroughly washed and rinsed. Clean and dry shells were manually
519 ground and then sieved through a 4 mm diameter filter. To remove any generated crab shell
520 dust, the filtrate was rinsed through a Buchner polypropylene filter. The shell pieces were
521 subjected to an extra round of autoclave sterilization. *Escherichia coli* strains were grown in
522 lysogenic broth (LB) (63). For growth analysis studies, strains were grown for 24 hours in 5
523 mL MOPS containing glucose. Then, an 18 mm test tube with 5 mL MOPS and a form of
524 chitin was inoculated with a 1:100 dilution of the prepared overnight cultures (23). The culture
525 time was dependent on the form of chitin used. All cultures were incubated at 30 °C with an
526 aeration of 225 RPM. For insoluble substrates, growth was measured as a function of the
527 optical density at 600 nm (OD₆₀₀) in a Spec20D+ spectrophotometer (Thermo Scientific) with
528 biomass containment as needed (64). In growth experiments with glucose or *N*-
529 acetylglucosamine growth was monitored using a Tecan M200Pro microplate reader (Tecan
530 Trading AG, Switzerland). All growth experiments were performed in biological triplicates.
531 Plate media were solidified with 1.5% agar. When required, kanamycin was used at a
532 concentration of 50 μ g/mL.

533 **Generation of Deletion Mutants.** Deletion mutants of the family GH18 chitinases of *C.*
534 *japonicus* were made and verified using previously published protocols (30, 42). A suicide
535 vector was generated by cloning \pm 500 bp up and downstream from the gene to be deleted into
536 the pk18mobsacB vector (65) at the EcoRI and the XbaI sites. The pk18mobsacB plasmids to
537 make the deletions of genes *chi18A*, *chi18B* and *chi18C* were synthesized by GeneWiz (South
538 Plainfield, NJ) while the plasmid to make the deletion of *chi18D* was amplified by PCR and
539 assembled via Gibson assembly (66). All vectors were chemically electroporated into *E.*
540 *coli* S17 λ _{PIR} strains. Through a tri-parental mating the deletion carrying plasmid was

541 conjugated into *C. japonicus* using an *E. coli* strain containing the plasmid pRK2013 (67).
542 Recombinant colonies were selected using MOPS/kanamycin plates. Then a counter selection
543 was carried out using MOPS/sucrose plates. The deletion mutants were confirmed by PCR.
544 Primers used for the construction and verification of the mutants are listed in Table S5.

545 **Visualization of Colloidal Chitin Degradation.** Colloidal chitin was prepared as described
546 in (68). Colloidal chitin plates were made using MOPS defined media supplemented with 2
547 % (w/v) colloidal chitin and 0.2 % (w/v) glucose. To assay for chitin degradation, 10 μ l of
548 overnight cultures of the *C. japonicus* strains to be analyzed were spotted onto the chitin plate
549 and incubated for 4 days at 30 °C. Then, plates were stained with a 0.1 % (w/v) Congo red
550 solution for 10 minutes followed by 10 minutes destaining with a 1 M NaCl solution, as
551 described previously for detection of degradation of carboxyl methyl cellulose (64).

552 **Transcriptomic Analysis.** A transcriptomic analysis was conducted for *C. japonicus* grown
553 on glucose or α -chitin. To prepare the samples, the protocol described in Gardner et al. (14,
554 30, 42) was followed. Briefly, 35 ml of cell culture were aliquoted into 50 ml conical tubes
555 with 5 ml of a stopping solution made with ethanol and saturated phenol (19:1). The cells
556 were pelleted by centrifugation at 8000 g for 5 min at 4 °C. The supernatant was decanted and
557 the pellets flash frozen in a dry ice/ethanol bath for 5 min before being stored at – 80 °C. For
558 each carbon source, samples were taken in triplicate and at two time points: the beginning of
559 the exponential phase ($0.1 > OD_{600} > 0.2$) and the beginning stationary phase. RNA extraction,
560 library preparation, and sequencing were performed by GeneWIZ (South Plainfield, NJ). For
561 this study, Illumina HiSeq2500 was performed in 50bp single-reads with at least 10 million
562 reads generated per sample. The raw sequence data generated was converted into FASTQ
563 files using Illumina CASSAVA 1.8.2, which were then imported into CLC Genomics
564 Workbench 7.5.1. Low quality base pairs were trimmed from the ends and the sequence reads
565 were mapped to the *C. japonicus* reference genome with RPKM values calculated for genes.
566 The \log_2 transformation and quantile normalization were performed for the RPKM values and
567 a Student's t-test conducted to compare gene expression between the glucose and the α -chitin.
568 An adjusted p-value > 0.01 and a \log_2 fold change > 1 were selected as significance cut-off
569 parameters.

570 **Bioinformatics Analysis.** We determined the predicted CAZy domains presented using the
571 Database for Automated Carbohydrate-active enzyme ANnotation (DbCAN)(40). Using
572 LipoP (31) and SignalP (32), we determined the putative location of the chitinolytic enzymes.
573 Three dimensional models of the GH18 chitinases were generated using PyMod 2.0 (69).

574 **Cloning and expression of chitinases.** Synthetic genes of the four full length chitinases,
575 optimized for expression in *Escherichia coli*, were purchased from GeneScript. These genes
576 encoded the catalytic domain for *CjChi18A* and the full length protein (without signal
577 peptide) for *CjChi18B*, *CjChi18C* and *CjChi18D*. Primers for amplification of the genes were
578 designed so that a 6xHis-tag followed by a TEV protease cleavage site was introduced at the
579 N-terminus of all proteins and a stop codon was introduced before the C-terminal His-tag
580 encoded in the vector. Using different primers for amplification, genes encoding different
581 versions (Table S6) of the modular proteins were generated, which were cloned into the pNIC-
582 CH vector (70) using ligation independent cloning (71). The DNA sequences of the genes
583 were confirmed by Sanger sequencing. Plasmids were transformed into *E. coli* BL21 cells
584 using heat shock. Protein expression was tested by inoculating 50 ml LB + 50 µg/ml
585 Kanamycin with 500 µl of an overnight culture. The cultures were grown at 37 °C and when
586 the OD reached 0.6, the cells were induced by adding IPTG to a final concentration of 0.2
587 mM. After growth overnight at 30 °C, the cells were harvested by centrifugation (6164 rpm,
588 12 min, 4 °C) and resuspended in 5 ml 20 mM Tris-HCl, 150 mM NaCl, 5 mM imidazole, pH
589 8.0. DNaseI and PMSF were added to final concentrations of 1.4 µg/ml and 0.1 mM,
590 respectively, before the cells were lysed by sonication (pulse 3 sec on, 3 sec off for 3-5 min),
591 using a VC750 VibraCell sonicator (Sonics & Materials, Inc, CT, USA). Following cell
592 disruption, the samples were centrifuged (11814 RCF, 12 min, 4 °C), after which the
593 supernatant was collected and filtrated (0.22 µm). Analysis by SDS-PAGE showed that only
594 *CjChi18A* was soluble. Except for the full length enzymes, which were not produced in
595 detectable levels, all other enzyme variants were produced in large amounts, but were
596 insoluble.

597 Production of *CjChi18A* was scaled up to a 500 ml culture following the same protocol
598 as above for production and harvesting. The filtrated supernatant containing *CjChi18A* was
599 used for protein purification by nickel affinity chromatography using a HisTrap HP 5 ml
600 column (GE Healthcare Life Sciences, Uppsala, Sweden) connected to an Äkta pure system
601 (GE Healthcare Life Sciences, Uppsala, Sweden). A continuous imidazole gradient ending at
602 300 mM imidazole was used to elute bound protein, using a flow rate of 1 ml/min.

603 *CjChi18B_{cat}*, *CjChi18B_{cat}+CBM5*, *CjChi18C_{cat}*, *CjChi18D_{cat}*, and *CjChi18D_{cat}+CBM5*
604 were produced using a denaturing and refolding method, starting with 500 ml cultures as
605 described above. After harvesting the cells (7025 RCF, 12 min, 4 °C), the pellet was
606 resuspended in 20 ml 50 mM Tris, pH 8.0, 0.2M NaCl and 50 % of the cell suspension was
607 used in the further steps. After another centrifugation (18459 RCF, 10 min, 4 °C) the pellet

608 was resuspended in 10 ml 50 mM Tris, pH 8.0, 0.2M NaCl and, after addition of lysozyme
609 (final concentration 200 μ g/ml) and DNaseI (final concentration 0.1 mM), the samples were
610 incubated on ice for 30 min. The samples were subsequently sonicated (pulse 3 sec on, 3 sec
611 off, 2-4 min) and inclusion bodies were harvested by centrifugation (18459 RCF, 10 min),
612 after which the protein pellet was resuspended in 25 ml washing buffer (20 mM Tris-HCl,
613 100 mM NaCl, 2% Triton X-100, pH 8.0). The centrifugation and subsequent resuspension
614 in washing buffer steps were repeated twice and after the final centrifugation, the pellet was
615 dissolved in 5 ml cold denaturing solution (50 mM Tris-HCl, pH 8.5, 100 mM NaCl, 5 M
616 Guanidine HCl, 1 mM EDTA, and 20 mM DTT), vortexed for 1 min, and incubated overnight
617 at 4 °C with slow rotation of the sample tubes. The samples were then centrifuged at 9000g,
618 for 10 min at 4 °C and the supernatant, containing the denatured protein, was collected for
619 refolding.

620 The sample containing the denatured protein (approximately 5 ml) was added to 250
621 ml cold refolding buffer (100 mM Tris-Cl, 0.4 M L-Arginine, 0.5 mM oxidized glutathione,
622 5 mM reduced Glutathione) at a rate of 1 ml/hour under intense stirring at 4 °C. After adding
623 all protein, the solution was stirred overnight at 4 °C, before centrifugation (9000 RCF, 10
624 min, 4 °C). The supernatant was collected and dialyzed (10 MWCO Snakeskin, Thermo
625 Fisher Scientific, MA, USA) against 2.5 L 50 mM Tris-HCl, 0.1M NaCl, pH 8.0 overnight.
626 The dialysis solution was changed once, after approximately 6 hours. After collecting the
627 dialyzed sample, the protein was purified with nickel affinity chromatography as described
628 above.

629 The catalytic domain of CjLPMO10A was expressed and purified as previously
630 described by Forsberg et al. (13).

631 **Activity assays.** Standard reactions contained 15 g/L α -chitin (extracted from *Pandalus*
632 *borealis*, Seagarden, Karmsund, Norway) or 2 mM chitooligosaccharides (MegaZyme, Bray,
633 Ireland), 0.1 mg/ml BSA and 20 mM BisTris, pH 6.5, unless stated otherwise. For chitin
634 degradation 0.5 μ M enzyme was used, while for chitooligosaccharide degradation 50 nM
635 enzyme was used, unless stated otherwise. Reaction mixtures were incubated in a
636 thermomixer at 30 °C, 800 rpm and enzyme activity was quenched by adding sulphuric acid
637 to a final concentration of 25 mM. A Rezex RFQ-Fast Acid H+ (8 %) ion-exclusion column
638 (Phenomenex, CA, USA) installed on a Dionex Ultimate 3000 HPLC with UV-detection
639 (194nm) was used to analyze and quantify degradation products as described by Mekasha et
640 al. (72). Reactions with (GlcNAc)₆ as substrate were analyzed using a Rezex ROA-organic
641 Acid H+ (8 %) ion exclusion column (Phenomenex, CA, USA) installed on a Dionex Ultimate

642 3000 HPLC, using a column temperature of 65 °C. An 8 µl sample was injected on the
643 column, and the mono/oligosaccharides were eluted isocratically at 0.6 ml/min with 5 mM
644 sulphuric acid as mobile phase. The chitooligosaccharides were monitored by measuring the
645 absorbance at 194 nm. Standards with known concentrations of GlcNAc (Sigma, MO, USA)
646 and (GlcNAc)₂₋₆ (MegaZyme, Bray, Ireland) were used to determine the concentrations of
647 (GlcNAc)₆ and the degradation products.

648 **ACKNOWLEDGEMENTS**

649 This work was supported by the U.S. Department of Energy, Office of Science, Office of
650 Biological and Environmental Research (DE-SC0014183). Additional support to E.C.M.
651 came from the Meyerhoff Graduate Fellows Program by an NIGMS Initiative for Maximizing
652 Student Development (2-R25-GM55036) and the Chemistry and Biology Interface Program
653 (T32- GM066706). TRT, GVK and VGHE were supported by the Research Council of
654 Norway through grant number 221576. We wish to thank C.E. Nelson for assistance with the
655 collection of the RNAseq samples for α -chitin and A.D. Blake for technical assistance.

656

657 **AUTHOR CONTRIBUTIONS**

658 **E.C.M.** generated *C. japonicus* mutants and performed the growth and physiology
659 experiments, generated and analyzed the RNAseq data, and contributed to writing the
660 manuscript.

661 **T.R.T.** Cloned and produced the chitinases. Planned and performed the biochemical
662 characterization, analyzed the data, and contributed to writing the manuscript.

663 **V.G.H.E** Participated in planning and analysis of the biochemical characterization of the
664 chitinases and contributed to writing the manuscript.

665 **G.V.K.** Participated in planning and analysis of the biochemical characterization of the
666 chitinases and contributed to writing the manuscript.

667 **J.G.G.** designed the study, provided overall project direction, and contributed to writing the
668 manuscript. All authors read and approved the final submitted version of the manuscript.

669 **DISCLAIMER**

670 This report was prepared as an account of work sponsored by an agency of the United States
671 Government. Neither the United States Government nor any agency thereof, nor any of their
672 employees, makes any warranty, expressed or implied, or assumes any legal liability or
673 responsibility for the accuracy, completeness, or usefulness of any information, apparatus,
674 product, or process disclosed, or represents that its use would not infringe privately owned
675 rights. Reference herein to any specific commercial product, process, or service by trade
676 name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its
677 endorsement, recommendation, or favoring by the United States Government or any agency
678 thereof. The views and opinions of authors expressed herein do not necessarily state or reflect
679 those of the United States Government or any agency thereof.

680

681 **COMPLIANCE WITH ETHICAL STANDARDS**

682 This article does not contain any studies using human participants or animals. In addition, the
683 authors declare that they have no conflict of interest.

684 **REFERENCES**

- 685 1. Gooday GW, Humphreys AM and McIntosh WH. 1986. Roles of chitinases in fungal
686 growth, p 83-91. *In* Muzzarelli, R, Jeuniaux C and Gooday GW (ed), Chitin in Nature
687 and Technology. Springer US.
- 688 2. Beier S and Bertilsson S. 2013. Bacterial chitin degradation-mechanisms and
689 ecophysiological strategies. *Front Microbiol* 4:149.
- 690 3. Khor E and Lim LY. 2003. Implantable applications of chitin and chitosan.
691 *Biomaterials* 24:2339-49.
- 692 4. Francesko A and Tzanov T. 2011. Chitin, chitosan and derivatives for wound healing
693 and tissue engineering. *Adv Biochem Eng Biotechnol* 125:1-27.
- 694 5. Matroodi S, Motallebi M, Zamani M and Moradyar M. 2013. Designing a new
695 chitinase with more chitin binding and antifungal activity. *World J Microbiol*
696 *Biotechnol* 29:1517-23.
- 697 6. Stoykov YM, Pavlov AI and Krastanov AI. 2015. Chitinase biotechnology:
698 Production, purification, and application. *Eng Life Sci* 15:30-38.
- 699 7. Chen X and Yan N. 2015. Sustainability: Don't waste seafood waste. *Nature* 524:155–
700 157.
- 701 8. Yan Q and Fong SS. 2015. Bacterial chitinase: nature and perspectives for sustainable
702 bioproduction. *Bioresources and Bioprocessing* 2:31.
- 703 9. Tharanathan RN and Kittur FS. 2003. Chitin-the undisputed biomolecule of great
704 potential. *Crit Rev Food Sci Nutr* 43:61-87.
- 705 10. Gortari MC and Hours RA. 2013. Biotechnological processes for chitin recovery out
706 of crustacean waste: A mini-review. *Electron J Biotechnol* 16.
- 707 11. Hemsworth GR, Dejean G, Davies GJ and Brumer H. 2016. Learning from microbial
708 strategies for polysaccharide degradation. *Biochem Soc Trans* 44:94-108.
- 709 12. Zargar V, Asghari M and Dashti A. 2015. A Review on Chitin and Chitosan Polymers:
710 Structure, Chemistry, Solubility, Derivatives, and Applications. *ChemBioEng*
711 *Reviews* 2:204-226.
- 712 13. Forsberg Z, Nelson CE, Dalhus B, Mekasha S, Loose JSM, Crouch LI, Røhr ÅK,
713 Gardner JG, Eijsink VGH and Vaaje-Kolstad G. 2016. Structural and functional
714 analysis of a lytic polysaccharide monooxygenase important for efficient utilization
715 of chitin in *Cellvibrio japonicus*. *J Biol Chem* 291:7300-7312.
- 716 14. Gardner JG, Crouch L, Labourel A, Forsberg Z, Bukhman YV, Vaaje-Kolstad G,
717 Gilbert HJ and Keating DH. 2014. Systems biology defines the biological significance

- 718 of redox-active proteins during cellulose degradation in an aerobic bacterium. *Mol*
719 *Microbiol* 94:1121–1133.
- 720 15. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørli M and Eijsink VGH.
721 2010. An oxidative enzyme boosting the enzymatic conversion of recalcitrant
722 polysaccharides. *Science* 330:219-222.
- 723 16. Tanaka T, Fukui T, Fujiwara S, Atomi H and Imanaka T. 2004. Concerted action of
724 diacetylchitobiose deacetylase and exo-beta-D-glucosaminidase in a novel
725 chitinolytic pathway in the hyperthermophilic archaeon *Thermococcus kodakaraensis*
726 KOD1. *J Biol Chem* 279:30021-30027.
- 727 17. Li X, Wang L-X, Wang X and Roseman S. 2007. The chitin catabolic cascade in the
728 marine bacterium *Vibrio cholerae*: characterization of a unique chitin oligosaccharide
729 deacetylase. *Glycobiology* 17:1377-1387.
- 730 18. Horn SJ, Sørbotten A, Synstad B, Sikorski P, Sørli M, Vårum KM and Eijsink VGH.
731 2006. Endo/exo mechanism and processivity of family 18 chitinases produced by
732 *Serratia marcescens*. *FEBS J* 273:491-503.
- 733 19. Meibom KL, Li XB, Nielsen AT, Wu CY, Roseman S and Schoolnik GK. 2004. The
734 *Vibrio cholerae* chitin utilization program. *Proc Natl Acad Sci USA* 101:2524-9.
- 735 20. Vaaje-Kolstad G, Bohle LA, Gaseidnes S, Dalhus B, Bjoras M, Mathiesen G and
736 Eijsink VGH. 2012. Characterization of the chitinolytic machinery of *Enterococcus*
737 *faecalis* V583 and high-resolution structure of its oxidative CBM33 enzyme. *J Mol*
738 *Biol* 416:239-54.
- 739 21. Vaaje-Kolstad G, Horn SJ, Sørli M and Eijsink VGH. 2013. The chitinolytic
740 machinery of *Serratia marcescens*-a model system for enzymatic degradation of
741 recalcitrant polysaccharides. *FEBS J* 280:3028-3049.
- 742 22. Gardner JG. 2016. Polysaccharide degradation systems of the saprophytic bacterium
743 *Cellvibrio japonicus*. *World J Microbiol Biotechnol* 32:1-12.
- 744 23. Gardner JG and Keating DH. 2012. Genetic and functional genomic approaches for
745 the study of plant cell wall degradation in *Cellvibrio japonicus*. *Methods Enzymol*
746 510:331-47.
- 747 24. DeBoy RT, Mongodin EF, Fouts DE, Tailford LE, Khouri H, Emerson JB, Mohamoud
748 Y, Watkins K, Henrissat B, Gilbert HJ and Nelson KE. 2008. Insights into plant cell
749 wall degradation from the genome sequence of the soil bacterium *Cellvibrio*
750 *japonicus*. *J Bacteriol* 190:5455-63.

- 751 25. Tuveng TR, Arntzen MØ, Bengtsson O, Gardner JG, Vaaje-Kolstad G and Eijsink
752 VGH. 2016. Proteomic investigation of the secretome of *Cellvibrio japonicus* during
753 growth on chitin. *Proteomics* 16:1904-1914.
- 754 26. Itoh Y, Kawase T, Nikaidou N, Fukada H, Mitsutomi M, Watanabe T and Itoh Y.
755 2002. Functional Analysis of the Chitin-binding Domain of a Family 19 Chitinase
756 from *Streptomyces griseus* HUT6037: Substrate-binding Affinity and cis-Dominant
757 Increase of Antifungal Function. *Biosci, Biotechnol, Biochem* 66:1084-1092.
- 758 27. Hoell IA, Dalhus B, Heggset EB, Aspmo SI and Eijsink VGH. 2006. Crystal structure
759 and enzymatic properties of a bacterial family 19 chitinase reveal differences from
760 plant enzymes. *FEBS J* 273:4889-4900.
- 761 28. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V and Henrissat B.
762 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for
763 Glycogenomics. *Nucleic Acids Res* 37:D233-D238.
- 764 29. Nelson CE, Rogowski A, Morland C, Wilhide JA, Gilbert HJ and Gardner AJ. 2017.
765 Systems analysis in *Cellvibrio japonicus* resolves predicted redundancy of beta-
766 glucosidases and determines essential physiological functions. *Mol Microbiol*
767 104:294–305.
- 768 30. Nelson CE and Gardner JG. 2015. In-frame deletions allow functional characterization
769 of complex cellulose degradation phenotypes in *Cellvibrio japonicus*. *Appl Environ*
770 *Microbiol* 81:5968-5975.
- 771 31. Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H and Krogh A. 2003.
772 Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci*
773 12:1652-62.
- 774 32. Petersen TN, Brunak S, von Heijne G and Nielsen H. 2011. SignalP 4.0:
775 discriminating signal peptides from transmembrane regions. *Nat Methods* 8:785-786.
- 776 33. Boraston AB, Bolam DN, Gilbert HJ and Davies GJ. 2004. Carbohydrate-binding
777 modules: fine-tuning polysaccharide recognition. *Biochem J* 382:769-781.
- 778 34. Wang G, Chen H, Xia Y, Cui J, Gu Z, Song Y, Chen YQ, Zhang H and Chen W. 2013.
779 How are the non-classically secreted bacterial proteins released into the extracellular
780 milieu? *Curr Microbiol* 67:688-95.
- 781 35. Bendtsen J, Jensen L, Blom N, von Heijne G and Brunak S. 2004. Feature based
782 prediction of non-classical protein secretion. *Protein Eng Des Sel* 17:349 - 356.
- 783 36. Bendtsen J, Kiemer L, Fausboll A and Brunak S. 2005. Non-classical protein secretion
784 in bacteria. *BMC Microbiol* 5:58.

- 785 37. Seydel A, Gounon P and Pugsley AP. 1999. Testing the '+2 rule' for lipoprotein
786 sorting in the *Escherichia coli* cell envelope with a new genetic selection. Mol
787 Microbiol 34:810-821.
- 788 38. Payne CM, Baban J, Horn SJ, Backe PH, Arvai AS, Dalhus B, Bjørås M, Eijsink
789 VGH, Sørliie M, Beckham GT and Vaaje-Kolstad G. 2012. Hallmarks of processivity
790 in glycoside hydrolases from crystallographic and computational studies of the
791 *Serratia marcescens* chitinases. The Journal of Biological Chemistry 287:36322-
792 36330.
- 793 39. Nakai K and Kanehisa M. 1991. Expert system for predicting protein localization sites
794 in gram-negative bacteria. Proteins 11:95-110.
- 795 40. Yin Y, Mao X, Yang J, Chen X, Mao F and Xu Y. 2012. dbCAN: a web resource for
796 automated carbohydrate-active enzyme annotation. Nucleic Acids Res 40:W445-
797 W451.
- 798 41. Black GW, Rixon JE, Clarke JH, Hazlewood GP, Theodorou MK, Morris P and
799 Gilbert HJ. 1996. Evidence that linker sequences and cellulose-binding domains
800 enhance the activity of hemicellulases against complex substrates. Biochem J 319 (Pt
801 2):515-20.
- 802 42. Gardner JG and Keating DH. 2010. Requirement of the type II secretion system for
803 utilization of cellulosic substrates by *Cellvibrio japonicus*. Appl Environ Microbiol
804 76:5079-5087.
- 805 43. Jalak J, Kurašin M, Teugjas H and Våljamäe P. 2012. Endo-exo synergism in cellulose
806 hydrolysis revisited. J Biol Chem 287:28802-28815.
- 807 44. Younes I, Hajji S, Rinaudo M, Chaabouni M, Jellouli K and Nasri M. 2016.
808 Optimization of proteins and minerals removal from shrimp shells to produce highly
809 acetylated chitin. Int J Biol Macromol 84:246-253.
- 810 45. Nikolov S, Fabritius H, Petrov M, Friák M, Lymperakis L, Sachs C, Raabe D and
811 Neugebauer J. 2011. Robustness and optimal use of design principles of arthropod
812 exoskeletons studied by ab initio-based multiscale simulations. Journal of the
813 Mechanical Behavior of Biomedical Materials 4:129-145.
- 814 46. Rawlings ND, Barrett AJ and Finn R. 2016. Twenty years of the MEROPS database
815 of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res 44:D343-
816 D350.
- 817 47. Keyhani NO and Roseman S. 1999. Physiological aspects of chitin catabolism in
818 marine bacteria. Biochim Biophys Acta 1473:108-122.

- 819 48. Perrakis A, Tews I, Dauter Z, Oppenheim AB, Chet I, Wilson KS and Vorgias CE.
820 1994. Crystal structure of a bacterial chitinase at 2.3 Å resolution. *Structure* 2:1169-
821 1180.
- 822 49. Van Aalten DMF, Synstad B, Brurberg MB, Hough E, Riise BW, Eijsink VGH and
823 Wierenga RK. 2000. Structure of a two-domain chitotriosidase from *Serratia*
824 *marcescens* at 1.9-Å resolution. *Proc Natl Acad Sci USA* 97:5842-5847.
- 825 50. Horn SJ, Sikorski P, Cederkvist JB, Vaaje-Kolstad G, Sørli M, Synstad B, Vriend G,
826 Vårum KM and Eijsink VGH. 2006. Costs and benefits of processivity in enzymatic
827 degradation of recalcitrant polysaccharides. *Proc Natl Acad Sci USA* 103:18089-
828 18094.
- 829 51. Zakariassen H, Aam BB, Horn SJ, Vårum KM, Sørli M and Eijsink VGH. 2009.
830 Aromatic residues in the catalytic center of chitinase A from *Serratia marcescens*
831 affect processivity, enzyme activity, and biomass converting efficiency. *J Biol Chem*
832 284:10610-10617.
- 833 52. Horn SJ, Sørli M, Vårum KM, Väljamäe P and Eijsink VGH. 2012. Measuring
834 Processivity. *Methods Enzymol* 510:69-95.
- 835 53. Koivula A, Kinnari T, Harjunpää V, Ruohonen L, Teleman A, Drakenberg T,
836 Rouvinen J, Jones TA and Teeri TT. 1998. Tryptophan 272: an essential determinant
837 of crystalline cellulose degradation by *Trichoderma reesei* cellobiohydrolase Cel6A.
838 *FEBS Lett* 429:341-346.
- 839 54. Watanabe T, Ariga Y, Urara S, Toratani T, Hashimoto M, Nikaidou N, Kezuka Y,
840 Nonaka T and Sugiyama J. 2003. Aromatic residues within the substrate-binding cleft
841 of *Bacillus circulans* chitinase A1 are essential for hydrolysis of crystalline chitin.
842 *Biochem J* 376:237-244.
- 843 55. Zakariassen H, Eijsink VGH and Sørli M. 2010. Signatures of activation parameters
844 reveal substrate-dependent rate determining steps in polysaccharide turnover by a
845 family 18 chitinase. *Carbohydr Polym* 81:14-20.
- 846 56. Larsbrink J, Zhu Y, Kharade SS, Kwiatkowski KJ, Eijsink VGH, Koropatkin NM,
847 McBride MJ and Pope PB. 2016. A polysaccharide utilization locus from
848 *Flavobacterium johnsoniae* enables conversion of recalcitrant chitin. *Biotechnology*
849 *for Biofuels* 9:260.
- 850 57. Howard MB, Ekborg NA, Taylor LE, Weiner RM and Hutcheson SW. 2004. Chitinase
851 B of “*Microbulbifer degradans*” 2-40 contains two catalytic domains with different
852 chitinolytic activities. *J Bacteriol* 186:1297-1303.

- 853 58. Hutcheson SW, Zhang H and Suvorov M. 2011. Carbohydrase systems of
854 *Saccharophagus degradans* degrading marine complex polysaccharides. *Mar Drugs*
855 9:645-665.
- 856 59. Igarashi K, Uchihashi T, Uchiyama T, Sugimoto H, Wada M, Suzuki K, Sakuda S,
857 Ando T, Watanabe T and Samejima M. 2014. Two-way traffic of glycoside hydrolase
858 family 18 processive chitinases on crystalline chitin. *Nature Communications* 5:3975.
- 859 60. Sikorski P, Sørbotten A, Horn SJ, Eijsink VGH and Vårum KM. 2006. *Serratia*
860 *marcescens* chitinases with tunnel-shaped substrate-binding grooves show endo
861 activity and different degrees of processivity during enzymatic hydrolysis of chitosan.
862 *Biochemistry* 45:9566-9574.
- 863 61. Nakagawa YS, Eijsink VGH, Totani K and Vaaje-Kolstad G. 2013. Conversion of α -
864 chitin substrates with varying particle size and crystallinity reveals substrate
865 preferences of the chitinases and lytic polysaccharide monooxygenase of *Serratia*
866 *marcescens*. *J Agric Food Chem* 61:11061-11066.
- 867 62. Neidhardt FC, Bloch PL and Smith DF. 1974. Culture Medium for *Enterobacteria*. *J*
868 *Bacteriol* 119:736-747.
- 869 63. Bertani G. 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic
870 *Escherichia coli*. *J Bacteriol* 62:293-300.
- 871 64. Nelson CE, Beri NR and Gardner JG. 2016. Custom fabrication of biomass
872 containment devices using 3-D printing enables bacterial growth analyses with
873 complex insoluble substrates. *J Microbiol Methods* 130:136-143.
- 874 65. Schafer A, Tauch A, Jager W, Kalinowski J, Thierbach G and Puhler A. 1994. Small
875 mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids
876 pK18 and pK19: selection of defined deletions in the chromosome of
877 *Corynebacterium glutamicum*. *Gene* 145:69-73.
- 878 66. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd and Smith HO.
879 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat*
880 *Methods* 6:343-5.
- 881 67. Figurski DH and Helinski DR. 1979. Replication of an origin-containing derivative of
882 plasmid RK2 dependent on a plasmid function provided in trans. *Proc Natl Acad Sci*
883 *USA* 76:1648-1652.
- 884 68. Murthy NKS and Bleakley BH. 2017. Simplified Method of Preparing Colloidal
885 Chitin Used For Screening of Chitinase- Producing Microorganisms. *The Internet*
886 *Journal of Microbiology* 10.

- 887 69. Janson G, Zhang C, Prado MG and Paiardini A. 2016. PyMod 2.0: improvements in
888 protein sequence-structure analysis and homology modeling within PyMOL.
889 Bioinformatics 33:444–446.
- 890 70. Savitsky P, Bray J, Cooper CD, Marsden BD, Mahajan P, Burgess-Brown NA and
891 Gileadi O. 2010. High-throughput production of human proteins for crystallization:
892 the SGC experience. J Struct Biol 172:3-13.
- 893 71. Aslanidis C and de Jong PJ. 1990. Ligation-independent cloning of PCR products
894 (LIC-PCR). Nucleic Acids Res 18:6069-6074.
- 895 72. Mekasha S, Byman IR, Lynch C, Toupalová H, Anděra L, Næs T, Vaaje-Kolstad G
896 and Eijsink VGH. 2017. Development of enzyme cocktails for complete
897 saccharification of chitin using mono-component enzymes from *Serratia marcescens*.
898 Process Biochem 56:132-138.

899

SUPPLEMENTARY MATERIAL

Chitin degradation by *Cellvibrio japonicus* is dependent on the non-redundant Chi18D chitinase

Estela C. Monge^a, Tina R. Tuveng^b, Gustav Vaaje-Kolstad^b, Vincent G. H. Eijsink^b, and Jeffrey G. Gardner^{a#}

Running Title

Chitin degradation in *C. japonicus*

Keywords:

Cellvibrio japonicus, chitin, chitinase, glycosyl hydrolase family 18, polysaccharide

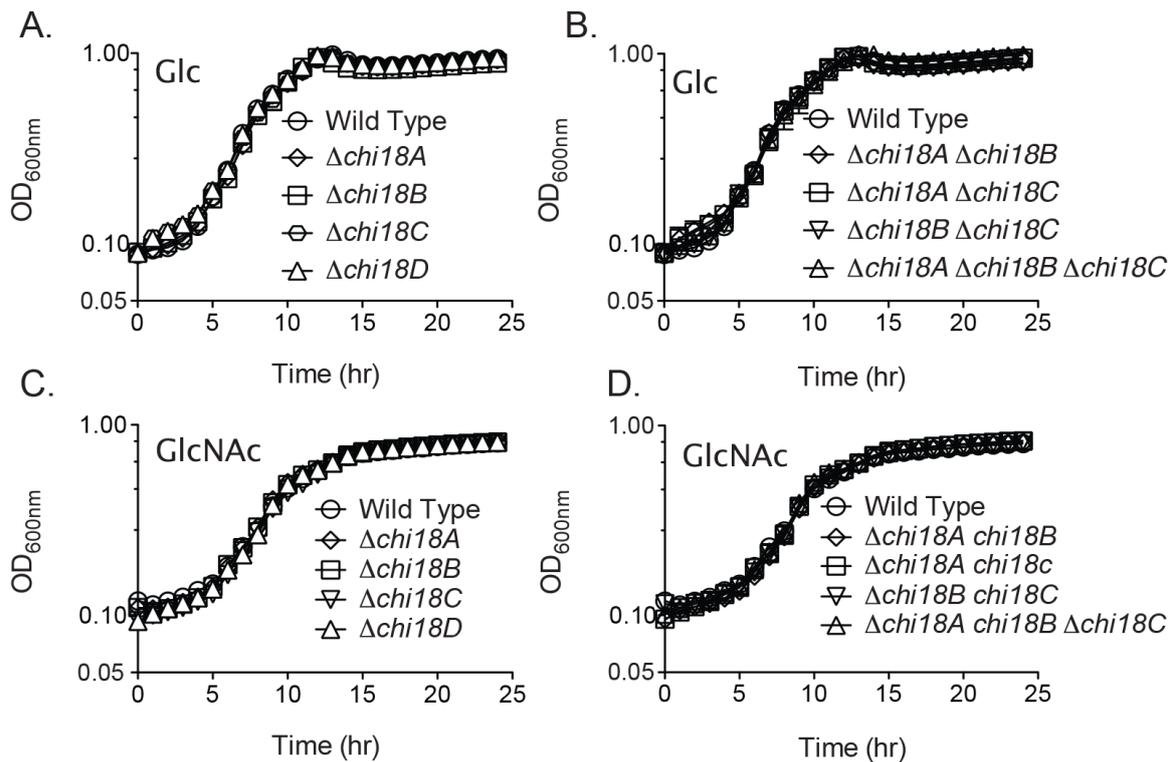
Author Affiliations

^a Department of Biological Sciences, University of Maryland - Baltimore County, Baltimore, Maryland, USA

^b Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences (NMBU), Ås, Norway

Correspondence

Jeffrey G. Gardner
Department of Biological Sciences
University of Maryland - Baltimore County
Email: jgardner@umbc.edu
Phone: 410-455-3613
Fax: 410-455-3875



47

48

49

50

51

52

53

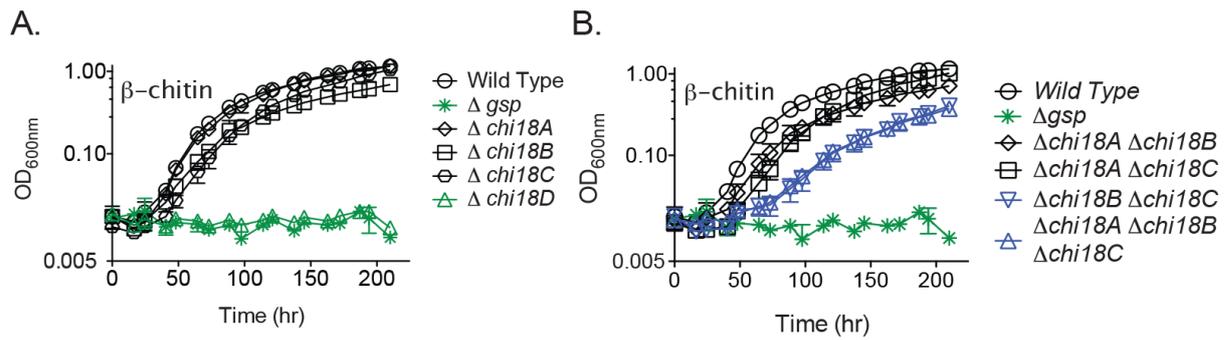
54

55

56

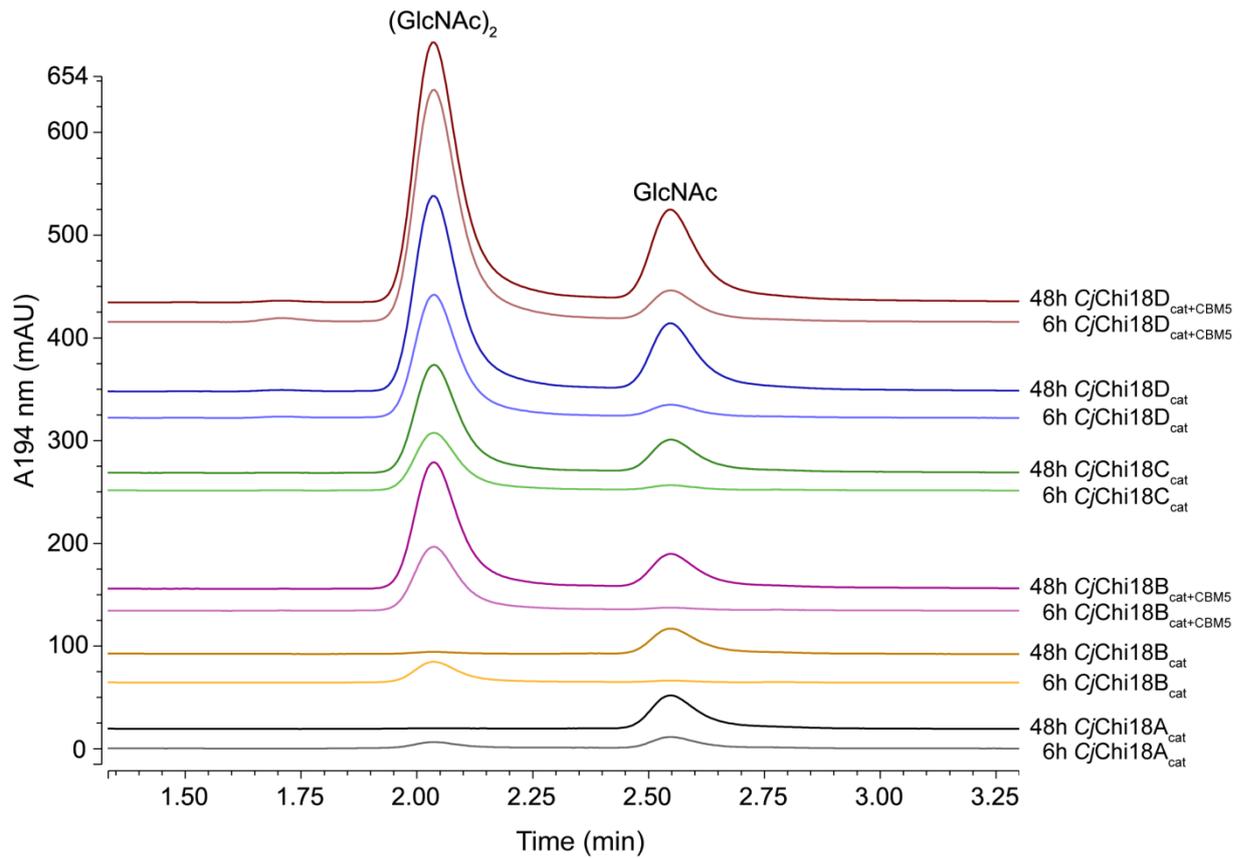
57

Fig. S1. Growth of *C. japonicus* mutants on Glc and GlcNAc. Growth analysis was conducted on MOPS minimal medium supplemented with glucose (0.2 %) (A-B) or GlcNAc (0.5 %) (C-D) as the sole source of carbon source. Although they all belong to the same experiment, the glucose experimental data is shown as two panels: single (A) and multiple (B) deletion mutants. The same applies to the GlcNAc experiments (panels C and D, respectively). All experiments were performed in biological triplicates; error bars represent standard deviations (note that in most cases they are not visible as they are smaller than symbols). These growth experiments were performed simultaneously, but are separated into multiple panels for clarity. As a consequence, the control strain (wild type) is repeated in each panel.



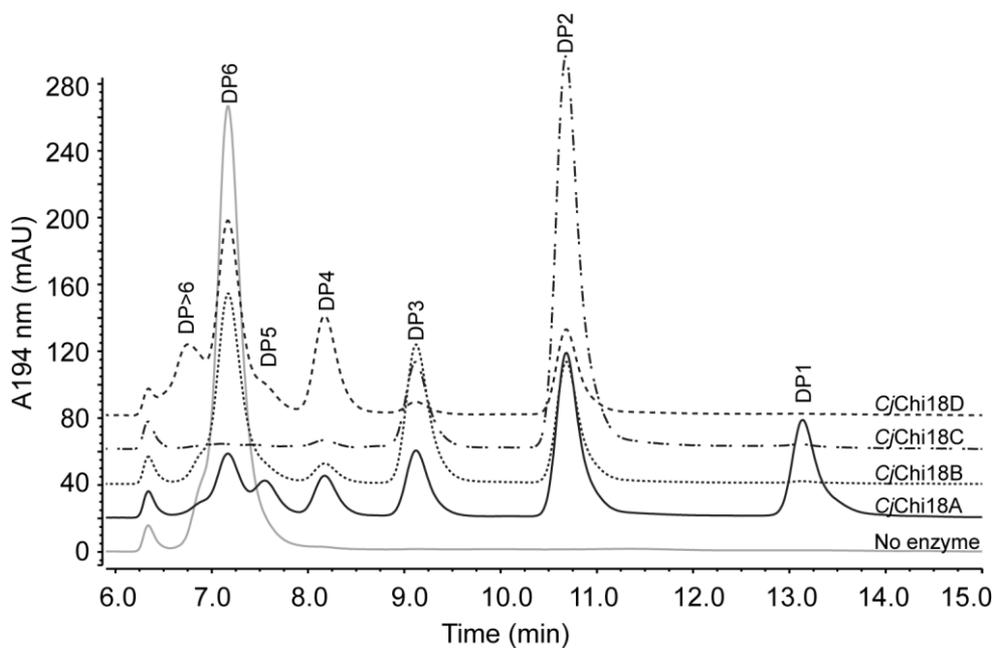
58

59 **Fig. S2. Growth of *C. japonicus* mutants on β -chitin.** Growth analysis of family GH18
60 deletion mutants of *C. japonicus* was performed in MOPS minimal medium supplemented
61 with 0.2 % of β -chitin as the sole source of carbon source. All experiments were performed
62 in biological triplicates; error bars represent standard deviations. The experimental data is
63 shown as two panels: (A) single and (B) double deletion mutants, although they all belong
64 the same experiment. As a consequence, the control strains (wild type and Δgsp) are repeated
65 in each panel.



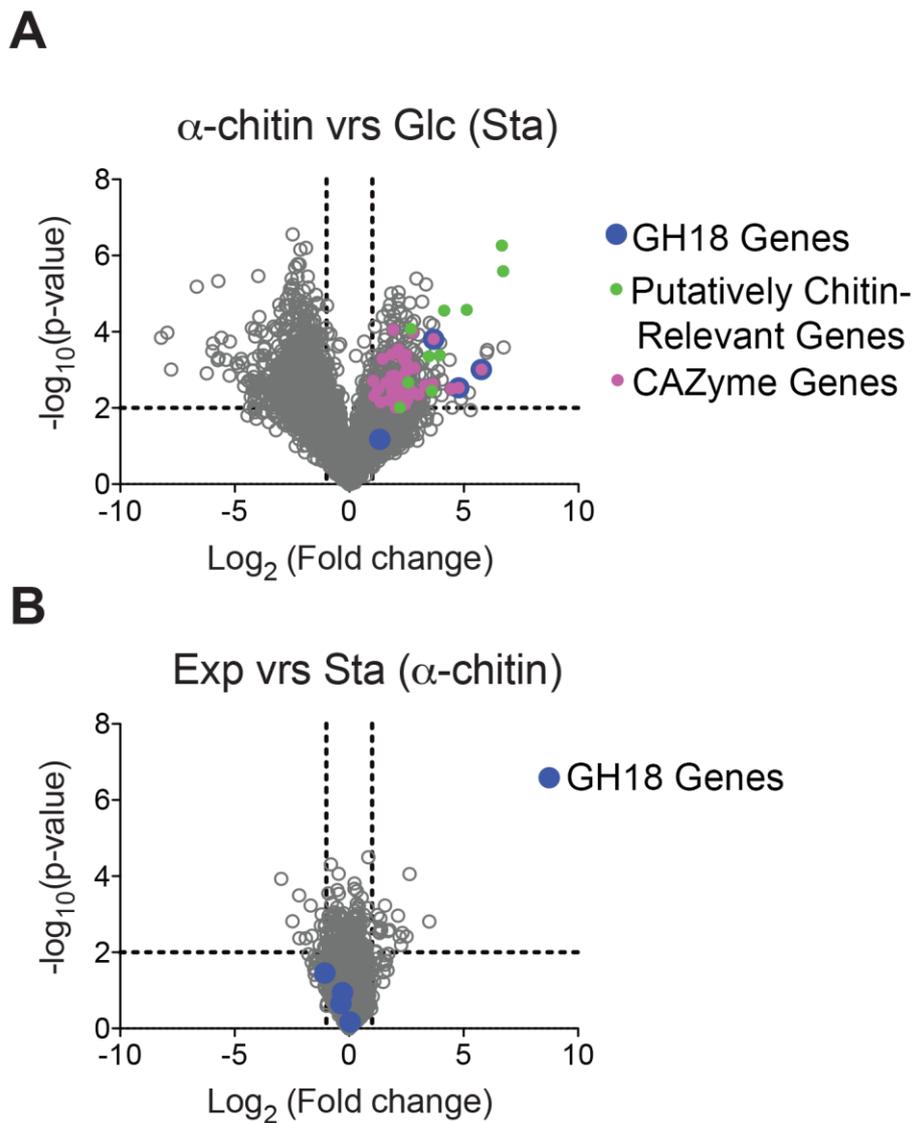
66

67 **Figure S3. Product profile after degradation of α -chitin.** Chromatograms of samples taken
 68 after 6h and 48h reaction time are shown, representing the product profile early and late in
 69 the degradation process. GlcNAc and/or $(\text{GlcNAc})_2$ are the main products for all enzymes.
 70 Reactions were done at 30 °C in 20 mM BisTris pH 6.5, 0.1 mg/ml BSA at 30 °C. The enzyme
 71 concentration was 0.5 μM .



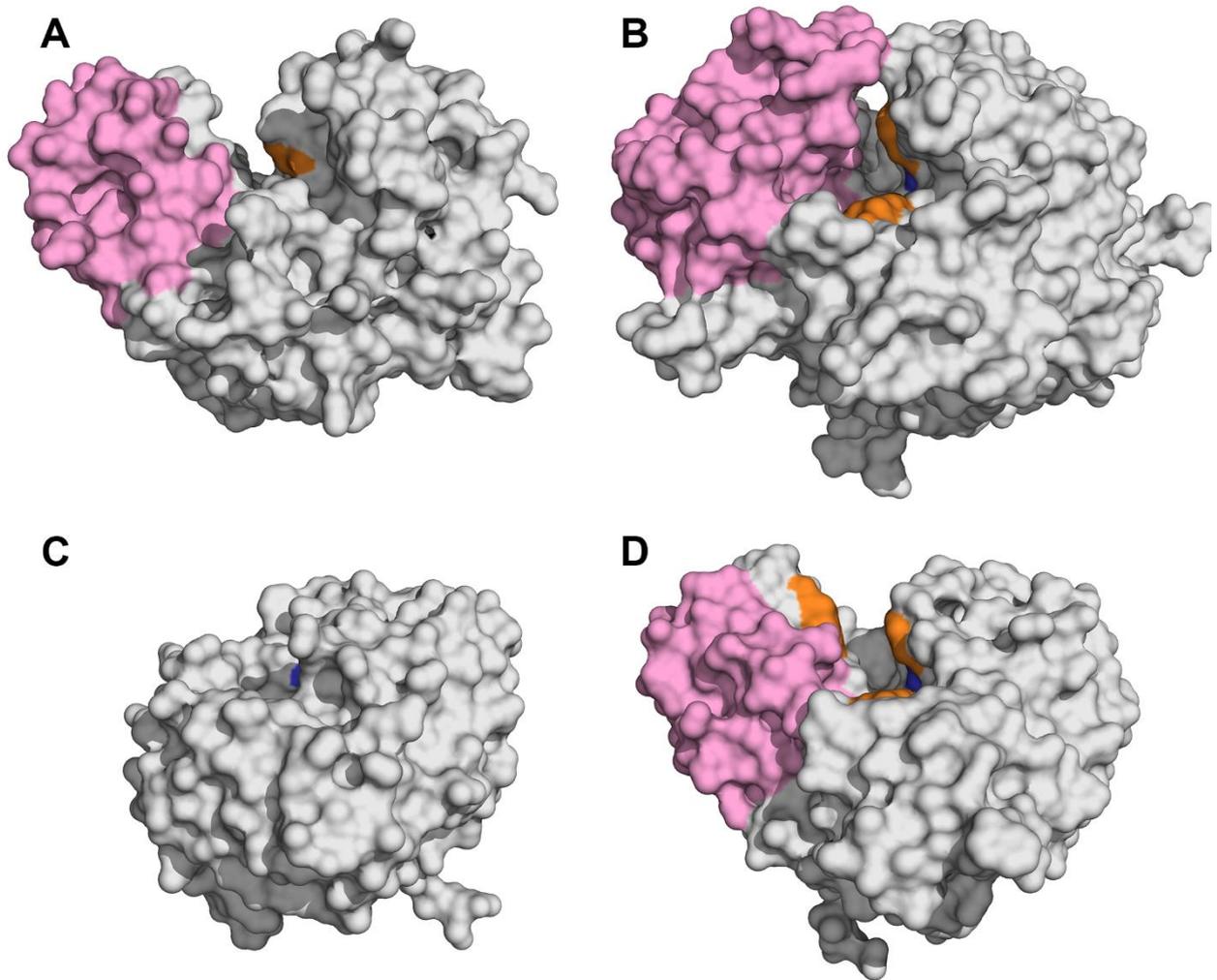
72

73 **Figure S4. Product profile after degradation of (GlcNAc)₆.** Chromatograms showing the
74 product profile obtained 60 minutes after mixing chitinases with substrate. DP1-6 represent
75 (GlcNAc)₁₋₆. Chromatogram for reaction without enzyme is also shown (grey line). These
76 experiments were done with the catalytic domains of the chitinases. The reactions contained
77 2 mM (GlcNAc)₆, 10 mM BisTris pH 6.5, 0.1 mg/ml BSA and 50 nM enzyme, and were done
78 in triplicates.



79

80 **Fig. S5. Differential gene expression in *C. japonicus*.** (A) Comparison of early stationary
 81 phase (Sta) transcriptomes during growth on glucose (Glc) or α -chitin (B) Comparison of
 82 early stationary phase and exponential phase (Exp) transcriptomes during growth on α -chitin.
 83 This volcano plot shows the \log_2 (fold change) plotted against the $-\log_{10}$ (p-value) of all
 84 expressed genes in *C. japonicus* and each gray circle represents the expression of a gene. The
 85 black dashed lines indicate the significance cut-off values: $-\log_{10}$ (p-value) >2 and \log_2 (fold
 86 change) >1 .



87

88 **Figure S6. Structural models of the catalytic domains of the four *C. japonicus* GH18**
 89 **chitinases. (A); *CjChi18A*, (B); *CjChi18B*, (C); *CjChi18C*, (D); *CjChi18D*.** Aromatic
 90 residues known to be important for processivity in other chitinases are shown in orange. The
 91 catalytic Glu is shown in blue, while the proposed $\alpha+\beta$ domain is colored pink. The models
 92 of the catalytic domains of the *C. japonicus* chitinases were build using PyMod 2.0 (1). For
 93 the *CjChi18A* model a chitinase from *Paenibacillus* sp. str. FPU-7 was used as template [35
 94 % sequence identity, PDB 5GZU (2)]. A chitinase from *Chromobacterium violaceum* (42 %
 95 sequence identity, PDB 4TXG) was used as template for *CjChi18B*. Modeling of *CjChiC* ws
 96 based on another chitinase from *Chromobacterium violaceum* (60 % sequence identity, PDB
 97 4TX8). For the *CjChiD* model, a *Streptomyces thermoviolaceus* chitinase (64 % sequence id,
 98 PDB 4W5U) was used as template.

```

SmChiA 159 -----VVGSYFVE*GVY-----GRNFTVDKIP---AQNLTLLYGEIPICGGNGIN-DSLKEIEGSPALQRS
SmChiB 1 -----MSTRKAVIGYYFIPTNQINNYTE'DT'SVVPFVSNIT-PAKAKQLTHINFSLELDINSNLECAWD-----
SmChiC 1 --MSTNNTINAAAD-----DAAIMPSIANKKILMGFWHNWAAGASDGYQQGQ-FANMNLTDIP---TEYNVVAVAEMKGGQGITPF-----
CjChiA 87 -----GYKAHEVKIIAYYM-----GDGSDLERYD---VSQIHTIYSEHLHLQGNKL-----
CjChiB 151 GNEACRPDGLYRTPGQDVPYCSVYDSEGREKMAIDRRIIGYFPS*RLGANG-----TPRYLANNIP---WGKITHINYAFAHIENNKISVGDVTSPSNPSTGLT---
CjChiC 212 -----GNARVPG-----DAA---SLPKHALVGYWHNFDNGSG-----LLRVADVD---PAWDVIVIAVDDAGNGNVEF-----
CjChiD 222 -----GGKVVGYFAQ*GVY-----ARNYHVKNIHSTSGSAAKLTHILYA*GNVQNGQCVIGDAYADYER-----

SmChiA 218 CQGREDFKVISIHDPFAALQKAQKGVTAWDDPYKGNFQGLMALKQAHPLDKILPS*IGWTLSDPFF-FM-----GDKVKRDRFVGSVKEFLQT---WKFFDGV
SmChiB 64 -----PATNDAKARDVNNRLTALKAHNPSLRIMFS*IGYYSNDLGVSHA-----NYVNAVKTASRAKFAQSCVRIMKD---Y-GFDGV
SmChiC 76 -----K-----PYNLSDETFRRQ---VGVLN--SQGRAVLSLGGADAHIEL-----KTGD-EDKLDKDEIRLREVEV---Y-GFDGL
CjChiA 130 -----AFDNEQDIRTFKQLVALKQQHPQLKVLSSLGG*WGCETCSAVF-----GSAENRQAFADSTLALLKQ---Y-SGDGL
CjChiB 247 -----WPDVPGAEMDPSLSYKGFHLLTKFKKQYQVKTMS*IVGG*WAE*TGGYFDGNGERVASGGFYTLTDSQANIDTFADSVVAFLRT---Y-HFDGA
CjChiC 270 -----RL-----DPGLNKAQFIAD---VAAKR--AQGNVLSYGGKGTVTL-----NNSTNLANFVNSTAAIINE---Y-GFDGV
CjChiD 281 -----AYTAADSVDGVADTWDQPLRGSFNQLRKLKQMPGLKVIW*SGG*WTWSGGFGQ-----ASQNTAFANSCYNLVEDPRWADVDFGI

SmChiA 311 *DIDW*YFPFGKGANPN-----LGSPODGETYVLLMKELRAMLDQLSAETGRK---YELTSASISAGKDK-ID-----KVAYNVAQNSM*HLFLMS*DFYGFAD
SmChiB 140 DIDW*YFQA*AEVD-----GFIAALQEIRTLNQQTIIDGRQALPYQLTIAGAGGAFPLSRY-----YSKLAQIVAPLBYINLMTYDLAGP*AE
SmChiC 137 DIDLE*QA*AI*GAANNKTVL-----PAALKKVKDHYAA---Q-GK---NFIISMAPEFPYLRT-----NGTYLDYINALEGY*YFIAPQY*INQQGD--
CjChiA 198 DIDW*YPTVPGYPGH-----AYGPQDKPNFTALQALRAT-----LGEH---YELSF*AAGGFDSYLEQ-----AVDWEVIMPLLSVNLMS*DMVNGGT
CjChiB 336 DIDY*EYATSMKDAGNPADWSIANARRAKLMSGYVALMKTLREKLD*DAASVQDGKH---YLLTVAAPSSGYL-LR-----GMDMYQVTQYLDYVNLMS*DLHG*AN
CjChiC 333 DIDLE*SGAGV-LHGAPVI-----QNMVSAIK-QLHA---M-FP---DLVSM*AP*EH*PYVQGGYVAYTGIW*AYLPMIDQLR*NEL*DL*LHVQL*MNNGGL--
CjChiD 362 DIDW*YFNECGLVCD-----TS---GFASYKTLMLQALRAR-----FGSS---NLVTS*AI*GAGA*AK-IN-----AADYAGA*QYV*FYMPMT*DFFG*AD

SmChiA 398 LKN-LGHQTALNAPAWKPD-----AYTTVNGVNALLAQGVKPKIV*VTAMYGRGWTGVNGYQNNIPFTGTATG-----
SmChiB 222 KV--TNHQALFGDAAGPTFYNALREANLGSWEELTRAFPSPF---SLTVDA*AVQQHLMMEGVPSAKIVM*GVPPFYGRAFKGVSGGNGGQYSSHSTPGED-----P
SmChiC 214 -----GIWVDELN---AWITQ*NNDAMKED-FLYYLTESLVTGTRGYA-KIPAAKFVIGLPSNNDAAATGYV*VNKQA-----VYNAF-----
CjChiA 279 -PH-TGHHTALEFSTDQ---Q-----K*SDT*NAV*RR*LL*RNV*VP*PK*IV*IGAAFYARVQVQAVEPTNNGLYQAGTHVD-----
CjChiB 431 EF--VGPQAPLFDGNDGEL-----QKWSFYSA-----YG--IGYLN*TDWGYHYFRGSM*PAGRINI*GV*PYYTRGWRDVS*GGTNGFWGKSP*T*SNQCAVGLT*VC*GKA
CjChiC 416 -----ATPYSGQ---AYAAGT*VDMVASARMLIEGFPLANGTAGFFQGLR*PDQVAL*GLPSG*PRAANS*GQATTANI-----NNAVNCLVLR*TGC-----
CjChiD 439 KTGPTAPH*SAL*EY*SG*PIA-----GYQSD*NAV*QLL*KSG*IPANK*ILL*GV*FYGRG*WNGV*QLAP-----GGSATG-----

SmChiA 467 -PVKGTWEN---G-----IVDYRQIAGQF-MSGEWQTYTDATAEAPYVFKPSTGDLITFDDARSVQAKGKYVL*DKQLG*GLFS*WEIDA*NDG
SmChiB 318 YPSTDYWL*V---GCEECVR-----DK-DPRIAS*YRQLEQ*MLQ*NGY*QRL*W*NDK*T*K*TPYLYHA*QNG*LFV*TYDDAES*FYK*KAKY*IKQ*QLG*V*M*F*HLG*Q*DN*RN
SmChiC 285 -----S---RLDAKN---LSIKGLMT*WSIN*WDNGK
CjChiA 343 -GVN-----FKDYETHFG---REQGFSYF*WDE*TAQ*APYV*NAAT*K*T*FAT*FD*DK*RSV*T*Q*M*RY*AK*QH*GLG*V*M*F*W*QLP*GD*KDK
CjChiB 523 VGI*DN*I*WH*DIENGA*EVAAGAN*PM*WH*AK*N*LERN*IA*ASYL*SAY*NLSAA---NLVGT*YQR*F*Y*DP*V*TS*PWL*W*NAT*K*VF*ISTE*DEES*LR*T*KAD*Y*VDK*IG*IG*IM*F*W*ELSG*DFAC
CjChiC 496 -----ATPYSGQ---AYAAGT*VDMVASARMLIEGFPLANGTAGFFQGLR*PDQVAL*GLPSG*PRAANS*GQATTANI-----NNAVNCLVLR*TGC-----
CjChiD 505 -AATGSYEA---G-----IEDYKVLK*T*KC-PT---TYTI-----AGTAYAKCGNE*WWS*FDT*P*QL*T*LG*M*DYV*NQ*NL*G*TFF*W*ELSG*TAN

```

99

100 **Figure S7. Sequence alignment.** The catalytic domains of four *C. japonicus* chitinases are
101 aligned with the catalytic domains of three well-characterized chitinases from *S. marcescens*.
102 Fully conserved residues are shown with green background, and a blue star indicates the
103 catalytic Glu acting as the catalytic acid/base. Trp residues in *SmChiA* and *SmChiB* known to
104 be important for processivity (3-5) are shown with a red background, and aligned Trp residues
105 in *C. japonicus* chitinases are shown on orange background.

106 **Table S1A. Growth of *C. japonicus* strains grown in MOPS defined media**
 107 **supplemented with α -chitin^a**

108

Strain	Growth Rate (gen hr^{-1})	Lag time (hrs)	Max OD ₆₀₀
Wild Type ^b	0.051±0.009	17.5	1.14±0.14
Δgsp	N/A	N/A	0.02±0.001
Δchi18A^c	0.059±0.016	17.5	1.20±0.21
Δchi18B^d	0.050±0.012	41	0.84±0.09
Δchi18C^d	0.058±0.009	49	1.08±0.16
Δchi18D	N/A	N/A	0.002±0.001
$\Delta\text{chi18A } \Delta\text{chi18B}^e$	0.046±0.008	17.5	0.88±0.05
$\Delta\text{chi18A } \Delta\text{chi18C}^e$	0.047±0.002	41	1.41±0.43
$\Delta\text{chi18B } \Delta\text{chi18C}^f$	0.032±0.007	41	0.50±0.06
$\Delta\text{chi18A } \Delta\text{chi18B } \Delta\text{chi18C}^g$	0.022 ±0.004	41	0.51±0.001

109

N/A = not available

^a Experiments were performed in biological triplicates; the Table shows average values and standard deviations

^b Time points used to calculate growth rate were $T_i=24$ and $T_f=66$

^c Time points used to calculate growth rate were $T_i=41$ and $T_f=66$

^d Time points used to calculate growth rate were $T_i=41$ and $T_f=72$

^e Time points used to calculate growth rate were $T_i=41$ and $T_f=96$

^f Time points used to calculate growth rate were $T_i=41$ and $T_f=137$

^g Time points used to calculate growth rate were $T_i=96$ and $T_f=120$

110 **Table S1B. Growth of *C. japonicus* strains grown in MOPS defined media**
 111 **supplemented with crab shell^a**
 112

Strain	Growth Rate (gen hr ⁻¹)	Lag time (hrs)	Max OD ₆₀₀
Wild Type ^b	0.011±0.001	25	0.707±0.011
<i>Δgsp</i>	N/A	N/A	0.011±0.004
<i>Δchi18A^c</i>	0.009±0.0009	25	0.480±0.104
<i>Δchi18B^d</i>	0.0134±0.004	25	0.514±0.051
<i>Δchi18C^d</i>	0.011±0.002	25	0.612±0.072
<i>Δchi18D</i>	N/A	N/A	0.014±0.008
<i>Δchi18A Δchi18B^b</i>	0.012±0.002	25	0.539±0.153
<i>Δchi18A Δchi18C^b</i>	0.023±0.018	25	0.732±0.182
<i>Δchi18B Δchi18C^b</i>	0.013±0.011	25	0.342±0.101
<i>Δchi18A Δchi18B Δchi18C^d</i>	0.025±0.008	25	0.485±0.236

113

N/A = not available

^a Experiments were performed in biological triplicates; the Table shows average values and standard deviations

^b Time points used to calculate growth rate were T_i=53 and T_f=125

^c Time points used to calculate growth rate were T_i=53 and T_f=148

^d Time points used to calculate growth rate were T_i=53 and T_f=103

114 **Table S2. Quantification of the zone of clearance generated by various *C. japonicus***
 115 **strains grown on colloidal chitin plates^a**
 116

Strain	Zone of Clearance (cm ²)
Wild Type	1.77±0.08
Δgsp	N/A
$\Delta chi18A$	1.93±0.14
$\Delta chi18B$	1.61±0.13
$\Delta chi18C$	1.54±0.22
$\Delta chi18D$	N/A
$\Delta chi18A \Delta chi18B$	1.54±0.00
$\Delta chi18A \Delta chi18C$	1.47±0.12
$\Delta chi18B \Delta chi18C$	0.89±0.10
$\Delta chi18A \Delta chi18B \Delta chi18C$	0.92±0.10

117

N/A = Not available

^a Experiments were performed in biological triplicates; the Table shows average values and standard deviations

118 **Table S3A. Up-regulated putative CAZyme-encoding genes during exponential growth on α -**
 119 **chitin compared to glucose^a**
 120

Substrate	CAZy name ^b	Fold change ^c	p-value ^d	Putative activity	Locus ID ^e
Chitin	<i>lpmo10A</i>	6.72	5.59	Lytic polyssaccharide mono-oxygenase	CJA_2191
Chitin	<i>chi18D</i>	6.66	6.26	Chitinase	CJA_2611
Chitin	<i>chi18C</i>	5.13	4.58	Chitinase	CJA_2993
Chitin	<i>hex20B</i>	4.14	4.56	Hexosaminidase	CJA_0287
Chitin	<i>chi18A</i>	3.95	3.38	Chitinase	CJA_1182
Chitin	<i>chi18B</i>	3.60	2.45	Chitinase	CJA_0988
Chitin	<i>nag9A</i>	3.47	3.36	Deacetylase	CJA_1163
Arabinan	<i>abf43L</i>	3.40	6.00	α -Arabinofuranidase	CJA_0806
Starch	<i>amy13A</i>	3.31	2.48	α -Amylase	CJA_2618
Arabinan	<i>abf43M</i>	3.02	3.23	α -Arabinofuranidase	CJA_0819
Transglycosylase	<i>lmt23D</i>	2.94	5.20	Transglycosylase	CJA_2884
Pectin	<i>bgl2C</i>	2.87	4.51	β -Galactosidase	CJA_2610
Arabinan	<i>gly43D</i>	2.77	2.05	α -Arabinofuranidase	CJA_0818
Arabinan	<i>gly43C</i>	2.73	3.61	α -Arabinofuranidase	CJA_0816
Cellulose	<i>cbp2E</i>	2.73	2.79	Predicted redox	CJA_2615
Chitin	<i>hex20A</i>	2.69	4.08	Hexosaminidase	CJA_0350
Carbohydrate binding protein	<i>cbp6B</i>	2.67	2.80	Carbohydrate binding protein	CJA_0276
Xylan	<i>abf62A</i>	2.61	3.22	α -Arabinofuranidase	CJA_3281
Chitin	<i>csn46F</i>	2.57	2.67	Chitosanase	CJA_2611
Pectin	<i>pga28A</i>	2.56	3.73	Polygalacturonase	CJA_0172
Cellulose	<i>cbp2D</i>	2.55	2.20	Predicted redox	CJA_2616
β -Glucan	<i>ebg98</i>	2.55	3.52	Endogalactosidase	CJA_3286
Arabinan	<i>arb43A</i>	2.47	2.47	α -Arabinofuranidase	CJA_0805
Glycosyl transferase	<i>gt5B</i>	2.34	2.13	Glycosyl transferase	CJA_3255
Starch	<i>pull13B</i>	2.32	3.21	Pullanase	CJA_3161
Xylan	<i>xyn11B</i>	2.29	2.62	Endoxylanase	CJA_3762
Chitin	<i>chi19A</i>	2.20	2.02	Chitinase	CJA_0996
Starch	<i>agd31A</i>	2.18	3.07	α -Glucosidase	CJA_3248
Xyloglucan	<i>gly74A</i>	2.14	4.07	Endoxyloglucanase	CJA_2477
Carbohydrate binding protein	<i>cbp6A</i>	2.04	2.32	Carbohydrate binding protein	CJA_1191
Starch	<i>gla15</i>	2.01	4.23	α -Glucosidase	CJA_0731

^a RNAseq sampling experiments were performed in biological triplicates

^b Names as described in DeBoy et al. (6) following the recommendations of Henrissat (7)

^c \log_2 of the fold change of the gene expression when grown in α -chitin versus glucose

^d The adjusted $-\log_{10}$ (p-value) was calculated using ArrayStar software. An adjusted p-value < 0.01 was selected as the significance cut-off value.

^e Locus IDs from DeBoy et al. (6)

Starch	<i>cbp26A</i>	1.99	4.33	Carbohydrate binding protein	CJA_2869
Xylan	<i>abf51A</i>	1.92	4.65	α -Arabinofuranidase	CJA_2769
Xylan	<i>xyn10C</i>	1.92	2.04	Endoxylanase	CJA_3066
Xylan	<i>gla67A</i>	1.89	3.91	α -Glucuronidase	CJA_2887
Arabinan	<i>gly43E</i>	1.84	2.01	α -Arabinofuranidase	CJA_0799
Xylan	<i>xyn10A</i>	1.83	2.02	Endoxylanase	CJA_2471
Mannan	<i>man26A</i>	1.74	3.81	Endomannanase	CJA_2770
Xylan	<i>xyn5A</i>	1.68	2.52	Endoxylanase	CJA_3279
Starch	<i>glu13A</i>	1.67	2.19	α -Glucosidase	CJA_0732
Starch	<i>mal77Q</i>	1.67	6.03	Amylomaltase	CJA_1882
Mannan	<i>aga27A</i>	1.64	2.01	α -Galactosidase	CJA_0246
Starch	<i>amy13J</i>	1.64	3.00	α -Amylase	CJA_0398
Cellulose	<i>LPMO10B</i>	1.63	2.48	Lytic polysaccharide mono-oxygenase	CJA_3139
Xylan	<i>xyn11A</i>	1.63	2.68	Endoxylanase	CJA_3763
Miscellaneous	<i>gly57A</i>	1.61	3.19	Glycoside hydrolase	CJA_1883
Mannan	<i>man5B</i>	1.61	2.07	Endomannanase	CJA_2480
Polysaccharide deacetylase	<i>pda4C</i>	1.60	2.70	Deacetylase	CJA_3428
Cellulose	<i>cel3B</i>	1.60	2.13	β -Glucosidase	CJA_1497
Xylan	<i>xyn10B</i>	1.59	3.28	Endoxylanase	CJA_3280
Pectin	<i>bgl2A</i>	1.58	3.08	β -Galactosidase	CJA_0496
Glycosyl transferases	<i>gt5A</i>	1.58	2.42	Glycosyl transferase	CJA_1886
Pectin	<i>gal53A-2</i>	1.55	3.20	Endogalactosidase	CJA_0491
Starch	<i>pul13A</i>	1.54	2.00	Pullanase	CJA_2160
Glycosyl transferases	<i>gt1B</i>	1.53	2.11	Rhamnosyltransferase	CJA_0772
Glycosyl transferases	<i>gt4A</i>	1.49	2.45	Glycosyl transferase	CJA_3411
Pectin	<i>pel1G</i>	1.48	2.30	Pectate lyase	CJA_3120
Starch	<i>amy13B</i>	1.48	2.67	α -Amylase	CJA_1522
Pectin	<i>pel10B</i>	1.47	3.80	Pectate lyase	CJA_2040
Pectin	<i>pel3B</i>	1.46	2.68	Pectate lyase	CJA_2413
Glycosyl transferases	<i>gt4B</i>	1.45	2.06	Glycosyl transferase	CJA_3410
Polysaccharide deacetylase	<i>pda4E</i>	1.41	2.87	Deacetylase	CJA_3408
Xylan	<i>cbp35A</i>	1.39	3.00	Carbohydrate binding protein	CJA_0020
Starch	<i>glc13A</i>	1.33	3.11	α -glucosidase	CJA_0257
Pectin	<i>bgl35A</i>	1.31	2.04	β -Galactosidase	CJA_2707
Cellulose	<i>cbp2A</i>	1.29	2.69	Carbohydrate binding protein	CJA_0007
Starch	<i>gbe13A</i>	1.25	2.33	Transglycosylase	CJA_1885
Cellulose	<i>cel45A</i>	1.24	2.23	Cellulase	CJA_0374
Cellulose	<i>cel6A</i>	1.24	2.77	Cellobiohydrolase	CJA_2473
Starch	<i>amy13F</i>	1.20	2.98	α -Amylase	CJA_0398
Transglycosylase	<i>lmt23B</i>	1.12	2.35	Transglycosylase	CJA_2053

Cellulose	<i>cel5D</i>	1.02	3.08	Cellulase	CJA_3010
Starch	<i>amy13D</i>	1.01	3.28	α -Amylase	CJA_0737

122 **Table S3B. Up-regulated putative CAZyme-encoding genes during early stationary growth on**
 123 **α -chitin compared to glucose^a**
 124

Substrate	CAZy name ^b	Fold change ^c	p-value ^d	Putative activity	Locus ID ^e
Chitin	<i>lpmo10A</i>	10.59	4.29	Lytic polyssaccharide mono-oxygenase	CJA_2191
Chitin	<i>chi18D</i>	5.77	3.01	Chitinase	CJA_2611
Chitin	<i>chi18C</i>	4.78	2.53	Chitinase	CJA_2993
Arabinan	<i>abf43L</i>	4.45	2.50	α -Arabinofuranidase	CJA_0806
Chitin	<i>chi18B</i>	3.68	3.81	Chitinase	CJA_0988
Pectin	<i>pel3B</i>	3.64	2.63	Pectate lyase	CJA_2413
Starch	<i>amy13A</i>	3.31	2.59	α -Amylase	CJA_2618
Xylan	<i>axe2C</i>	3.02	2.35	Acetylxy lan esterase	CJA_0450
Arabinan	<i>gly43C</i>	2.84	3.05	α -Arabinofuranidase	CJA_0816
β -Glucans	<i>ebg98</i>	2.76	3.98	Endogalactosidase	CJA_3286
Arabinan	<i>gly43D</i>	2.73	2.45	α -Arabinofuranidase	CJA_0818
Glycosyl Transferase	<i>gt9B</i>	2.66	2.40	Glycosyl Transferase	CJA_1369
Chitin	<i>nag9A</i>	2.58	3.06	Deacetylase	CJA_1163
Pectin	<i>bgl2C</i>	2.56	2.30	β -Galactosidase	CJA_2610
Glycosyl Transferase	<i>gt5B</i>	2.53	2.22	Glycosyl Transferase	CJA_3255
Starch	<i>amy13H</i>	2.50	3.17	α -Amylase	CJA_3247
Starch	<i>cgt13B</i>	2.47	3.39	Glucanotransferase	CJA_3263
Transglycosylase	<i>lmt23D</i>	2.47	2.95	Transglycosylase	CJA_2884
Pectin	<i>pme8C</i>	2.43	2.09	Pectinesterase	CJA_0181
Starch	<i>agd31A</i>	2.35	3.29	α -Glucosidase	CJA_3248
Xylan	<i>cpb35A</i>	2.34	2.15	Carbohydrate binding protein	CJA_0020
β -Glucan	<i>cgs94A</i>	2.32	2.37	Glucan synthetase	CJA_0849
Pectin	<i>pme8A</i>	2.31	2.59	Pectin methylesterase	CJA_0041
β -Glucan	<i>glu16A</i>	2.25	2.21	β -Glucanase	CJA_0225
Arabinan	<i>gly43G</i>	2.23	2.19	α -Arabinofuranidase	CJA_3070
Arabinan	<i>abf43M</i>	2.23	2.42	α -Arabinofuranidase	CJA_0819
Arabinan	<i>gly43J</i>	2.21	2.03	α -Arabinofuranidase	CJA_3067
Mannan	<i>man5C</i>	2.14	3.54	Endomannanase	CJA_3470
Chitin	<i>hex20A</i>	2.12	2.80	Hexosaminidase	CJA_0350
Chitin	<i>chi19A</i>	2.04	2.65	Chitinase	CJA_0996

^a RNAseq sampling experiments were performed in biological triplicate

^b Names as described in DeBoy et al. (6) following the recommendations of Henrissat (7)

^c \log_2 of the fold change of the gene expression when grown in α -chitin versus glucose

^d The adjusted $-\log_{10}$ (p-value) was calculated using ArrayStar software. An adjusted p-value < 0.001 was selected as the significance cut-off value.

^e Locus IDs from DeBoy et al. (6)

Arabinan	<i>afc95A</i>	2.03	2.01	α -fucosidase	CJA_2710
Arabinan	<i>abf51A</i>	1.98	2.85	α -fucosidase	CJA_2769
Pectin	<i>pel10B</i>	1.98	2.18	Pectate lyase	CJA_2040
Pectin	<i>pel1D</i>	1.97	2.31	Pectate lyase	CJA_2040
Polysaccharide deacetylase	<i>pda4C</i>	1.91	4.06	Deacetylase	CJA_3428
Cellulose	<i>cel5D</i>	1.89	2.73	Cellulase	CJA_3010
Xylan	<i>abf62A</i>	1.86	2.46	α -Arabinofuranidase	CJA_3281
Cellulose	<i>cbp2D</i>	1.86	3.40	Predicted redox	CJA_2616
Xylan	<i>gla67A</i>	1.85	2.84	α -Glucuronidase	CJA_2887
Xylan	<i>cbp35B</i>	1.83	2.71	Carbohydrate binding protein	CJA_0559
Pectin	<i>bgl35A</i>	1.74	2.20	β -Galactosidase	CJA_2707
Cellulose	<i>cel6A</i>	1.61	2.66	Cellulase	CJA_2473
Starch	<i>amy13B</i>	1.46	3.29	α -Amylase	CJA_1522
Mannan	<i>man5D</i>	1.36	2.15	Endomannanase	CJA_0244
Mannan	<i>man5B</i>	1.25	2.48	Endomannanase	CJA_2475
Carbohydrate Binding Protein	<i>cbp35C</i>	1.09	2.32	Carbohydrate binding protein	CJA_0494
Cellulose	<i>cel3B</i>	1.04	2.71	β -Glucosidase	CJA_1497

126 **Table S4. Product ratios after degradation of α -chitin and (GlcNAc)₆.**
 127

Chitinase	Chitin degr.	(GlcNAc)₆ degr.
	(GlcNAc) ₂ /GlcNAc ratio, 12h	(GlcNAc) ₂ /(GlcNAc) ₄ ratio, 2 min
<i>CjChi18A</i> _{cat}	0.2 ± 0.002	1.8 ± 0.048
<i>CjChi18B</i> _{cat}	7.0 ± 0.409	13.4 ± 0.692
<i>CjChi18C</i> _{cat}	7.3 ± 0.148	1.5 ± 0.067
<i>CjChi18D</i> _{cat}	2.8 ± 0.031	1.1 ± 0.048
<i>CjChi18B</i> _{cat} +CBM5	19.1 ± 0.034	N/A
<i>CjChi18D</i> _{cat} +CBM5	3.4 ± 0.109	N/A

128 N/A=not available

129 **Table S5. Strains, plasmids and primers used in this study**
130

Strains, plasmid or primer	Genotype	Source or Reference
Strains		
<i>E. coli</i> DH5 α	λ - Φ 80dlacZ Δ M15 Δ (lacZYA-argF)U169 recA1 endA1 hsdR17(rk-mk) supE44 thi-1gyrA relA1	Laboratory collection
<i>E. coli</i> S17 λ pir	Tpr Smr recA thi pro hsdR hsdM+ RP4-2-TC::Mu::Km Tn7 λ pri	Laboratory collection
<i>C. japonicus</i> Ueda 107	Wild Type	Laboratory collection
<i>C. japonicus</i> Δ gsp	Ueda 107 Δ gsp	(8)
<i>C. japonicus</i> Δ chi18A	Ueda 107 Δ chi18A ^a	This study
<i>C. japonicus</i> Δ chi18B	Ueda 107 Δ chi18B ^b	This study
<i>C. japonicus</i> Δ chi18C	Ueda 107 Δ chi18C ^c	This study
<i>C. japonicus</i> Δ chi18D	Ueda 107 Δ chi18D ^d	This study
<i>C. japonicus</i> Δ chi18A Δ chi18B	Ueda 107 Δ chi18A Δ chi18B	This study
<i>C. japonicus</i> Δ chi18A Δ chi18C	Ueda 107 Δ chi18A Δ chi18C	This study
<i>C. japonicus</i> Δ chi18B Δ chi18C	Ueda 107 Δ chi18B Δ chi18C	This study
<i>C. japonicus</i> Δ chi18A Δ chi18B Δ chi18C	Ueda 107 Δ chi18A Δ chi18B Δ chi18C	This study
Plasmids		
pK2013	ColE1 RK2-Mob ⁺ RK2-Tra ⁺ ; Km ^r	(9)
pK18mobsacB	pMB1 ori mob ⁺ sacB ⁺ ; Km ^r	(10)
pK18/ Δ chi18A	Contains 500bp upstream and downstream of <i>chi18A</i> cloned into pK18mobsacB; Km ^r	This study
pK18/ Δ chi18B	Contains 500bp upstream and downstream of <i>chi18B</i> cloned into pK18mobsacB; Km ^r	This study
pK18/ Δ chi18C	Contains 500bp upstream and downstream of <i>chi18C</i> cloned into pK18mobsacB; Km ^r	This study
pK18/ Δ chi18D	Contains 1000bp upstream and downstream of <i>chi18D</i> cloned into pK18mobsacB; Km ^r	This study
Primer		
(5') to amplify 750 bp upstream of <i>chi18D</i>	GCTATGACATGATTACGGGTGGTTA TACGCGTAATAACCTTC	This study
(3') to amplify 750 bp upstream of <i>chi18D</i>	GAATTAGCGTTTCATAGTGTTTTCTC CAACGTTTTTATATAAATACG	This study
(5') to amplify 750 bp downstream of <i>chi18D</i>	CACTATGAAACGCTAATTCATGATT ACCGGAAGC	This study

^a BioCyc accession number CJA_1182^b BioCyc accession number CJA_0988^c BioCyc accession number CJA_2993^d BioCyc accession number CJA_2611

(3') to amplify 750 bp downstream of <i>chi18D</i>	GCCTGCAGGTCGACTGGTGATATCG ATATAGCTGGCGTTG	This study
Δ <i>chi18A</i> _CONF_ (5')	ATCATGGGCAGCTTTC	This study
Δ <i>chi18A</i> _CONF_ (3')	AGCAGGAGCCTGGTA	This study
Δ <i>chi18B</i> _CONF_ (5')	CAATTGGAAATTGGTAATC	This study
Δ <i>chi18B</i> _CONF_ (3')	ATATAGTCACGCCCTATTTTG	This study
Δ <i>chi18C</i> _CONF_ (5')	AAGGGCATCTGGTTATT	This study
Δ <i>chi18C</i> _CONF_ (3')	GTATTTCTATCTGCGTTCAC	This study
Δ <i>chi18D</i> _CONF_ (5')	CTGATTGTCCCCTATCTGC	This study
Δ <i>chi18D</i> _CONF_ (3')	ATTTCCCAGCGATTGTTAC	This study
<i>chi18A</i> INT_(5')	GGTGGTTCTAGAGCTTGTATCAGTG CG	This study
<i>chi18A</i> INT_(3')	GGTGGTGAATTCCAAGCATCCTTCA CATC	This study
<i>chi18B</i> INT_(5')	GGTGGTGAATTCGCTATGTGGCGTT GA	This study
<i>chi18B</i> INT_(3')	GGTGGTTCTAGACTATGTCGTGCCA AATA	This study
<i>chi18C</i> INT_(5')	GGTGGTAAGCTTAGTTTGGGACAAC TG	This study
<i>chi18C</i> INT_(3')	GGTGGTTCTAGATGGAGTTATTCAG CG	This study
<i>chi18D</i> INT_(5')	GGTGGTAAGCTTCGACATCCTCTGT TG	This study
<i>chi18D</i> INT_(3')	GGTGGTTCTAGAATAGGCATCACCA ATA	This study

132 **Table S6. Name and description of expressed and characterized versions of the *C.***
 133 ***japonicus* GH18 chitinases.**

Name	Description
<i>CjChi18A_{cat}</i>	Chi18A catalytic domain (residues 87-432 of totally 432)
<i>CjChi18B_{cat+CBM5}</i>	Chi18B catalytic domain + CBM5 domain (residues 151-890 of totally 890)
<i>CjChi18B_{cat}</i>	Chi18B catalytic domain (residues 151-808 of totally 890)
<i>CjChi18C_{cat}</i>	Chi18C catalytic domain (residues 218-537 of totally 537)
<i>CjChi18D_{cat+CBM5}</i>	Chi18D CBM5 + catalytic domain (residues 119-588 of totally 588)
<i>CjChi18D_{cat}</i>	Chi18D catalytic domain (residues 222-588 of totally 588)

134

135 **References**

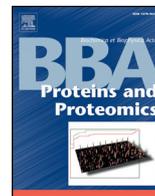
- 136 1. Janson G, Zhang C, Prado MG and Paiardini A. 2016. PyMod 2.0: improvements in
137 protein sequence-structure analysis and homology modeling within PyMOL.
138 *Bioinformatics* 33:444–446.
- 139 2. Itoh T, Hibi T, Suzuki F, Sugimoto I, Fujiwara A, Inaka K, Tanaka H, Ohta K, Fujii
140 Y, Taketo A and Kimoto H. 2016. Crystal Structure of Chitinase ChiW from
141 *Paenibacillus* sp. str. FPU-7 Reveals a Novel Type of Bacterial Cell-Surface-
142 Expressed Multi-Modular Enzyme Machinery. *PLoS One* 11:e0167310.
- 143 3. Horn SJ, Sikorski P, Cederkvist JB, Vaaje-Kolstad G, Sørлие M, Synstad B, Vriend G,
144 Vårum KM and Eijsink VGH. 2006. Costs and benefits of processivity in enzymatic
145 degradation of recalcitrant polysaccharides. *Proc Natl Acad Sci USA* 103:18089-
146 18094.
- 147 4. Payne CM, Baban J, Horn SJ, Backe PH, Arvai AS, Dalhus B, Bjørås M, Eijsink
148 VGH, Sørлие M, Beckham GT and Vaaje-Kolstad G. 2012. Hallmarks of processivity
149 in glycoside hydrolases from crystallographic and computational studies of the
150 *Serratia marcescens* chitinases. *The Journal of Biological Chemistry* 287:36322-
151 36330.
- 152 5. Zakariassen H, Aam BB, Horn SJ, Vårum KM, Sørлие M and Eijsink VGH. 2009.
153 Aromatic residues in the catalytic center of chitinase A from *Serratia marcescens*
154 affect processivity, enzyme activity, and biomass converting efficiency. *J Biol Chem*
155 284:10610-10617.
- 156 6. DeBoy RT, Mongodin EF, Fouts DE, Tailford LE, Khouri H, Emerson JB, Mohamoud
157 Y, Watkins K, Henrissat B, Gilbert HJ and Nelson KE. 2008. Insights into plant cell
158 wall degradation from the genome sequence of the soil bacterium *Cellvibrio*
159 *japonicus*. *J Bacteriol* 190:5455-63.
- 160 7. Henrissat B. 1998. Glycosidase families. *Biochem Soc Trans* 26:153.
- 161 8. Nelson CE and Gardner JG. 2015. In-Frame Deletions Allow Functional
162 Characterization of Complex Cellulose Degradation Phenotypes in *Cellvibrio*
163 *japonicus*. doi:10.1128/AEM.00847-15.
- 164 9. Figurski DH and Helinski DR. 1979. Replication of an origin-containing derivative of
165 plasmid RK2 dependent on a plasmid function provided in trans. *Proc Natl Acad Sci*
166 *USA* 76:1648–1652.
- 167 10. Schafer A, Tauch A, Jager W, Kalinowski J, Thierbach G and Puhler A. 1994. Small
168 mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids
169 pK18 and pK19: selection of defined deletions in the chromosome of
170 *Corynebacterium glutamicum*. *Gene* 145:69-73.

171

Paper IV

Genomic, proteomic and biochemical analysis of the chitinolytic machinery of *Serratia marcescens* BJL200.

Tuveng, T. R., Hagen, L. H., Mekasha, S., Frank, J., Arntzen, M. Ø., Vaaje-Kolstad, G. & Eijsink, V. G. H. 2017. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1865, 414-421.



Genomic, proteomic and biochemical analysis of the chitinolytic machinery of *Serratia marcescens* BJL200



Tina R. Tuveng, Live Haldal Hagen, Sophanit Mekasha, Jeremy Frank, Magnus Øverlie Arntzen, Gustav Vaaje-Kolstad, Vincent G.H. Eijsink *

Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, 1432 Ås, Norway

ARTICLE INFO

Article history:

Received 27 October 2016

Received in revised form 10 January 2017

Accepted 23 January 2017

Available online 24 January 2017

Keywords:

Proteomics

Chitinases

Serratia marcescens

Secretome

Biomass degradation

ABSTRACT

The chitinolytic machinery of *Serratia marcescens* BJL200 has been studied in detail over the last couple of decades, however, the proteome secreted by this Gram-negative bacterium during growth on chitin has not been studied in depth. In addition, the genome of this most studied chitinolytic *Serratia* strain has until now, not been sequenced. We report a draft genome sequence for *S. marcescens* BJL200. Using label-free quantification (LFQ) proteomics and a recently developed plate-method for assessing secretomes during growth on solid substrates, we find that, as expected, the chitin-active enzymes (ChiA, B, C, and CBP21) are produced in high amounts when the bacterium grows on chitin. Other proteins produced in high amounts after bacterial growth on chitin provide interesting targets for further exploration of the proteins involved in degradation of chitin-rich biomasses. The genome encodes a fourth chitinase (ChiD), which is produced in low amounts during growth on chitin. Studies of chitin degradation with mixtures of recombinantly produced chitin-degrading enzymes showed that ChiD does not contribute to the overall efficiency of the process. ChiD is capable of converting *N,N'*-diacetyl chitobiose to *N*-acetyl glucosamine, but is less efficient than another enzyme produced for this purpose, the Chitobiase. Thus, the role of ChiD in chitin degradation, if any, remains unclear.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Chitin, found in crustaceans, insects, and fungal cell walls, is often considered as the second most abundant biopolymer in nature, after cellulose. The shells of crustaceans consist, in addition to chitin, of calcium carbonate and protein. The extraction of chitin from these resources is today done by use of harsh, environmentally unfriendly chemicals, and further processing of the chitin, e.g. to make chitosan, also involves such chemicals [1,2]. It is therefore desirable to make this process more environmental friendly, for instance by use of enzymes. Many microorganisms are known to degrade chitin and by studying the proteins secreted by an organism growing on chitin, a better understanding of the enzymatic degradation of chitin-rich biomasses can be obtained.

The chitinolytic machinery of *Serratia marcescens* BJL200, a Gram-negative soil bacterium, is well studied [3–5]. It is already known that, during growth on chitin, this bacterium produces three chitinases (ChiA, ChiB, ChiC), belonging to glycosyl hydrolase (GH) family 18, a lytic polysaccharide monooxygenase (LPMO), belonging to auxiliary activity (AA) family 10, named CBP21, and a β -hexosaminidase (Chitobiase), belonging to the GH20 family [6–11]. The latter enzyme

is located in the periplasm and converts *N,N'*-diacetyl chitobiose, the primary product of the chitinases, into *N*-acetyl glucosamine for further utilization by the cell [12]. ChiA, ChiB and ChiC are known to be secreted by *S. marcescens*, but ChiB and ChiC do not harbor a conventional signal peptide [6,10,13]. Both ChiB and ChiC are known to be located in the periplasm before export to the extracellular space [6,13], possibly through a Type 2 secretion system commonly found in Gram-negative bacteria [14,15]. However, Hamilton et al. [13] showed that *S. marcescens* Db10/Db11 lacks a typical Type 2 secretion system, and demonstrated that a holin-like protein (ChiW) and an endopeptidase (ChiX) are essential for secretion of chitinolytic enzymes. In addition, a LysR-type transcription regulator (ChiR) is essential for production of the chitinolytic machinery of *S. marcescens* 2170 [16]. The *chiR*, *chiW* and *chiX* genes are located in the same region as the *chiB* and *cbp21* genes on the chromosome of *S. marcescens* Db10/Db11 [13].

The genome of *S. marcescens* BJL200 likely encodes at least one more GH18 chitinase (ChiD), since this enzyme is encoded in the genomes of other members of the *Serratia* genus [17,18]. Chitinase D from *Serratia proteamaculans* (SpChiD) has been characterized in detail and displays both hydrolytic and transglycolytic activities [17,19–21], whereas, similar properties have recently been described for a chitinase from *S. marcescens* GPS5 [18]. However, a quantitative characterization of ChiD activity has not been described and little is known about how the activity of ChiD relates to the activity of the other chitinolytic enzymes. To obtain more insight into chitin degradation by *S. marcescens*,

Abbreviations: CAZymes, carbohydrate active enzymes; GH, glycosyl hydrolase; LPMO, lytic polysaccharide monooxygenase; LFQ, label free quantification.

* Corresponding author.

E-mail address: vincent.eijsink@nmbu.no (V.G.H. Eijsink).

and potentially discover additional proteins involved in chitin conversion, we have carried out a quantitative proteomics study of the secretome of *S. marcescens* BJL200 during growth on chitin. Identified proteins were mapped on the draft genome sequence of the bacterium which was determined as part of this work. The genome indeed contained a *chiD* gene, and the proteomic data showed that the ChiD protein is much less abundant during growth on chitin compared to the other chitinases and CBP21. We have therefore performed an in-depth characterization of ChiD to investigate its possible role in chitin degradation by *S. marcescens*.

2. Materials and methods

2.1. DNA extraction

DNA extraction was done essentially as described by Rosewarne et al. [22], with minor modifications. Quantification of DNA was done using the Qubit™ fluorometer and the Quant-iT™ dsDNA BR Assay Kit (Invitrogen, CA, USA), before sequencing.

2.2. Genome sequencing and annotations

Illumina MiSeq sequencing was performed at the Norwegian Sequencing Centre (Oslo, Norway), using TruSeq sample preparation and a 2 × 300 paired-end sequencing kit (Illumina, CA, USA). Low quality base calls (Phred quality score < 20) and short sequences (length < 20 bp) were trimmed using Sickle PE [23], and the sequences were *de novo* assembled using IDBA UD [24]. Paired-end reads files were merged prior to assembly using the fq2fa program implemented the IDBA package. Annotation of the contigs was done using Rapid Annotation Subsystem Technology version 2.0 (RAST, <http://rast.nmpdr.org/>) [25–27]. The raw sequencing data has been uploaded to the NCBI Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/SRA/>) under accession number SRP076778. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession MSEC00000000. The version described in this paper is version MSEC01000000.

The subcellular location of proteins was predicted using the LipoP 1.0 server (<http://www.cbs.dtu.dk/services/LipoP/>) [28], and the PRED-TAT software (<http://www.compgen.org/tools/PRED-TAT/>) [29]. Proteins annotated as cytosolic (CYT) by the LipoP server were further analyzed using the SecretomeP 2.0 server (<http://www.cbs.dtu.dk/services/SecretomeP/>) [31] to detect possible non-classical secretion. We considered proteins with a SecretomeP score of 0.5 or higher as secreted in a non-classical (NC) fashion. Annotation of carbohydrate active enzymes (CAZymes) according to the CAZy database (<http://www.cazy.org/Citing-CAZy.html>) [11] was done using dbCAN (<http://csbl.bmb.uga.edu/dbCAN/index.php>) [30] with hidden Markov models version 3.0.

2.3. Proteomics

S. marcescens BJL200 was grown on agarose (1%) plates with 1% wt/vol α-chitin (extracted from *Pandalus borealis*, Seagarden, Husøyvegen 278, Karmsund Fiskerihavn, 4262 Avaldsnes, Norway), 1% wt/vol β-chitin (extracted from squid pen, Batch 20140101, France Chitin, Chemin de Porte Claire, F-84100 Orange, France), or 0.2% wt/vol glucose (VWR International) as sole carbon source in M9 minimal medium. The M9 minimal medium was supplemented with 1 mM MgSO₄ and 0.1 mM CaCl₂. The plates were prepared essentially as described by Tuveng et al. [31], using a sterile Grade QM-A Quartz Filter, circle, 47 mm (GE Healthcare Life Sciences, Oslo, Norway). The plates comprise two layers of identically composed solid medium; the filter is located between the two layers and separates cells (growing on the top of the plate) from the bottom of the plate, where protein samples were collected [32]. Glass petri dishes with a diameter of 80 mm were used and the incubation temperature for bacterial growth was 30 °C.

Secretomes from cells grown on plates were collected at different time points using biological triplicates for each time point. Samples were prepared as described by Bengtsson et al. [32] with the exception that trypsinated samples were dried under vacuum (Concentrator plus, Eppendorf, Denmark) to concentrate the samples, and dissolved in 0.1% (vol/vol) trifluoro acetic acid (TFA), before purification of peptides using ZipTip C18 pipette tips (Merck Millipore, Cork, Ireland). To produce samples for analysis of intracellular proteins, we collected cells by scraping them off the plates and suspending them in lysis buffer (50 mM Tris, 0.1% triton X-100, 200 mM NaCl, 1 mM DTT, pH 7.1). Glass beads (acid washed, ≤106 μm, Sigma, Oslo, Norway) were then added and the cells were disrupted using a FastPrep-24 (MP Biomedicals, CA, USA) for 3 × 1 min, followed by centrifugation. Proteins in the supernatant were precipitated by adding trichloro acetic acid (TCA) to a final concentration of 16% (vol/vol). After collecting the precipitated proteins by centrifugation at 15,000 × g, the proteins were dissolved in SDS-buffer and loaded onto a SDS-PAGE gel for standard electrophoretic separation. In-gel digestion was performed essentially as described by Shevchenko et al. [33] and peptides were purified using ZipTips. All peptide samples were dried under vacuum, dissolved in 10 μl 2% (vol/vol) acetonitrile (ACN), 0.1% (vol/vol) TFA, and analyzed by LC-MS/MS.

2.4. Mass spectrometry

Mass spectrometry analysis of peptides was done essentially as described by Tuveng et al. [31]. In brief, peptides were analyzed using two technical replicates for each biological replicate or one technical replicate for the cell lysis samples. The system used was a nanoHPLC-MS/MS system consisting of a Dionex Ultimate 3000 RSLCnano (Thermo Scientific, Bremen, Germany) connected to a Q-Exactive hybrid quadrupole-orbitrap mass spectrometer (Thermo Scientific, Bremen, Germany), equipped with a nano-electrospray ion source. Samples were loaded onto a trap column (Acclaim PepMap100, C₁₈, 5 μm, 100 Å, 300 μm i.d. × 5 mm, Thermo Scientific, Bremen, Germany) and back-flushed onto an analytical column (Acclaim PepMap RSLC C₁₈, 2 μm, 100 Å, 75 μm i.d., Thermo Scientific, Bremen, Germany). In order to isolate and fragment the 10 most intense peptide precursor ions at any given time throughout the chromatographic elution, the mass spectrometer was operated in data-dependent mode to switch automatically between orbitrap-MS and higher-energy collisional dissociation (HCD) orbitrap-MS/MS acquisition. The selected precursor ions were then excluded for repeated fragmentation for 20 s. The resolution was set to $R = 70,000$ and $R = 35,000$ for MS and MS/MS, respectively. For optimal acquisition of MS/MS spectra, automatic gain control (AGC) target values were set to 50,000 charges and a maximum time of 128 milliseconds.

2.5. Data analysis

MS raw files were analyzed using MaxQuant version 1.4.1.2 [34,35], and proteins were identified and quantified using the MaxLFQ algorithm [36]. Technical replicates, i.e. HPLC re-injections, were combined during MaxQuant analysis. The data were searched against a custom database of the predicted proteome of *S. marcescens* BJL200 (5202 protein sequences), supplemented with common contaminants such as keratins, trypsin and BSA. In addition, reversed sequences of all protein entries were concatenated in order to estimate the false discovery rate (FDR). The tolerance levels for matching to the database were 6 ppm for MS, 20 ppm for MS/MS. Trypsin was used as digestion enzyme, and two missed cleavages were allowed. N-terminal acetylation, oxidation of methionine, conversion of glutamine to pyro glutamic acid, and deamidation of asparagine and glutamine were set as variable modifications (carbamidomethylation of cysteine residues was set as fixed modification in analysis of cell lysis samples). The ‘match between runs’ feature of MaxQuant was enabled with default parameters, in order to transfer identifications between samples based on accurate mass and

retention time [36]. The settings allowed only transfer of peptides between samples from the same carbon source. All identifications were filtered to achieve a 1% FDR. The results were further processed using Perseus version 1.5.5.3 as described previously by Tuveng et al. [31]. In brief, after log10 transformation of LFQ values, missing values were imputed from a normal distribution (width of 0.2 and downshifted 1.9 standard deviations from the original distribution) in total matrix mode. A protein was considered present if it was identified in at least two of three biological replicates in at least one time point on at least one substrate.

2.6. Cloning and expression of *ChiD*

Genomic DNA from *S. marcescens* BJL200 was extracted from an overnight culture by boiling and centrifugation. Amplification of the *chiD* gene was done by PCR, using the following primers; forward 5'-TTAAGAAGGAGATATACTATGATGAAAATAACCCGAAAACCTG-3' and reverse 5'-AATGGTGGTATGATGGTGGCCGCTTCTCGCCTTTTATCCC-3'. The amplified gene, encoding full-length ChiD including its signal peptide, was subsequently cloned into a pNIC-CH vector [37] using ligation-independent cloning. The pNIC-CH vector containing a gene encoding ChiD with a C-terminal 6x-His tag was then transformed into *Escherichia coli* BL21 and positive transformants were selected on LB-plates containing kanamycin and 5% (wt/vol) sucrose. For expression, a pre-culture was made by inoculating 5 ml LB medium supplemented with kanamycin. The pre-culture was then used to inoculate 0.5 L TB-medium supplemented with kanamycin and Antifoam 204 (Sigma-Aldrich, Oslo, Norway), followed by incubation at 37 °C in a Harbinger system (LEX-48 Bioreactor, Harbinger biotech, Markham, Canada). At $OD_{600} \approx 0.6$, the culture was induced with IPTG (final concentration 0.2 mM) and grown over night at 30 °C before harvesting cells by centrifugation. Exported ChiD was isolated from the periplasm. To prepare periplasmic extracts, the cells were resuspended in 50 ml ice-cold spheroplast buffer (0.5 mM EDTA, 0.1 M Tris HCL pH 8, 17.1% (wt/vol) sucrose, 0.125 mM PMSF). Cells were pelleted by centrifugation (8000 rpm, 10 min at 4 °C), before warming the pellet to room temperature. After resuspending the pellet in 42 ml ice-cold dH₂O and addition of 2.1 ml 20 mM MgCl₂, the cells were incubated on ice for 45 s and then the cells were pelleted by centrifugation as above. The supernatant, i.e. the periplasmic extract, was filtered (0.45 µm) and the buffer was exchanged to 20 mM Tris-HCl, 150 mM NaCl and 10 mM imidazole, pH 8.0, using Amicon Ultra-15 centrifugal filters with 10,000 NMWL (Merck Millipore, Cork, Ireland). The protein was purified by nickel affinity chromatography on a HisTrap HP 5 ml column (GE Healthcare Life Sciences, Oslo, Norway) connected to an Äkta pure system (GE Healthcare Life Sciences, Oslo, Norway). A stepwise imidazole gradient ending at 300 mM imidazole was used to elute bound protein. After checking protein purity by SDS-PAGE the buffer was exchanged to 20 mM Tris-HCl, 100 mM NaCl, pH 8.0, and the protein was concentrated using Amicon Ultra-15 centrifugal filters. The protein concentration was measured using the Bradford micro assay supplied by Bio-Rad (CA, USA).

2.7. Enzymatic assays

Standard reactions contained 15 g/l α -chitin, 15 g/l β -chitin, or different concentrations of *N,N'*-diacetyl chitobiose as substrate, in 20 mM BisTris pH 6.0. Purified ChiD was obtained as described above, whereas all other enzymes were recombinantly produced and purified in-house, using previously published methods [6,7,10,38]. For activity assays with multiple enzymes, the total enzyme concentration was 1 µM. In the experiments for Michaelis-Menten kinetics the enzyme concentrations were 1 µM and 0.2 nM for ChiD and Chitobiase, respectively and in this case, the reaction mixtures contained 0.1 mg/ml BSA. Reactions were incubated in a thermomixer at 40 °C, 800 rpm and degradation products were analyzed using a Rezex RFQ-Fast Acid H+ (8%)

Table 1

Genome sequencing. The table shows the genomic size (base pairs), the GC content, the number of contigs, the number of predicted coding sequences (CDSs), and predicted CAZymes in the genome of *S. marcescens* BJL200. The predicted subcellular location of proteins is also provided, in absolute numbers and percentage of the total (in parenthesis). Abbreviations of subcellular locations; Spl, signal peptidase I cleavage site; SplI, signal peptidase II cleavage site; NC, non-classically secreted; TAT, twin arginine signal peptide; CYT, cytosolic; TMH, trans membrane helix; TS, total secreted (sum of Spl, SplI, TAT and NC).

Genome		Subcellular location	
Genome size (bp)	5,850,214	Spl	499 (9.6%)
GC content (%)	59.02	SplI	113 (2.2%)
Contigs	1632	TAT	46 (0.9%)
CDSs	5198	NC	446 (8.6%)
CAZymes	160	CYT	3364 (64.7%)
		TMH	734 (14.1%)
		TS	1104 (21%)

ion-exclusion column (Phenomenex, CA, USA) installed on a Dionex Ultimate 3000 HPLC with UV-detection, as previously described by Hamre et al. [39].

2.8. Deglycosylation of *RNaseB*

20 µg RNaseB [New England Biolab (NEB), MA, USA] was denatured by a 10 min incubation at 100 °C in 1 µl 10× Glycoprotein denaturing buffer (NEB, MA, USA) and 9 µl dH₂O. Subsequently, 2 µl 10× Glycobuffer 3 (NEB, MA, USA), dH₂O and the desired enzyme [PNGaseF (NEB, MA, USA), Endo H (NEB, MA, USA) or ChiD] were added to reach a total reaction volume of 20 µl. Incubation of the reaction mixture at 37 °C for 1 h was followed by SDS-PAGE analysis. The gel was stained with Coomassie Brilliant Blue R250 (BioRad, CA, USA) to visualize protein bands.

3. Results and discussion

3.1. Draft genome of *S. marcescens* BJL200

The draft genome of *S. marcescens* BJL200 was obtained by Illumina MiSeq sequencing, resulting in 9,761,690 paired end sequences, and annotated using RAST 2.0 [25] (Table 1). The genomic size (5.85 Mb), and hence the number of CDSs, are slightly higher than for other *S. marcescens* genomes, whose usual size is in the range of 4.9 to 5.2 Mb [40].

The identified CDSs were translated into protein sequences, before further annotation in terms of the putative location of the gene

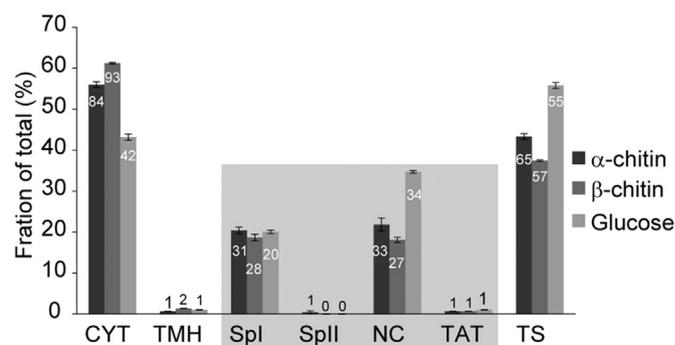


Fig. 1. Prediction of subcellular locations. Proteins identified in the secretomes during growth on different carbon sources are categorized by predicted subcellular location: CYT, cytosolic; TMH, trans membrane helix; Spl, signal peptidase I cleavage site; SplI, signal peptidase II cleavage site; NC, non-classically secreted; TAT, twin arginine signal peptide; TS, total secreted (summing up the categories in the grey box). The bars represent the average of three biological replicates for time points indicated in the main text, with standard deviation. Average of total numbers of proteins per category is indicated in the figure. The average total numbers of identified proteins were 138 ± 5 , 152 ± 7 , and 100 ± 3 for α -chitin, β -chitin and glucose, respectively.

substrates, we selected the 20-h time point for glucose and β -chitin, while we used the 48-h time point for α -chitin. These time points represent roughly similar amount of growth, based on color development.

3.3. Subcellular location of proteins

Of the 232 proteins identified at the selected time points (for all substrates taken together), 103 were predicted as extracellular proteins. Fig. 1 shows an overview of the number of proteins belonging to the different subcellular locations found in the secretomes for the three carbon sources at the selected time points. It is evident that the fractions of cytosolic proteins are high, ranging from around 45% for glucose to around 60% for α - and β -chitin. This is somewhat surprising, as a previous study on *Cellvibrio japonicus* showed that the plate method allows harvesting of secretomes that are more strongly enriched for secreted proteins, with cytosolic fractions down to about 30% [31].

The relatively high fraction of cytosolic proteins in the secretomes may reflect how *S. marcescens* secretes proteins. As already mentioned, ChiB and ChiC do not have a conventional sec-signal peptide, yet they are known to be secreted, by a mechanism that is not fully understood [6,13]. Both secretion of proteins by unknown mechanisms and appearance of cytoplasmic proteins are well known from studies on secretomes [41,42]. There are several plausible explanations for the presence of cytosolic proteins in the secretome: (1) Some “cytoplasmic” proteins may in fact be secreted and could be so-called moon-lightning proteins, i.e. proteins with two functions, one relevant in the extracellular space and one relevant in the cytoplasm [41]. (2) Gram-negative bacteria are known to produce outer membrane vesicles, and these are often found to contain cytoplasmic proteins [43]; indeed, the production of outer membrane vesicles has been observed in *S. marcescens* [44]. (3) Some cells may actually lyse on purpose and by that release cytosolic content into the surrounding to benefit the remaining cells [45].

Interestingly, the periplasmic Chitobiase [12,46] is absent from all secretomes, suggesting that cell lysis or damage was limited. To investigate this further, we also analyzed the intracellular proteomes and observed that the fraction of cytosolic proteins was higher than in the secretomes (above 70% in all cases, the predicted theoretical maximum being 65%; Fig. S3). In these samples, the Chitobiase was readily detected in the α - and β -chitin samples but not in glucose samples (Fig. S4), which shows that the Chitobiase indeed was induced and which adds confidence to the notion that the lack of Chitobiase in the secretomes reflects that cell lysis was limited. Overall, these observations indicate that the secretome samples to a large extent represent the true secretome of *S. marcescens*.

Notably, the other GH20 protein encoded in the genome, which is predicted to have a signal peptide, was also detected in the samples of intracellular proteins, including in the glucose samples. This enzyme does not seem to be related to chitin conversion and was not considered further.

3.4. Secreted, chitin-active proteins

Fig. 2 shows a heat map for all the 103 putatively secreted proteins, which are clustered into 7 clusters according to the substrate-dependency of their abundance. The four well-known chitin-degrading CAZymes (ChiA, ChiB, ChiC and CBP21) are all clearly more abundant in the chitin samples, compared to the glucose sample (Fig. 2, 3). Interestingly, the secretome is almost devoid of additional CAZymes [11], indicating that the response to the chitin polysaccharide is very specific. The known chitin-degrading enzymes occur in two clusters, 1 and 4. These two clusters contain 14 additional proteins, including three proteins with putative function, which are interesting targets for future investigations on chitin metabolism in *S. marcescens*. Clusters 5 and 6 also contain proteins showing higher abundance in chitin samples, and in these two clusters we find five hypothetical proteins, which also are interesting targets for further investigations of their role. ChiD is part of

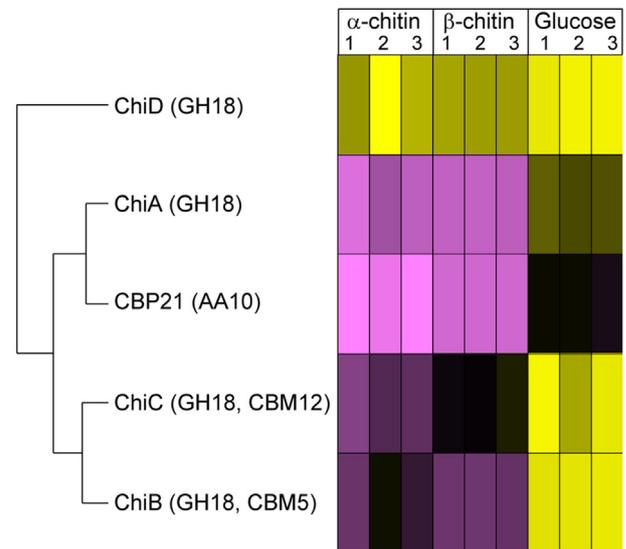


Fig. 3. Comparative analysis of the expression of chitinases and CBP21. Heat map comparing the abundance of ChiA, B, C, D, and CBP21 after growth on the different carbon sources (three biological replicates for each), for the selected time points (see text for details). The colors in the heat map indicate protein abundance, ranging from high (pink color, \log_{10} MaxLFQ = 9.0) to low abundance (yellow color, \log_{10} MaxLFQ = 5.5). Missing values were replaced by imputation as described in the Materials and methods section.

the large cluster 3 containing low abundance proteins; low amounts of ChiD were only detected in the chitin samples.

Three proteins in cluster 1 (containing ChiA and CBP21; Fig. 2) are annotated as secreted alkaline metalloproteinases, and a BLAST against the MEROPS database (<http://merops.sanger.ac.uk/index.shtml>) [47] showed that they belong to metallo-peptidase family M10B containing the serralsins. *A priori*, we expected to see upregulated proteases in the chitin samples, since chitin-rich biomasses comprise a co-polymeric

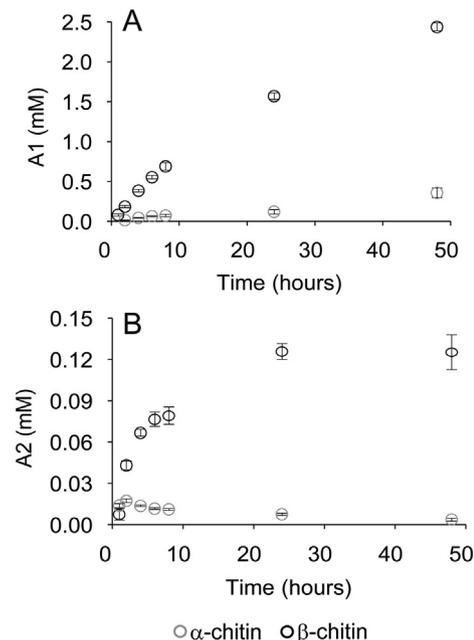


Fig. 4. Degradation of α - and β -chitin by ChiD. The graphs show product formation during incubation of 15 g/l α - or β -chitin with 1 μ M ChiD in 20 mM BisTris pH 6.0, at 40 °C. Panel A shows *N*-acetyl glucosamine (A1), whereas panel B shows *N,N'*-diacetylchitobiose (A2). In the α -chitin samples, small amounts of trimer and tetramer were also observed (less than 10% of total product formation). All data points are the average of three replicates with error bars representing standard deviations.

network of chitin and protein. Serralysins are sometimes considered as virulence factors in pathogenic *Serratia* species [48], but one study has shown that a serralysin increased the chitinolytic activity of ChiA from *Serratia* sp. KCK toward crude chitin [49]. One of the serralysins found in cluster 1 (contig-100_30_3497_36484) shows 93% sequence identity with the serralysin from *Serratia* sp. KCK, while the other two have 50–60% sequence identity to this protein. These three proteases appear in the same cluster as ChiA and CBP21, both of which, notably, were also detected at low levels in the glucose cultures. It is not certain that these proteases are regulated by the presence of chitin, since they were also detected at high levels in the glucose samples, whereas two other detected proteases (Cluster 2 and 7) were clearly most abundant in the glucose samples.

Fig. 3 shows a close-up view of the expression of ChiA, B, C, D, and CBP21, underpinning that all are more abundant in the chitin samples compared to the glucose samples. ChiD is much less abundant relative to the other chitinases in the chitin samples. In the chitin samples, the abundance of most chitinolytic enzymes, including ChiD, increased over time (Fig. S5). At all time-points ChiD LFQ values were 10–100 times lower compared to the other chitinases. ChiD was not observed in the glucose samples at any time point (Fig. S5; Table S1).

3.5. Characterization of SmChiD

The proteomics data suggest that ChiD is not crucial for the degradation of chitin, but the fact that the enzyme was detected in the chitin samples only, indicates that it might play a role at a later stage of the chitin degradation. In order to explore the role of ChiD we expressed and purified the enzyme and performed a deeper characterization of this chitinase, with the aim of complementing previous work on similar enzymes and assessing its biological role.

Fig. 4 shows that ChiD is able to degrade α - and β -chitin, albeit with low efficiency compared to other *Serratia* chitinases (comparative studies appear further below). Unexpectedly for a true chitinase, *N*-acetyl glucosamine is the main product of ChiD, whereas smaller amounts of *N,N'*-diacetylchitobiose are also produced. The amount of *N,N'*-diacetylchitobiose decreases over time, implying that ChiD converts *N,N'*-diacetylchitobiose to *N*-acetyl glucosamine. The latter observation is consistent with previous findings for homologous proteins [18,21].

To investigate the effect of ChiD on the product profiles of ChiA, ChiB and ChiC, and to see if ChiD would add degradative power to a cocktail of these better known chitinases, α -chitin was incubated with various chitinase combinations. The main product of the three other chitinases

is *N,N'*-diacetylchitobiose [50] and the results obtained upon combining each of these enzymes with ChiD show that ChiD increased the formation of *N*-acetyl glucosamine (Fig. 5). Notably, Fig. 5 shows that the amount of product generated by ChiD alone is much lower than for the other chitinases. These findings support the proteomics data, which suggest that ChiD does not have an essential function in chitin degradation. The data in Fig. 5 suggest that ChiD may have a role in converting *N,N'*-diacetylchitobiose to *N*-acetyl glucosamine, but Fig. 5 also shows that the enzyme thought to be responsible for this task, the GH20 Chitobiase, a periplasmic enzyme which is induced by chitin (Fig. S4) is much more efficient. Adding ChiD to a cocktail of the three other chitinases did not yield synergistic effects, whereas addition of the Chitobiase did yield such effects (Fig. 5).

To further compare ChiD and the Chitobiase, we performed Michaelis-Menten kinetics to determine K_M and k_{cat} for degradation of *N,N'*-diacetylchitobiose (Fig. 6). The kinetic data confirm that Chitobiase is a much more effective degrader of *N,N'*-diacetylchitobiose than ChiD, performing around 1000-fold more cuts per second, making it unlikely that *N,N'*-diacetylchitobiose conversion by ChiD is biologically relevant. Notably, in contrast to Chitobiase, ChiD operates outside the cell; possibly, the ability to slowly convert *N,N'*-diacetylchitobiose to *N*-acetyl glucosamine in the extracellular space is important for the bacterium, perhaps for regulatory purposes.

In a final attempt to allocate a putative biological function to ChiD, we considered the fact that *N,N'*-diacetylchitobiose type of structures are found in glycoproteins, and that chitinases have been proposed to have a role in virulence of pathogenic bacteria, possibly targeting host glycoproteins containing *N,N'*-diacetylchitobiose [3]. Indeed, some GH18 enzymes are known to cleave the bond between the two *N*-acetyl glucosamine units in *N*-linked glycans, e.g. in the model substrate RNase B [51]. We therefore tested if ChiD was able to deglycosylate RNaseB and the results showed that this was not the case for ChiD (Fig. 7), nor for the other *Serratia* chitinases (results not shown).

4. Concluding remarks

The genome sequence of *Serratia marcescens* BJL200 shows that the extensively studied chitinolytic machinery of this bacterium comprises one more GH18 enzyme, ChiD, next to the already well-known three chitinases, A, B and C, the Chitobiase and a single LPMO, CBP21. The proteomic data show that in terms of secreted CAZymes, the response to chitin is very specific in the sense that all the well-known secreted chitin-active enzymes are found in much higher amounts after growth on

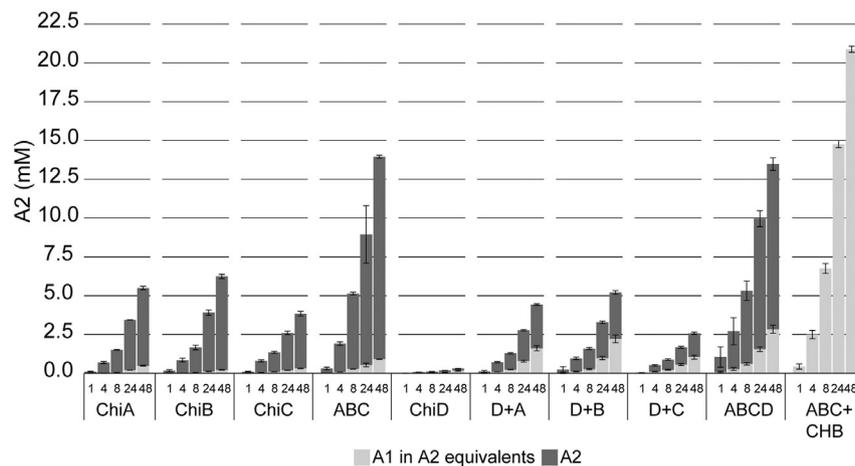


Fig. 5. Product profiles generated from α -chitin by individual chitinases and their combinations. The figure shows the product profiles obtained at different time points (indicated in hours) during incubation of 15 g/l α -chitin with a total enzyme concentration of 1 μ M in 20 mM BisTris pH 6.0, at 40 °C. The total enzyme concentration was 1 μ M in each reaction and enzyme fractions in mixtures were always equal (on a molar basis). The chitinases are indicated by ChiA, ChiB, ChiC and ChiD or, in reactions with more than one enzyme, A, B, C, and D. CHB indicates the Chitobiase. The amount of A1, i.e. *N*-acetyl glucosamine, is presented as A2 (*N,N'*-diacetylchitobiose) equivalents (i.e. the measured concentrations of A1 were divided by two prior to making the graph). The bars represent the average of three replicates with error bars representing the standard deviation.

chitin compared to growth on glucose. In addition, the other CAZymes predicted to be secreted (Fig. 2), except the CBM50, are only present in the chitin samples. The rather limited response of CAZymes in general, only 12 at the selected time points (10 of which are considered as secreted), may reflect the nature of the substrates on which the bacterium was grown. Chitin, although recalcitrant, is a simple substrate compared to e.g. hemicellulose.

The secretome studies revealed several proteins that are co-regulated with the chitinases, which could indicate that they play a role in chitin conversion. However, other explanations are thinkable; for example, growth on a substrate difficult to degrade, such as α -chitin, may induce a general stress response, e.g. due to poor access to nutrition. Most of the 14 proteins found in the same clusters as the well-known chitin-degrading enzymes (Fig. 2), except for the serralysins that may have a role in degradation of crude chitin, have no obvious role in chitin-conversion. The three putative proteins among these 14, together with the hypothetical proteins in cluster 5 and 6, may be worth further investigations. Overall, the data indicate that the response to chitin is very limited in the sense that this response is dominated by expression of the known chitinolytic machinery.

A.R. Podile and his team have previously studied enzymes homologous to *SmChiD* found in the *S. marcescens* BJL200 genome [17–21]. Both hydrolytic and transglycosylating activities against different chito-oligosaccharides have been described, but so far the biological role of this chitinase has not been addressed, and there are no studies on the interplay between ChiD and other *Serratia* chitinases. The only previously described kinetic parameter for a ChiD homologue is a k_{cat} of 0.36 s^{-1} for *SpChiD* acting on *N,N'*-diacetylchitobiose. This value is very similar to the k_{cat} of 0.27 s^{-1} described here for *SmChiD* and three orders of magnitude lower than the k_{cat} for Chitobiase acting on the same substrate. Fig. 5 shows that ChiD is much less active on α -chitin, compared to the other *Serratia* chitinases. Compared to these other chitinases, *SpChiD* has an extra loop that occludes part of the substrate

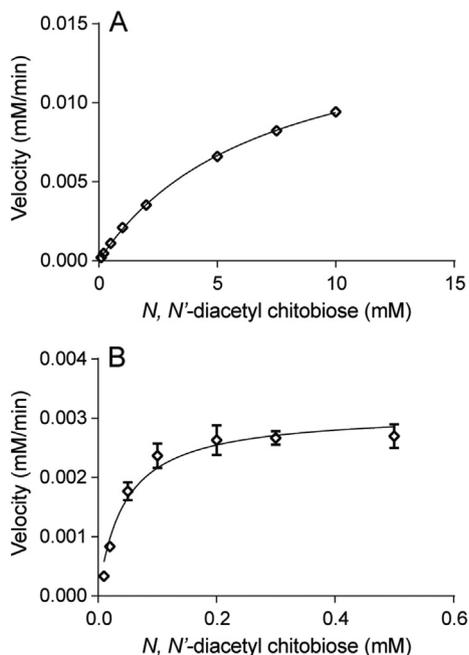


Fig. 6. Michaelis-Menten kinetics for degradation of *N,N'*-diacetylchitobiose by for ChiD (A) and Chitobiase (B). Initial enzyme velocities are plotted against substrate concentration (mM) and the curves were generated by nonlinear fitting using GraphPad Prism 7. K_M values were determined to be $6.94 \pm 0.28 \text{ mM}$ and $0.044 \pm 0.006 \text{ mM}$ for ChiD and Chitobiase, respectively, while k_{cat} values were determined to be $0.265 \pm 0.005 \text{ s}^{-1}$ and $259 \pm 10 \text{ s}^{-1}$ for ChiD and Chitobiase, respectively. All reactions were done in triplicates and the error bars represent standard deviations. Some of the error bars are not visible because they are smaller than the data point. The enzyme concentrations were $1 \mu\text{M}$ and $0.0002 \mu\text{M}$ for ChiD and Chitobiase, respectively.

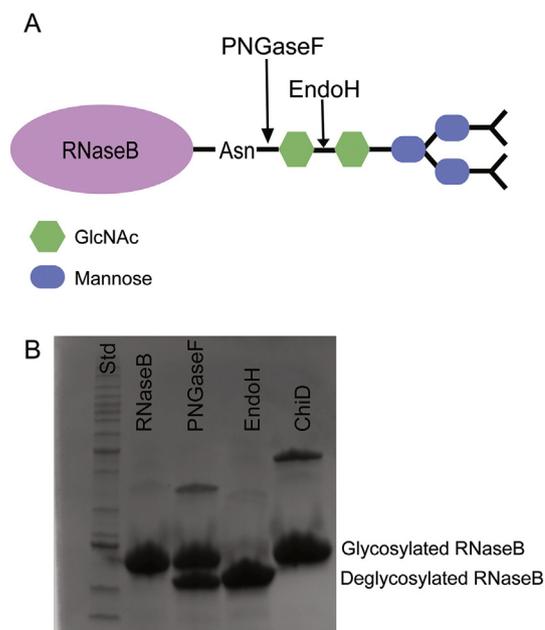


Fig. 7. Deglycosylation of RNaseB. A) Schematic representation of the glycosylation in RNaseB (GlcNAc is *N*-acetyl glucosamine). B) SDS-PAGE gel stained with Coomassie brilliant blue after incubation of different enzymes with RNaseB. From left; lane 1: BenchMark protein ladder, lane 2: RNaseB incubated without other enzyme, lane 3: RNaseB incubated with PNGaseF, lane 4: RNaseB incubated with EndoH, lane 5: RNaseB incubated with ChiD.

binding cleft [19], likely affecting the catalytic properties [20]. While ChiD may be a useful enzyme for production of chito-oligosaccharides, due to its well-studied transglycosylating activity [17,18,20,21] the previous and present biochemical data, as well as the current proteomic study indicate that this enzyme is not involved in chitin conversion by *S. marcescens*. The enzyme could perhaps play a role in virulence, for example by cleaving other glycans than the one tested in this study, or its transglycosylating abilities may be biologically relevant.

The following are the supplementary data related to this article.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.bbapap.2017.01.007>.

Transparency Document

The Transparency document associated with this article can be found, in online version.

Acknowledgements

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [52] partner repository with the dataset identifier PXD005225. This work was funded by the research council of Norway, through the MARPOL project (221576). We thank Ida R. Byman for assistance in cloning and purification of *SmChiD*.

References

- [1] R. Hackman, Studies on chitin I. Enzymic degradation of chitin and chitin esters, *Aust. J. Biol. Sci.* 7 (1954) 168–178.
- [2] K.M. Vårum, M.H. Ottøy, O. Smidsrød, Acid hydrolysis of chitosans, *Carbohydr. Polym.* 46 (2001) 89–98.
- [3] G. Vaaje-Kolstad, S.J. Horn, M. Sorlie, V.G. Eijsink, The chitinolytic machinery of *Serratia marcescens* – a model system for enzymatic degradation of recalcitrant polysaccharides, *FEBS J.* 280 (2013) 3028–3049.
- [4] S. Takanao, S. Honma, T. Miura, C. Ogawa, H. Sugimoto, K. Suzuki, T. Watanabe, Construction and basic characterization of deletion mutants of the genes involved in

- chitin utilization by *Serratia marcescens* 2170, *Biosci. Biotechnol. Biochem.* 78 (2014) 524–532.
- [5] K. Suzuki, M. Shimizu, N. Sasaki, C. Ogawa, H. Minami, H. Sugimoto, T. Watanabe, Regulation of the chitin degradation and utilization system by the ChiX small RNA in *Serratia marcescens* 2170, *Biosci. Biotechnol. Biochem.* 80 (2016) 376–385.
- [6] M.B. Brurberg, V.G.H. Eijsink, A.J. Haandrikman, G. Venema, I.F. Nes, Chitinase B from *Serratia marcescens* BJL200 is exported to the periplasm without processing, *Microbiology* 141 (1995) 123–131.
- [7] B. Synstad, G. Vaaje-Kolstad, F.H. Cederkvist, S.F. Saua, S.J. Horn, V.G.H. Eijsink, M. Sørlie, Expression and characterization of endochitinase C from *Serratia marcescens* BJL200 and its purification by a one-step general chitinase purification method, *Biosci. Biotechnol. Biochem.* 72 (2008) 715–723.
- [8] G. Vaaje-Kolstad, B. Westereng, S.J. Horn, Z. Liu, H. Zhai, M. Sørlie, V.G. Eijsink, An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides, *Science* 330 (2010) 219–222.
- [9] K. Suzuki, M. Suzuki, M. Taiyoshi, N. Nikaidou, T. Watanabe, Chitin binding protein (CBP21) in the culture supernatant of *Serratia marcescens* 2170, *Biosci. Biotechnol. Biochem.* 62 (1998) 128–135.
- [10] M.B. Brurberg, V.G. Eijsink, I.F. Nes, Characterization of a chitinase gene (chiA) from *Serratia marcescens* BJL200 and one-step purification of the gene product, *FEMS Microbiol. Lett.* 124 (1994) 399–404.
- [11] V. Lombard, H. Golaconda Ramulu, E. Drula, P.M. Coutinho, B. Henrissat, The carbohydrate-active enzymes database (CAZy) in 2013, *Nucleic Acids Res.* 42 (2014) D490–D495.
- [12] T. Toratani, T. Shoji, T. Ikehara, K. Suzuki, T. Watanabe, The importance of chitobiase and N-acetylglucosamine (GlcNAc) uptake in N,N'-diacetylchitobiose [(GlcNAc)₂] utilization by *Serratia marcescens* 2170, *Microbiology* 154 (2008) 1326–1332.
- [13] J.J. Hamilton, V.L. Marlow, R.A. Owen, A. Costa Mde, M. Guo, G. Buchanan, G. Chandra, M. Trost, S.J. Coulthurst, T. Palmer, N.R. Stanley-Wall, F. Sargent, A holin and an endopeptidase are essential for chitinolytic protein secretion in *Serratia marcescens*, *J. Cell Biol.* 207 (2014) 615–626.
- [14] J.G. Gardner, D.H. Keating, Requirement of the type II secretion system for utilization of cellulosic substrates by *Cellvibrio japonicus*, *Appl. Environ. Microbiol.* 76 (2010) 5079–5087.
- [15] N.P. Cianciotto, Type II secretion: a protein secretion system for all seasons, *Trends Microbiol.* 13 (2005) 581–588.
- [16] K. Suzuki, T. Uchiyama, M. Suzuki, N. Nikaidou, M. Regue, T. Watanabe, LysR-type transcriptional regulator ChiR is essential for production of all chitinases and a chitin-binding protein, CBP21, in *Serratia marcescens* 2170, *Biosci. Biotechnol. Biochem.* 65 (2001) 338–347.
- [17] P. Purushotham, A.R. Podile, Synthesis of long-chain chitooligosaccharides by a hypertransglycosylating processive endochitinase of *Serratia proteamaculans* 568, *J. Bacteriol.* 194 (2012) 4260–4271.
- [18] P.R. Vaikuntapu, S. Rambabu, J. Madhuprakash, A.R. Podile, A new chitinase-D from a plant growth promoting *Serratia marcescens* GP55 for enzymatic conversion of chitin, *Bioresour. Technol.* 220 (2016) 200–207.
- [19] J. Madhuprakash, A. Singh, S. Kumar, M. Sinha, P. Kaur, S. Sharma, A.R. Podile, T.P. Singh, Structure of chitinase D from *Serratia proteamaculans* reveals the structural basis of its dual action of hydrolysis and transglycosylation, *Int. J. Biochem. Mol. Biol.* 4 (2013) 166–178.
- [20] J. Madhuprakash, K. Tanneeru, B. Karlapudi, L. Guruprasad, A.R. Podile, Mutagenesis and molecular dynamics simulations revealed the chitooligosaccharide entry and exit points for chitinase D from *Serratia proteamaculans*, *Biochim. Biophys. Acta* 1840 (2014) 2685–2694.
- [21] J. Madhuprakash, K. Tanneeru, P. Purushotham, L. Guruprasad, A.R. Podile, Transglycosylation by chitinase D from *Serratia proteamaculans* improved through altered substrate interactions, *J. Biol. Chem.* 287 (2012) 44619–44627.
- [22] C.P. Rosewarne, P.B. Pope, S.E. Denman, C.S. McSweeney, P. O'Cuiv, M. Morrison, High-yield and phylogenetically robust methods of DNA recovery for analysis of microbial biofilms adherent to plant biomass in the herbivore gut, *Microb. Ecol.* 61 (2011) 448–454.
- [23] N.A. Joshi, J.N. Fass, Sickle: A Sliding-window, Adaptive, Quality-based Trimming Tool for FastQ Files (Version 1.33) [Software], 2011.
- [24] Y. Peng, H.C. Leung, S.-M. Yiu, F.Y. Chin, IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth, *Bioinformatics* 28 (2012) 1420–1428.
- [25] R.K. Aziz, D. Bartels, A.A. Best, M. DeJongh, T. Disz, R.A. Edwards, K. Formsma, S. Gerdes, E.M. Glass, M. Kubal, F. Meyer, G.J. Olsen, R. Olson, A.L. Osterman, R.A. Overbeek, L.K. McNeil, D. Paarmann, T. Paczian, B. Parrello, G.D. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein, A. Wilke, O. Zagnitko, The RAST server: rapid annotations using subsystems technology, *BMC Genomics* 9 (2008) 75.
- [26] T. Brettin, J.J. Davis, T. Disz, R.A. Edwards, S. Gerdes, G.J. Olsen, R. Olson, R. Overbeek, B. Parrello, G.D. Pusch, M. Shukla, J.A. Thomson 3rd, R. Stevens, V. Vonstein, A.R. Wattam, F. Xia, RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes, *Sci. Rep.* 5 (2015) 8365.
- [27] R. Overbeek, R. Olson, G.D. Pusch, G.J. Olsen, J.J. Davis, T. Disz, R.A. Edwards, S. Gerdes, B. Parrello, M. Shukla, V. Vonstein, A.R. Wattam, F. Xia, R. Stevens, The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST), *Nucleic Acids Res.* 42 (2014) D206–D214.
- [28] A.S. Juncker, H. Willenbrock, G. Von Heijne, S. Brunak, H. Nielsen, A. Krogh, Prediction of lipoprotein signal peptides in Gram-negative bacteria, *Protein Sci.* 12 (2003) 1652–1662.
- [29] P.G. Bagos, E.P. Nikolaou, T.D. Liakopoulos, K.D. Tsirogos, Combined prediction of Tat and Sec signal peptides with hidden Markov models, *Bioinformatics* 26 (2010) 2811–2817.
- [30] Y. Yin, X. Mao, J. Yang, X. Chen, F. Mao, Y. Xu, dbCAN: a web resource for automated carbohydrate-active enzyme annotation, *Nucleic Acids Res.* 40 (2012) W445–W451.
- [31] T.R. Tuveng, M.Ø. Arntzen, O. Bengtsson, J.G. Gardner, G. Vaaje-Kolstad, V.G.H. Eijsink, Proteomic investigation of the secretome of *Cellvibrio japonicus* during growth on chitin, *Proteomics* 16 (2016) 1904–1914.
- [32] O. Bengtsson, M.Ø. Arntzen, G. Mathiesen, M. Skaugen, V.G.H. Eijsink, A novel proteomics sample preparation method for secretome analysis of *Hypocrea jecorina* growing on insoluble substrates, *J. Proteomics* 131 (2016) 104–112.
- [33] A. Shevchenko, H. Tomas, J. Havli, J.V. Olsen, M. Mann, In-gel digestion for mass spectrometric characterization of proteins and proteomes, *Nat. Protoc.* 1 (2006) 2856–2860.
- [34] J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification, *Nat. Biotechnol.* 26 (2008) 1367–1372.
- [35] J. Cox, N. Neuhauser, A. Michalski, R.A. Scheltema, J.V. Olsen, M. Mann, Andromeda: a peptide search engine integrated into the MaxQuant environment, *J. Proteome Res.* 10 (2011) 1794–1805.
- [36] J. Cox, M.Y. Hein, C.A. Lubner, I. Paron, N. Nagaraj, M. Mann, Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ, *Mol. Cell. Proteomics* 13 (2014) 2513–2526.
- [37] C. Aslanidis, P.J. de Jong, Ligation-independent cloning of PCR products (LIC-PCR), *Nucleic Acids Res.* 18 (1990) 6069–6074.
- [38] G. Vaaje-Kolstad, D.R. Houston, A.H. Riemen, V.G. Eijsink, D.M. van Aalten, Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21, *J. Biol. Chem.* 280 (2005) 11313–11319.
- [39] A.G. Hamre, S.B. Lorentzen, P. Våljamäe, M. Sørlie, Enzyme processivity changes with the extent of recalcitrant polysaccharide degradation, *FEBS Lett.* 588 (2014) 4620–4624.
- [40] P. Li, A.H. Kwok, J. Jiang, T. Ran, D. Xu, W. Wang, F.C. Leung, Comparative genome analyses of *Serratia marcescens* FS14 reveals its high antagonistic potential, *PLoS One* 10 (2015), e0123061.
- [41] J. Bendtsen, L. Kiemer, A. Fausboll, S. Brunak, Non-classical protein secretion in bacteria, *BMC Microbiol.* 5 (2005) 58.
- [42] P. Siljämäki, P. Varmanen, M. Kankainen, A. Sukura, K. Savijoki, T.A. Nyman, Comparative exoprotein profiling of different *Staphylococcus epidermidis* strains reveals potential link between nonclassical protein export and virulence, *J. Proteome Res.* 13 (2014) 3249–3261.
- [43] E.-Y. Lee, D.-S. Choi, K.-P. Kim, Y.S. Cho, Proteomics in Gram-negative bacterial outer membrane vesicles, *Mass Spectrom. Rev.* 27 (2008) 535–555.
- [44] K.J. McMahon, M.E. Castelli, E.G. Vescovi, M.F. Feldman, Biogenesis of outer membrane vesicles in *Serratia marcescens* is thermoregulated and can be induced by activation of the Rcs phosphorelay system, *J. Bacteriol.* 194 (2012) 3241–3249.
- [45] N. Allocati, M. Masulli, C. Di Ilio, V. De Laurenzi, Die for the community: an overview of programmed cell death in bacteria, *Cell Death Dis.* 6 (2015), e1609.
- [46] I. Tews, R. Vincentelli, C.E. Vorgias, N-Acetylglucosaminidase (chitobiase) from *Serratia marcescens*: gene sequence, and protein production and purification in *Escherichia coli*, *Gene* 170 (1996) 63–67.
- [47] N.D. Rawlings, A.J. Barrett, R. Finn, Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors, *Nucleic Acids Res.* 44 (2016) D343–D350.
- [48] F.X. Gomis-Rüth, Structural aspects of the metzincin clan of metalloendopeptidases, *Mol. Biotechnol.* 24 (2003) 157–202.
- [49] H.-S. Kim, P.N. Golyshin, K.N. Timmis, Characterization and role of a metalloprotease induced by chitin in *Serratia* sp. KCK, *J. Ind. Microbiol. Biotechnol.* 34 (2007) 715–721.
- [50] M.B. Brurberg, I.F. Nes, V.G.H. Eijsink, Comparative studies of chitinases A and B from *Serratia marcescens*, *Microbiology* 142 (1996) 1581–1589.
- [51] L.A. Bøhle, G. Mathiesen, G. Vaaje-Kolstad, V.G. Eijsink, An endo-β-N-acetylglucosaminidase from *Enterococcus faecalis* V583 responsible for the hydrolysis of high-mannose and hybrid-type N-linked glycans, *FEMS Microbiol. Lett.* 325 (2011) 123–129.
- [52] J.A. Vizcaíno, R.G. Côté, A. Csordas, J.A. Dienes, A. Fabregat, J.M. Foster, J. Griss, E. Alpi, M. Birim, J. Contell, The PRoteomics IDentifications (PRIDE) database and associated tools: status in 2013, *Nucleic Acids Res.* 41 (2013) D1063–D1069.

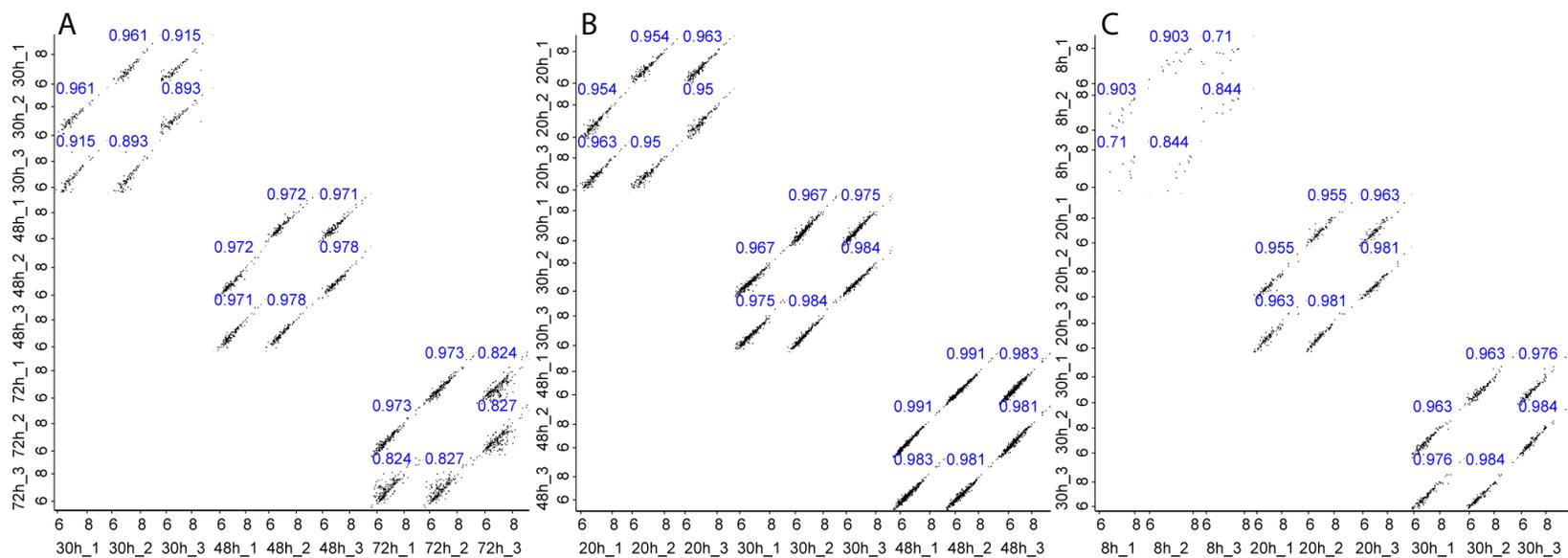


Figure S2. Pearson correlation between replicates for each substrate. A) α -chitin, **B)** β -chitin, **C)** Glucose. Graphs were generated in Perseus, and blue numbers indicate the correlation coefficients. The axes show log₁₀ of LFQ values. The graphs show correlations for three sets of triplicates, each set applying to a different time point, as indicated.

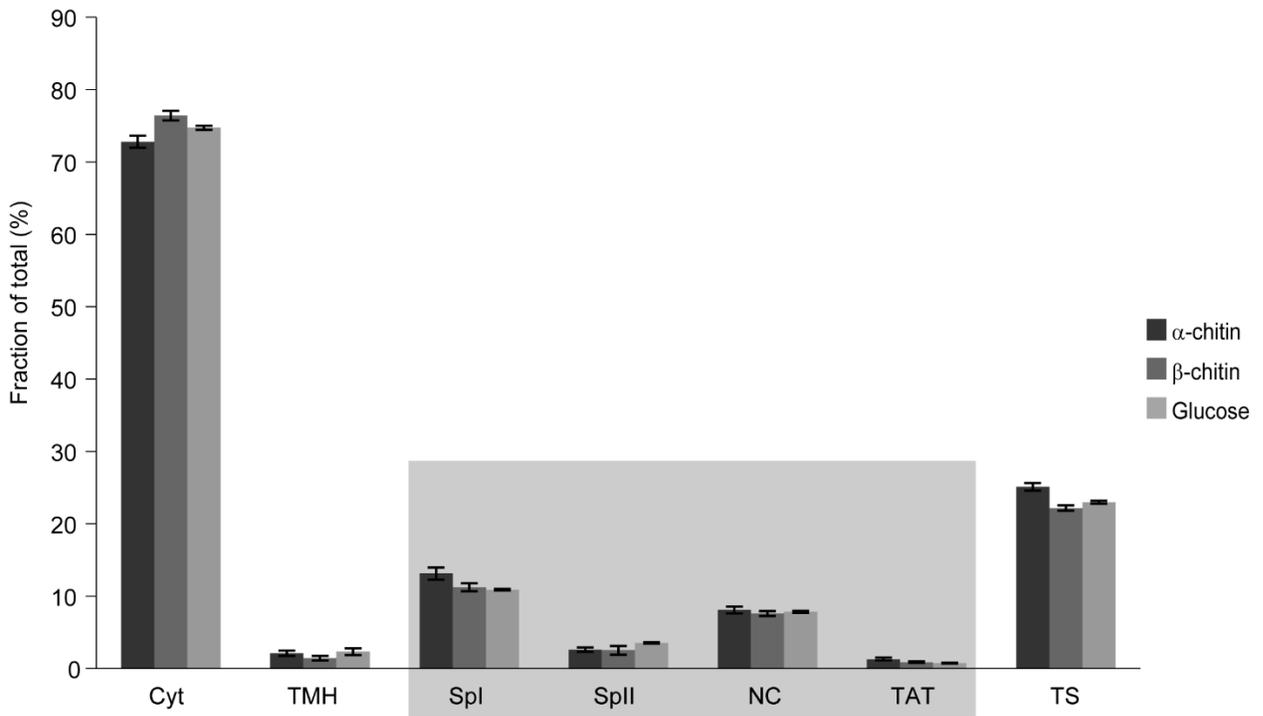


Figure S3. Prediction of subcellular location for proteins detected in cell lysates. The graph shows data for three different carbon sources, α -chitin, β -chitin and glucose. Cyt, cytosolic; TMH, trans-membrane helix; Spl, signal peptidase I cleavage site; SplI, signal peptidase II cleavage site; NC, non-classically secreted; Tat: twin arginine signal peptide; TS, Total secreted (summing up the categories in the grey box). The bars represent the average of three biological replicates, with standard deviation.

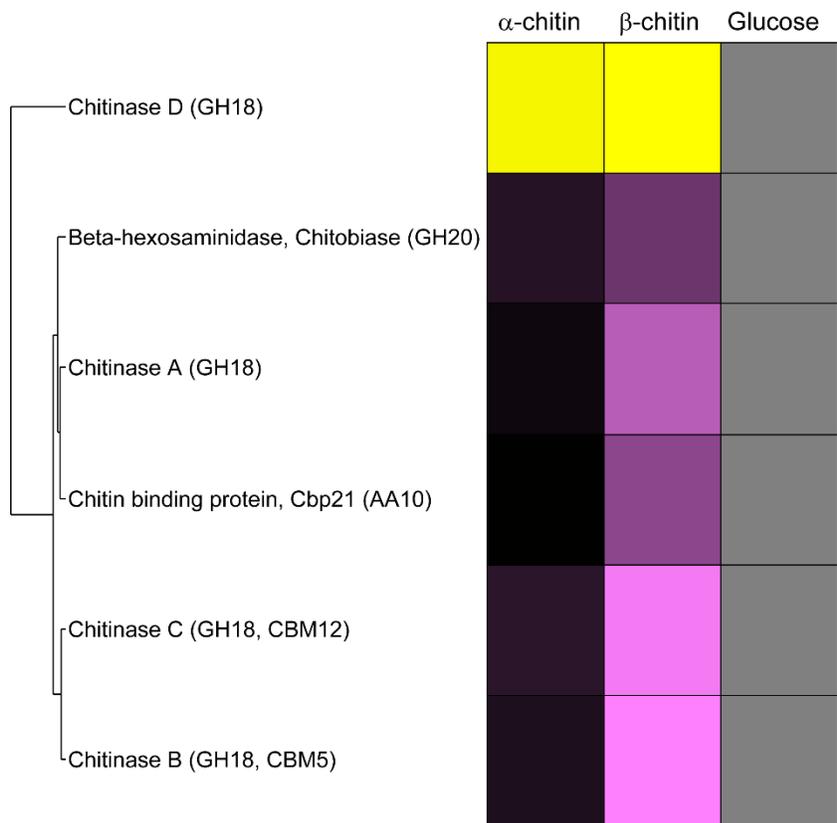


Figure S4. Heat map for chitin-active proteins found in cell lysis samples. The color gradient represents protein abundance expressed as the average of three biological replicates, ranging from \log_{10} LFQ= 6.6 (yellow) to \log_{10} LFQ= 9.1 (pink). Grey fields indicate that the protein was not identified in the sample.

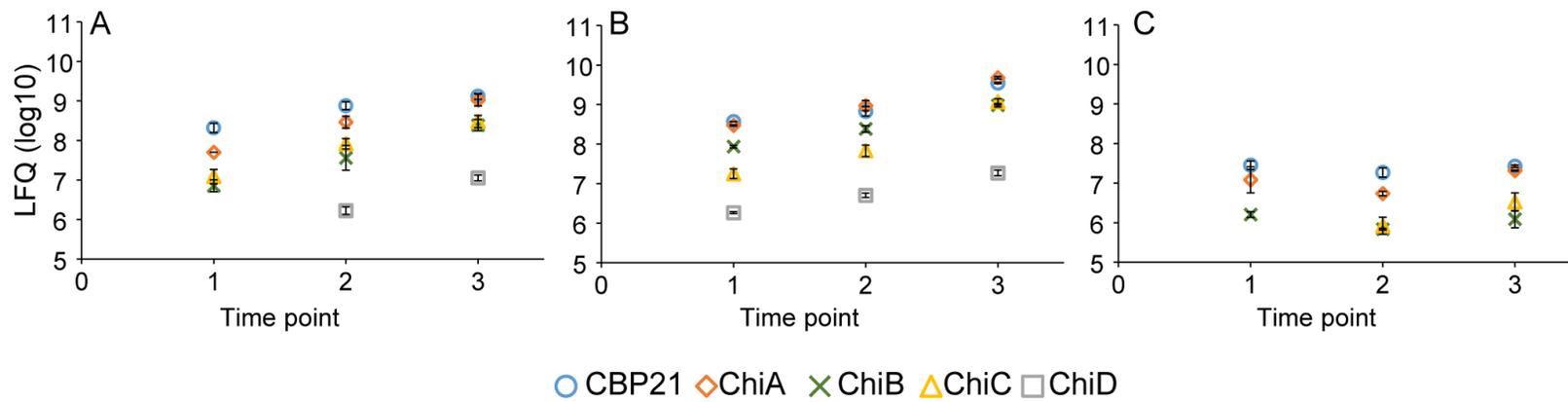


Figure S5. LC-MS LFQ intensities (log₁₀) for chitin-active proteins at the various time points. A) α -chitin, B) β -chitin, C) Glucose. The figure shows how the LFQ intensities vary over time for the different chitin-active proteins found in the secretome, i.e. chitinase A, B, C, D, and CBP21.

Table S1 is available from the publisher's website through the following link:

<http://dx.doi.org/10.1016/j.bbapap.2017.01.007>

References

[1] P. Di Tommaso, S. Moretti, I. Xenarios, M. Orobitg, A. Montanyola, J.M. Chang, J.F. Taly, C. Notredame, T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension, *Nucleic Acids Res.*, 39 (2011) W13-17.